Measurement of the CP Violation Parameter $\sin 2\beta$

by

Kenneth Francis Kelley

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1999

Author . . . .
Department of Physics
September 29, 1998

Certified by . . . .
Paraskevas A. Sphicas
Professor
Thesis Supervisor

Accepted by . . . .
Thomas J. GreyTak
Professor, Associate Department Head for Education

Measurement of the CP Violation Parameter $\sin 2\beta$

by

Kenneth Francis Kelley

Submitted to the Department of Physics
on September 29, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis presents a measurement of the time-dependent asymmetry in the rate of $\overline{B}_d^0$ versus B_d^0 decays to $J/\psi K_S^0$. In the context of the Standard Model this is interpreted as a measurement of the CP violation parameter $\sin(2\beta)$. A total of 198 ± 17 $B_d^0/\overline{B}_d^0$ decays were observed in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV by the CDF detector at the Fermilab Tevatron. The initial B flavor (whether B^0 or $\overline{B}^0$) is determined by a same-side flavor tagging technique. The analysis results in $\sin(2\beta) = 1.8 \pm 1.1\,(\text{stat.}) \pm 0.3\,(\text{syst.})$.

This analysis demonstrates the feasibility of studying CP violation in the B^0–$\overline{B}^0$ system at a hadron collider. By applying the methods used in this analysis, future, higher-statistics experiments should be able to tightly constrain the parameters of the Standard Model.

Portions of this analysis have been submitted for Publication. The article "Measurement of the B_d^0-$\overline{B}_d^0$ flavor oscillation frequency and study of same side flavor tagging of B mesons in $p\overline{p}$ collisions" has been submitted to *Phys. Rev. D*; the preprint for this article is **FERMILAB-Pub-98/188-E**. The article "Measurement of the CP-Violation Parameter $\sin(2\beta)$ in $B_d^0/\overline{B}_d^0 \to J/\psi K_S^0$ Decays" has been submitted to *Phys. Rev. Lett.*; the preprint for this article is **FERMILAB-Pub-98/189-E**.

Thesis Supervisor: Paraskevas A. Sphicas
Title: Professor

Acknowledgements

First of all, I would like to thank my wife, Nancy Chase. Without the strength, love, and support she has given me, I would never have had the endurance to complete this analysis.

I would like to thank my advisor, Paris Sphicas, who provided guidance, wisdom, and support throughout my graduate career. This analysis was originally his conception, and without his suggestions and criticisms, it would never have happened. I would also like to thank Gerry Bauer and Jeff Tseng for their invaluable assistance with performing and writing up the analysis.

I would like to thank Petar Maksimovic for his help. The results of this analysis depend heavily on those of his thesis; if he had not done such a thorough job, this analysis would not have been possible.

In addition, I would like to thank my other MIT collaborators, Troy Daniels, Wasiq Bokhari, Bob Mattingly, Tushar Shah, Dejan Vucinic, Paul Ngan, Steve Pavlon, Sham Sumorok, Steve Tether, and Elizabeth Hafen, who have all provided valuable assistance and support.

I would like to thank the people who built and ran the Fermilab Tevatron and the CDF detector; without them, the data analyzed in this thesis would not exist. In particular, I would like to thank the CDF B physics group for their guidance and suggestions.

I would like to thank my "godparents," Barry Wicklund, Mel Schochet, Larry Nodulman, Donatella Lucchesi, and Colin Gay, whose input has been a vital part of the process of preparing this analysis for publication. I would also like to thank $\mathrm{B_c}$ Todd Huffman, David Stuart, Jonathan Lewis, Joe Kroll, Michael Schmidt, Randall Hans, Steve Pappas, Craig Blocker, Fritz Dejongh, Manfred Paulini, and Slawomir Tkaczyk for their invaluable comments and recommendations.

I would like to thank my friends for being patient with me and helping to keep me sane. Specifically, I'd like to thank Heather Grove, Marian Waldman, Joe Steele, Ed Kuns, Benn Tannenbaum, Anna Breig, Elizabeth Ditchburn, Arthur Lue, Brian

Johnson, Jim Hoff, Maureen Murphy, Noreen Huffman, Kara Hoffman, Carol Hawk, Carol Anway-Wiese, Tim Broberg, Judy Anderson, J. B. Sweeney, Andrew Kursar, Richard Higbee, Carol Hannagan, and Carolingians too numerous to name. I would also like to thank my family for the love and support they have given me all my life.

Finally, I would like to acknowledge my father, Travis Kelley, and my friend, Grover Bethards, who both died during my tenure at MIT. They were both great men, and their influence on my life has been profound.

Contents

List of Figures

List of Tables

Part I

Introduction:
What is CP Violation?
How is it Measured?

Preface

In Lewis Carrol's story *Through the Looking Glass*, a young girl, Alice, climbs through a mirror into the (mirror image) world on the other side. She had expected everything on the other side to be identical to what she saw in the "real" world, but reversed. She was surprised to discover that things were quite different: clocks had actual faces, chess pieces walked about, and flowers talked.

One of the things which makes this story such an excellent piece of fiction is the certainty in the reader's mind that things should be identical (but reversed) in the mirror-image world. When the story was written (1872), it was the unquestioned opinion of nearly everyone on earth that it would be so. Our mirror-image selves would see a mirror-image world. No experiment could tell the one world from the other. This symmetry is called *Parity.*

This century has seen quite a number of experiments whose results have shattered previously unquestioned beliefs. In 1956, C. S. Wu *et al.* [1] performed an experiment which indicated there would be a measurable difference between this world and its mirror-image. They polarized a set of radioactive Cobalt-60 atoms, and observed their decays. The beta-rays produced by these decays tended to be in the same direction as the polarization. In the mirror-image of this reaction, front-and-back are reversed, so the direction of the spin, and hence the polarization would be reversed. But the beta-rays would still be produced in the same direction.

In the mirror image world, the beta rays and polarizations would be opposite, while in the real world, they would be the same. Therefore, parity is not conserved. The effect is much more subtle than in the story by Lewis Carrol, but it exists.

Many scientists were disturbed by this development. If Parity were not conserved, then how many of their other fundamental assumptions about the universe might also be flawed? There was a great deal of relief when it was quickly shown that if, in the mirror-image world, all the matter was replaced by antimatter, then the experiment would agree with the normal-matter experiment in the real world. Anti-Cobalt produces anti-beta rays in the opposite direction to its polarization. The antimatter experiment would look exactly like a mirror-image of the matter experiment. *Charge-Conjugation* is the formal name for replacing all matter with antimatter (and vice-versa). Thus, the combined symmetry Charge-Conjugation–Parity reversal (labeled CP for convenience) appeared to be a conserved symmetry of the universe.

This relief did not last long. Eight years later, in 1964, Christenson, Cronin, Fitch, and Turlay [2] found that CP is violated in the decays of neutral kaons. In short, they produced a stream of particles in one CP state[1] and looked at their decay. Most of the particles decayed to an equivalent CP state, but a few, about 1 in 500, decayed in ways inconsistent with CP conservation.

In the same year, M. Gell-Mann proposed a new theory for elementary particle interactions, one involving three new particles, called *quarks*. In 1970 [3], S. L. Glashow, J. Iliopoulos, and L. Maiani proposed that the addition of a fourth type of quark would explain the lack of certain decays of neutral kaons. In 1973 [4], M. Kobayashi and T. Maskawa proposed that a theory with six types of quarks could accommodate CP violation, while a theory with only four quarks could not. This proposal was made before evidence for a fourth type of quark had even been observed!

Evidence for the existence of a fourth type of quark came in 1975 [5, 6]. In 1977 [7] and 1995 [8], the fifth and sixth quarks were discovered as well. The 6 quark model is called the *Standard Model*, and it has been extremely successful at predicting/explaining all forms of elementary particle interactions. It even provides a possible explanation for the observed CP violation in the decays of neutral kaons. One side-effect of this explanation is the prediction that there should also be CP violation in the decays of other particles, like B^0 mesons.

Neutral kaons are abundantly produced by cosmic rays, and can also be produced in accelerator laboratories. Unfortunately, neutral B mesons are ten-times as massive as neutral kaons, and therefore require much more energy to produce. Also the large

---

[1] Some particles are their own antiparticles, and thus the CP transformation leaves them unchanged. Also, if one has a particle-antiparticle pair, then Charge Conjugation will switch them. Thus it is possible to have a collection of particles which is unchanged by CP, even though the individual particles would be changed by CP.

B meson mass means that most B mesons decay to non CP eigenstates, which would show no asymmetry. This is why nearly 34 years passed with no observation of CP violation in neutral B meson decays.[2]

The Fermilab Tevatron has produced the world's largest collection of neutral B mesons. Roughly half were produced at the CDF interaction region, and the other half at the D0 interaction region. By studying the neutral B mesons produced at CDF, we measure the CP asymmetry in $B^0(\overline{B}^0) \rightarrow J/\psi K^0_S$ decays.

---

[2] The OPAL experiment in Europe has recently made public their results on a study of CP violation in neutral B mesons, using the same method used in this analysis [9].

Chapter 1

The Discovery of CP Violation

This chapter provides a brief historical background for the discovery of CP violation. A more detailed discussion can be found in [10]. Most of the historical background covered by the first chapters of this thesis is taken from this reference. Chapter 2 describes the development of the Standard Model and the currently favored explanation for CP violation.

1.1 Symmetries

Symmetries like invariance under rotations, spatial translations, or temporal translations have been familiar in physics for hundreds of years. In 1918, Emmy Noether [11] proved that for every continuous symmetry there is an associated conserved quantity. The three symmetries above lead to the conservations of angular momentum, linear momentum, and energy, respectively. As F. Gursey [12] notes, "Before Noether's Theorem the principle of conservation of energy was shrouded in mystery."

Symmetries like these are called *continuous*, because they correspond to invariance under a continuous group of transformations. For each symmetry, any element of this group of transformations can be derived from the identity element through a series of applications of an infinitesimal *generator*. For example, the generator of the group of two-dimensional rotations is $\boldsymbol{G_R}$:

$$\boldsymbol{G_R} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{1.1}$$

Any rotation $R(\theta)$ can be expressed in terms of $\boldsymbol{G_R}$:

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \exp\{\theta\,\boldsymbol{G_R}\} = \lim_{N\to\infty}\left(I + \frac{\theta\,\boldsymbol{G_R}}{N}\right)^N . \tag{1.2}$$

Thus, $R(\theta)$ can be reached from the identity by successive applications of an infinitesimal transformation.

Not all symmetries fall into this category. In addition to the continuous symmetries, there are *discrete* symmetries, like invariance under parity (the reversal of the three spatial dimensions). The group of transformations holds only two elements: the parity operator (P) and the identity element (I).[1] There is no infinitesimal operator which can be applied multiple times to yield P.

Discrete symmetries do not lead to conserved quantities in the way that continuous ones do. In 1927, E. Wigner [13] explained that while in quantum mechanics the parity operator leads to the definition of a conserved parity phase, there is no analog of this phase in classical mechanics. In other words the concept of "conservation of parity" is not really meaningful in classical physics. This is in contrast to the conserved quantities from Noether's theorem which apply to classical physics as well as quantum mechanics (which was not available until seven years after Noether's theorem was published).

Wigner went on to remark [14] "Man wird ihn aber nur selten gebrauchen können da er nur zwie Eigenwerte (±) hat und so zu wenig auszusagen vermag," which translates to "Only rarely will one be able to use [parity] since it has only two eigenvalues (±1) and therefore has too little predictive power." About this particular point, Wigner was wrong. Symmetries under discrete operators have proven to be valuable tools in the understanding of elementary particle interactions. In addition to the parity operator, there are two other key operators, time reversal (T) and charge

---

[1] The parity operator is obviously its own inverse, so $P^2 = I$.

conjugation (C). These three are key because the combination of them, CPT, is an inherent symmetry of all quantum field theories, as will be discussed in Section 1.1.4.

1.1.1 Parity (P)

Parity (P), the reversal of the three spatial dimensions, is arguably the most well-known discrete symmetry. The assumption of parity-conservation is so ingrained that few people even realize they make it. Indeed, as stated above, before the invention of quantum mechanics, it was not possible to create a theory which contained interactions which did not conserve parity.

With quantum mechanics, it becomes possible to formulate an interaction which violates parity. Every particle state has a parity phase (*parity* for short) associated with it, which can be either $+1$ or -1. If particle A with parity P_A decays to two particles, B and C, with parities P_B and P_C, then parity conservation requires $P_A = P_B \times P_C \times (-1)^{\ell}$, where ℓ is the quantum number for the orbital angular momentum of the final state. Any interaction where the above correlation does not hold would be said to *violate* parity.

In 1924, Otto Laporte [15] discovered that the energy levels of iron atoms could be separated into two subsets, and that radiative decays could only occur between levels in different subsets. Unfortunately, quantum mechanics was not available until 1925, so no (accurate) explanation of the observed effect was forthcoming. In 1927, E. Wigner [13] explained that Laporte's results could be described by parity conservation, if the two subsets correspond to states with different parities. Because a photon has negative parity, emission of a photon requires transition to a state of opposite parity. Therefore decays between the two subsets would be allowed, but decays within the subsets would not. Wigner's explanation of Laporte's results was the first use of the assumption of the invariance of a discrete symmetry to explain an observed phenomenon.

1.1.2 Time Reversal (T)

Another commonly assumed symmetry is that of *time reversal* (T). The laws of classical mechanics are known to be invariant under time reversal. In quantum mechanics, the operator T changes each state into its complex conjugate, so a complex Hamiltonian would yield interactions which violate T.

In addition to being the first person to propose conservation under parity as an explanation of an observed phenomenon, Wigner was also the first to propose T conservation as an explanation for an observed effect. In 1932 [16], he showed how T invariance predicts Kramers' degeneracy theorem, which states [17] "the energy eigenstates of an odd number of spin-$\frac{1}{2}$ particles are at least doubly degenerate in the absence of an external magnetic field."

1.1.3 Charge Conjugation (C)

In 1927, Dirac formulated an equation for describing the time-evolution of a free electron. One consequence of this equation was the prediction that, in addition to the electron, there should be another particle, one with mass equal to the electron mass and charge opposite that of the electron. This particle was named the *positron*, and its prediction was considered a problem with the theory, as no positrons had been observed.

In 1933 [18], Anderson discovered the positron, vindicating the Dirac theory. The existence of antimatter gave rise to a new operator: the *charge conjugation* operator, which replaces all particles with their antiparticles. This operator, like P and T, is its own inverse and has an associated discrete symmetry.

C Invariance and Neutral Kaons

One of the most interesting applications of the assumption of C invariance came in 1955 by M. Gell-Mann and A. Pais [19]. They showed that a photon must have $C = -1$, while a π^0 must have $C = +1$. In addition, a state with a π^+ and a π^- would have $C = (-1)^\ell$, where ℓ is the relative orbital angular momentum.

They then proposed that the K^0 particle (known then as θ^0) should have an antiparticle, $\overline{K}^0$, distinct from itself. Unlike the photon and the π^0, the K^0 and $\overline{K}^0$ would not be eigenstates of C: the C operator would interchange them. However, mixtures of the two states would be eigenstates of C:

$$|K_1^0\rangle = \frac{|K^0\rangle + |\overline{K}^0\rangle}{\sqrt{2}} \tag{1.3}$$

$$|K_2^0\rangle = \frac{|K^0\rangle - |\overline{K}^0\rangle}{\sqrt{2}}. \tag{1.4}$$

The K_1^0 would have $C = +1$ and the K_2^0 would have $C = -1$.

These states would not be interesting if it were not possible for a K^0 state to change into a $\overline{K}^0$ state, as there would be no way for a pure K^0 state to turn into one of the above mixed states. However, the K^0 is known to decay weakly to $\pi^+\pi^-$, so C invariance would require that the $\overline{K}^0$ also decay to $\pi^+\pi^-$. Thus, through the double weak interaction, $K^0 \leftrightarrow \pi^+\pi^- \leftrightarrow \overline{K}^0$, the two states can mix. The mixed states (K_1^0 and K_2^0) can arise from pure states (K^0 and $\overline{K}^0$).

Gell-Mann and Pais then showed[2] that decay modes of the K_1^0 and K_2^0 to a collection of pions and photons have no overlap. The $\pi^+\pi^-$ state with $\ell = 0$ has $CP = +1$, and the $\pi^+\pi^-\pi^0$ state has $CP = -1$. Whatever the K^0 spin, only one of the two states K_1^0 or K_2^0 could decay to two pions, and only the other state could decay to three pions.

If the decay modes do not overlap, then the lifetimes of the K_1^0 and K_2^0 states are independent. Gell-Mann and Pais continue with:

> "While we have seen that the θ_1^0 and θ_2^0 may each be assigned a lifetime, this is evidently not true of the θ^0 or $\overline{\theta}^0$. Since we should properly reserve the word 'particle' for an object with a unique lifetime, it is the θ_1^0 and θ_2^0 quanta that are the true 'particles'. The θ^0 and $\overline{\theta}^0$ must, strictly speaking, be considered as 'particle mixtures.' " [19]

---

[2] To make this derivation, they also assumed P invariance. Only the weaker assumption that combined symmetry CP was conserved was necessary; assuming C and P were separately conserved was not required. This was not important at the time, but became very important once C and P were found to be violated in such a way that CP appeared conserved.

The notion that the particles produced in strong interactions are actually "particle mixtures." was certainly a radical concept. Gell-Mann and Pais made three predictions which could be tested experimentally. They first predicted "that *not more than half of all* θ^0*'s* can undergo the familiar decay to two pions." They also predicted that, if the spin and parity of the K^0 are even, then the decay $K^0 \rightarrow \pi^0\pi^0$ should be allowed as well. Thirdly, they predicted there should be another state with lifetime $\tau' \gg \tau$ (where τ is the lifetime of the observed $K^0 \rightarrow \pi^+\pi^-$), which does not decay to two pions.

Evidence of the long-lived K^0 decay was found in 1956 [20]. In 1957, it was found [21] that not more than half of neutral kaons decay to $\pi^+\pi^-$, and in 1958 [22], the $\pi^0\pi^0$ decay of the short-lived K^0 state was observed. This last observation also indicated that the K_1^0 is the short-lived state and the K_2^0 is the long-lived one.

1.1.4 Combined CPT Symmetry

The combination of the three symmetries CPT holds a special status. In 1957, R. Jost and F. J. Dyson [23, 24] presented proofs which showed that CPT invariance is a theorem, one which is derivable from the axioms of quantum field theory.[3] Thus, while it was possible to create a quantum field theory in which any of the three symmetries, C, P, or T, was violated, the combined symmetry CPT must hold. This did not prove that CPT must be inviolate in the real world, only that, if CPT were violated, no quantum field theory could explain it.

1.2 Violations of Symmetries

In 1956, particle physicists were plagued by the "τ-θ puzzle." Two particles, the τ^+ and θ^+, appeared to have the same mass and lifetime, but opposite parities. The τ^+ decayed to $\pi^+\pi^+\pi^-$, a $P = -1$ state, and the θ^+ decayed to $\pi^+\pi^0$, a $P = +1$ state.

---

[3] In 1954 [25], Lüders and Pauli showed that, assuming P invariance, C conservation was equivalent to T conservation. Jost and Dyson presented the first rigorous axiomatic proofs of CPT invariance for quantum field theories.

It would have been a nice simplification to identify the two as a single particle, but the assumption of parity conservation led to the conclusion that they could not be.

T. D. Lee and C. N. Yang proposed [26] that the assumption of parity conservation might not be appropriate. They pointed out that while there was copious evidence that parity was conserved in strong and electromagnetic interactions, parity conservation had never been demonstrated in weak interactions. They proposed two experiments that could be performed to study parity violation in weak interactions.

1.2.1 Violation of P

The first experiment proposed by Lee and Yang was to study the angular dependence of β rays from the decays of polarized Co^{60} nuclei. The distribution of β rays in θ, where θ is the angle between the direction of the β ray and the direction of polarization, should follow:

$$I(\theta) \quad \propto \quad (1+\alpha\cos\theta)\sin\theta, \tag{1.5}$$

where $\alpha \neq 0$ would indicate P violation.

C. S. Wu *et al.* [1] quickly performed the experiment. Due to systematic uncertainties, determination of the exact evaluation of α was difficult, but they were able to exclude values of α below 0.4. This result was clear evidence of parity violation in weak interactions.

1.2.2 Violation of C

The other experiment proposed by Lee and Yang was to study the decay $\pi^+ \to \mu^+\nu$. If parity were conserved, then the muons from this decay would not exhibit any polarization along their direction of motion. This muon polarization could be studied by looking at their subsequent decays to $e\nu\nu$. The distribution of the angle (θ) between the muon direction of motion and the direction of the electron momentum (in the muon rest frame) should be symmetric about $\theta = \pi/2$.

Two groups [27, 28] rushed to perform this experiment, publishing shortly after Wu *et al.*. They each found that the muons from $\pi^+ \to \mu^+\nu$ were highly polarized, thus confirming parity violation. They also observed that in $\pi^- \to \mu^-\overline{\nu}$, the polarization was opposite to that for the π^+ decay, demonstrating C violation.

1.2.3 Violation of CP

While these experiments demonstrated that neither C and P is conserved, the experiments did not indicate that the combined symmetry, CP, was violated. In contrast, the observed violations of C and of P cancel exactly in the combination CP. The obvious place to search for CP violation was in the decay of the K_2^0, the long-lived, CP-negative K^0 state. *Any* 2–pion decays of the K_2^0 would indicate CP violation.

Lots of experiments [29, 30] sought, but did not find, evidence of CP violation in weak interactions. They did exclude branching fractions for $K_2^0 \to \pi^+\pi^-$ higher than 1/300, though. In 1964, Christenson, Cronin, Fitch, and Turlay published the results of their search [2]. They found 45 $K_2^0 \to \pi^+\pi^-$ decays out of roughly 22,000 K_2^0 mesons produced, for a branching fraction of $(2.0 \pm 0.4) \times 10^{-3}$.

The first implication of this was that the CP eigenstates, K_1^0 and K_2^0, are not strictly "particles," either. The real particles are the K_S^0 (K-short) and K_L^0 (K-long):

$$|K_S^0\rangle = \frac{p_K|K^0\rangle - q_K|\overline{K}^0\rangle}{\sqrt{|p_K|^2 + |q_K|^2}} = \frac{|K_1^0\rangle + \epsilon|K_2^0\rangle}{\sqrt{1+|\epsilon|^2}} \tag{1.6}$$

$$|K_L^0\rangle = \frac{p_K|K^0\rangle + q_K|\overline{K}^0\rangle}{\sqrt{|p_K|^2 + |q_K|^2}} = \frac{\epsilon|K_1^0\rangle + |K_2^0\rangle}{\sqrt{1+|\epsilon|^2}} \tag{1.7}$$

where $|\epsilon| \simeq 2.3 \times 10^{-3}$ was also measured in [2].

The next implication was that an explanation for CP violation was needed. Three types of explanations were available:

1. Weak interactions violate CP.
2. Either strong or electromagnetic interactions violate CP.
3. There exists a new *super-weak* interaction which violates CP and has coupling strength $\sim 10^{-9}$ times weaker than that of weak interactions.

Option 2) was disfavored, as both C and P appeared to be very well conserved in both strong and electromagnetic interactions. Option 3) was disfavored by some, as it requires adding in a new field to the theory, a field which would be very difficult to study since its only effects are seen in neutral meson systems. The currently favored explanation is option 1), as introduced in the Standard Model. Chapter 2 contains a brief description of interactions in the Standard Model and how it can provide a mechanism for the explanation of CP violation.

Chapter 2

The Standard Model

The *Standard Model* of particle physics consists of two quantum field theories (Quantum Chromodynamics and the Electroweak Theory) which describe the strong, electromagnetic and weak interactions. It contains several free *parameters* which describe the strengths of the three interactions. These parameters are not predicted by the theory and must therefore be determined experimentally. Once these parameters are measured, the Standard Model can be used to predict the outcome of any experiment involving any combination of the above three interactions. In this sense, the Standard Model has been extremely successful in describing essentially all experimental results since its inception.

The violation of CP observed in the neutral K^0–$\overline{K}^0$ system can be accommodated in the context of the Standard Model, via the introduction of an additional free parameter, η. With this parameter constrained, the Standard Model predicts that there should also be CP violation in neutral B mesons; the magnitude of CP violation in neutral B mesons is also defined by η. Thus, measurements of CP violation in neutral B mesons provide powerful tests of the Standard Model. Results inconsistent with the predictions would indicate some fundamental weakness in the Standard Model formalism.

This chapter provides a brief description of the Standard Model. It explains how CP violation arises in the context of the Standard Model, and what predictions the Standard Model makes for CP violation in neutral B mesons.

2.1 Interactions in the Standard Model

The elementary particles of the Standard Model [31] are leptons, quarks, and gauge bosons; there are six "flavors" of lepton, six flavors of quark, and four flavors of gauge boson. Table 2.1 lists them, their masses, and their electric charges.

Leptons		
Particle	Mass (MeV/c^2)	Charge
e	0.511	-1
ν_e	$< 15 \times 10^{-6}$	0
μ	106	-1
ν_μ	< 0.17	0
τ	1777	-1
ν_τ	< 18	0

Quarks		
Particle	Mass (MeV/c^2)	Charge
d	3–9	$-1/3$
u	1.5–5	2/3
s	60–170	$-1/3$
c	1100–1400	2/3
b	4100–4400	$-1/3$
t	174×10^3	2/3

Gauge Bosons		
Particle	Mass (MeV/c^2)	Charge
Photon (γ)	$< 2 \times 10^{-22}$	0
Gluon (g)	0	0
$W^\pm$	80.4×10^3	1
Z^0	91.2×10^3	0

Table 2.1: Elementary particles of the Standard Model. Photon and neutrino mass numbers are 95% CL upper limits. The gluon mass is the theoretical prediction; masses as large as a few MeV/c^2 are not excluded by experiment. Each of these particles has an antiparticle with opposite charge, except for the γ, Z, and g, which are their own antiparticles. All of these quantities are from Ref. [32].

The leptons and quarks all have spin 1/2, and are called *fermions*. The gauge bosons all have spin 1. Standard Model interactions connect two fermions with one boson, three bosons with one another, or four bosons with one another, as shown in figure 2-1.

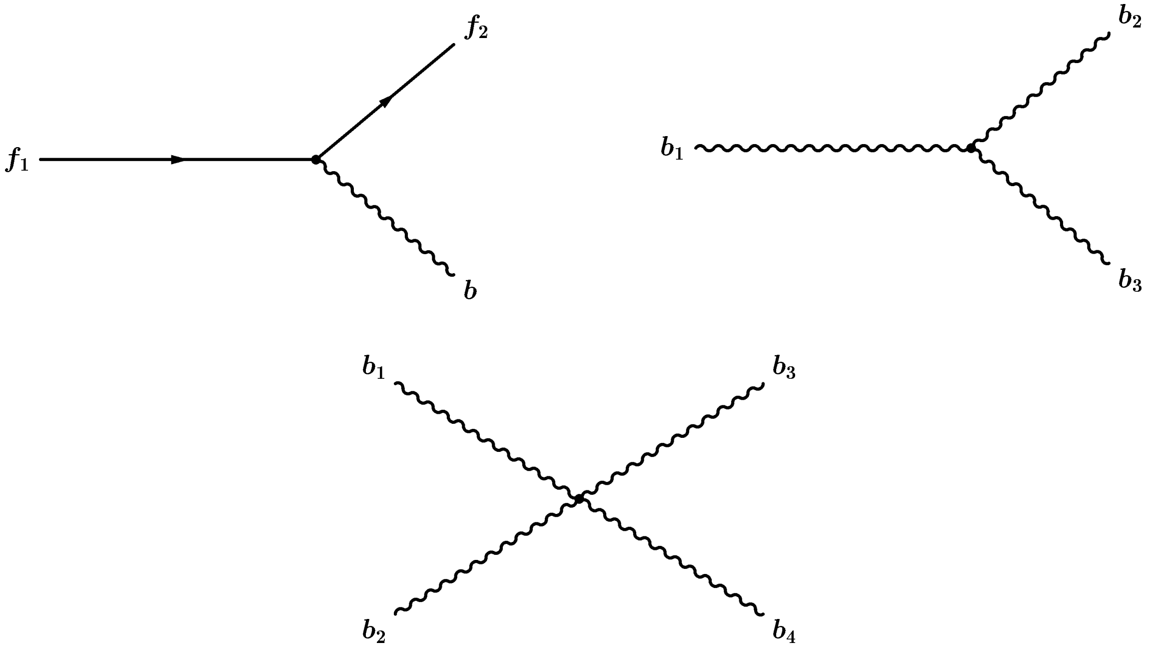

Figure 2-1: Types of interactions in the Standard Model; fermion-fermion-boson, 3-boson, and 4-boson. Lines labeled with an "f" indicate fermions, and those labeled with a "b" indicate bosons.

The coupling strength of each of these interactions is provided by the Standard Model Lagrangian. Many combinations, such as $eu\gamma$ are *forbidden* and have an amplitude of zero. The interactions that are allowed are separated into three categories, based on the type of boson involved.

2.1.1 Strong Interactions

Interactions involving gluons are called *strong interactions.* Historically, this name arose because it is the strong interaction that holds atomic nuclei together, overcoming the electromagnetic repulsion between protons.

The only allowed interactions involving gluons are those involving quarks or other gluons. In addition, in the interaction between a gluon and two quarks, the incoming and outgoing quarks must have the same flavor. The strengths of all strong interactions are described by a single Standard Model parameter: α_s.

2.1.2 Electromagnetic Interactions

Interactions involving photons are called *electromagnetic*, as they give rise to electromagnetic phenomena. The amplitudes of these interactions are all defined by the single parameter:

$$\alpha = \frac{e^2}{\hbar c} \tag{2.1}$$

where e is the magnitude of the charge of the electron. The strength of the interaction is proportional to the charge of the fermion or boson coupled to (see Table 2.1). As with strong interactions, the only $ff\gamma$ interactions allowed are those where the incoming and outgoing fermions have the same flavor.

There are also electromagnetic interactions coupling one or two photons to a W^+W^- pair. The strengths of these interactions are also defined by α.

2.1.3 Weak Interactions

Weak Interactions were first postulated to exist as an explanation of nuclear β decay. The coupling had to be weak, to account for the long observed half-lives of radioactive particles. As will be described in Section 2.2, the decays of charged pions, muons, and several other particles have also been attributed to weak interactions. Two parameters are required to describe the strengths of weak interactions, α (same as for electromagnetic) and $\sin\theta_W$.

The couplings for the three-boson weak interaction ($Z^0W^+W^-$) is $\alpha\cot\theta_W$. The couplings for the four-boson weak interactions are: $\alpha^2\cot^2\theta_W$ for $W^+W^-Z^0Z^0$, $\alpha^2/\sin^2\theta_W$ for $W^+W^-W^+W^-$, and $\alpha^2\cot\theta_W$ for $W^+W^-Z^0\gamma$.

Neutral Weak Interactions

Fermion-fermion-boson interactions involving Z^0 bosons are similar to strong and electromagnetic interactions in that they are only allowed when the incoming and outgoing fermions have the same flavor. The corresponding "charge" which defines the coupling strength is:

$$Q_Z = \frac{e}{\sin\theta_W \cos\theta_W}\left(T - Q\sin^2\theta_W\right) \tag{2.2}$$

where Q is the electric charge (in units of e); T is zero for all right-handed particles,[1] $T = +1/2$ for left-handed $\nu_e, \nu_\mu, \nu_\tau, u, c, t$, and $T = -1/2$ for left-handed e, μ, τ, d, s, b. The coupling for anti-particles is the same as that for the equivalent particle with the opposite helicity. Thus, $Q_Z(u_L) = Q_Z(\overline{u}_R)$.

Charged Weak Interactions

Unlike the other bosons, W bosons carry electric charge. Thus, the two fermions in a ffW interaction necessarily have different flavors, or charge could not be conserved. Table 2.1 separates leptons and quarks into "generations": $\{e, \nu_e\}$, $\{\mu, \nu_\mu\}$, $\{\tau, \nu_\tau\}$, $\{d, u\}$, $\{s, c\}$, and $\{b, t\}$. To first order, ffW interactions occur only when the two fermions are different members of the same generation.[2] The charges corresponding to the coupling strengths of these interactions are:

$$Q_W = \frac{e\,(1-\gamma_5)}{\sin\theta_W}, \tag{2.3}$$

where $\gamma_5 = +1$ for right-handed particles and left-handed antiparticles, and $\gamma_5 = -1$ for left-handed particles and right-handed antiparticles.[3]

Parity violation is included directly in the Standard Model via the $(1-\gamma_5)$ coupling factor. Right-handed particles and left-handed antiparticles have $\gamma_5 = +1$, so they have $Q_W = 0$, and they do not participate in charged weak interactions. This discrimination between right and left-handed particles is the effect at the heart of the result observed by C.S. Wu [1], mentioned in the previous chapter. Since the

---

[1] "Right-handed" particles are those with "positive helicity", or, in other words, those whose spins are parallel to their momenta. "Left-handed" particles are those whose spins are antiparallel to their momenta.

[2] In addition, there are ffW interactions with two quarks from separate generations. These will be discussed in Section 2.3.

[3] Technically, γ_5 is a matrix, and $(1-\gamma_5)$ is an operator which is a projection onto left-handed particles and right-handed antiparticles.

couplings are opposite for antiparticles, these couplings do not violate the combined symmetry CP.

2.1.4 Summary

The Standard Model needs only 3 parameters (α_s, α, and $\sin\theta_W$) to define the strengths of all these interactions, to first order. However, there is an additional complication which has not been discussed. The W boson coupling is not strictly restricted to the pairs listed above; interactions involving quarks from different generations are also allowed. Description of these additional interactions requires 4 more parameters, one of which is η, the parameter which defines the magnitude of CP violation in neutral kaons and B mesons. The next section provides a brief history of the development of the Lagrangian of the weak interaction, showing how these *intergenerational* couplings have been added. The section after describes the modern weak Lagrangian, and how CP violation arises within it.

2.2 Historical Development of the Weak Lagrangian

Weak interactions were first postulated to be the source of nuclear β decay. When pions were discovered to decay to muons and muons were discovered to decay to electrons, these decays were also attributed to weak interactions. The lifetimes of charged pions and muons were much too long for the decays to be from strong or electromagnetic interactions.

Today's understanding of weak interactions is formulated in the Standard Model weak Lagrangian, which involves quarks and leptons. Thus, the development of this Lagrangian was closely tied to the development of the theory of quarks.

2.2.1 Quark Theory

In 1964 [33], Gell-Mann proposed that hadrons, instead of being elementary particles, were composite particles, made up of *quarks*. The proposal included three *flavors* of quarks: *up* (u), *down* (d), and *strange* (s). The u quark would have electric charge $Q = +2/3$, and the d and s quarks would have $Q = -1/3$. The proton would contain uud and the neutron would contain udd. The π^+ would contain $u\overline{d}$, the π^-: $d\overline{u}$, the K^0: $d\overline{s}$, and the $\overline{K}^0$: $s\overline{d}$. Other assignments can be found in [32].

Citing a paper [34] written a year earlier by Cabibbo, Gell-Mann proposed that weak interactions would couple the u quark to a combined state, the $|d'\rangle = \cos\theta|d\rangle + \sin\theta|s\rangle$, where $\theta \approx 0.26$ was measured by Cabibbo. With this definition, all (known) weak interactions could be described by a single coupling constant. Since the u quark couples to this combined state, interactions involving a $s \to u$ transition would be suppressed by a factor of $\tan\theta$ relative to those involving a $d \to u$ transition. This was exactly the result observed by Cabibbo.

2.2.2 The GIM Mechanism

In 1970 [3], S. L. Glashow, J. Iliopoulos, and L. Maiani proposed that a fourth flavor of quark exists, which they labeled *charm* (c). They did this to explain why the K^0 meson was not observed to decay to $\mu^+\mu^-$. Figure 2-2 shows a diagram of how a K^0 decays to $\mu^+\mu^-$.

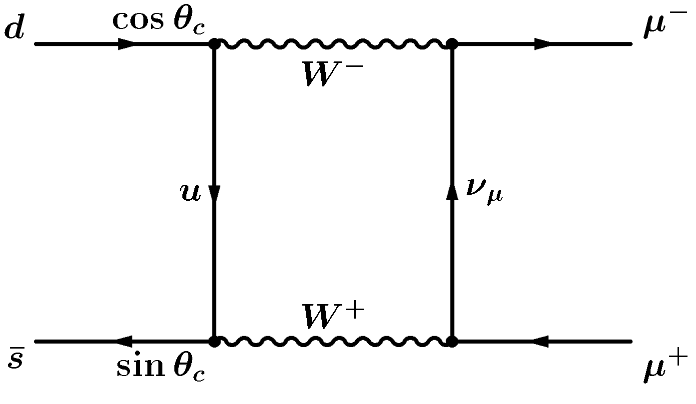

Figure 2-2: One path for the decay $K^0 \to \mu^+\mu^-$.

They proposed that the c quark would have a weak interaction partner $|s'\rangle = -\sin\theta|d\rangle + \cos\theta|s\rangle$. This would add another decay path, as shown in figure 2-3. The first diagram would have an amplitude proportional to $\cos\theta\sin\theta$ and the second would have amplitude $-\cos\theta \sin\theta$, so they would cancel exactly. Only the difference between the c and u masses would allow the decay to occur.

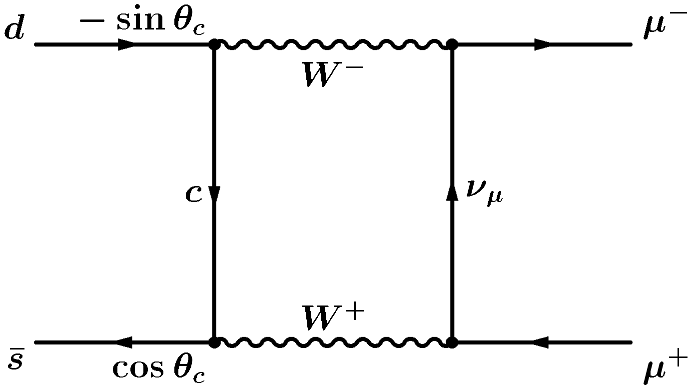

Figure 2-3: Another path for the decay $K^0 \to \mu^+\mu^-$.

With this addition, the d' and s' states are simply mixtures of the d and s states. The "mixing matrix" would be:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{2.4}$$

As a change of basis, V would necessarily be unitary (*i.e.* $V^\dagger = V^{-1}$), which confirms that the $|s'\rangle$ state must be $-\sin\theta|d\rangle + \cos\theta|s\rangle$.

2.2.3 The Kobayashi-Maskawa Mechanism

In 1973 [4], M. Kobayashi and T. Maskawa proposed extending the quark model to six quarks, instead of four. The reason for this extension was that, with a six-quark model, CP violation could be explained within the Standard Model, while a four-quark model would require introduction of another interaction.

They explained that, while an $n\times n$ mixing matrix contains n^2 real numbers, $2n-1$ of them can be removed by redefining the quark phases.[4] Thus, for a 2×2 matrix,

---

[4] If u is a weak eigenstate, then so obviously must be $e^{i\phi}u$ for any real ϕ. Redefining the u quark to have a different complex phase would not affect anything observable, but it would change the representation of the mixing matrix. Thus, one can remove one of the free parameters of the matrix by redefining the phase of the u state. The same can be dome for each of the quark phases, but

only 1 free parameter would remain, and that would be the real angle θ. Therefore, a mixing matrix in two dimensions is always expressible as a real matrix. A 3×3 mixing matrix would have 4 parameters, but a 3×3 orthogonal matrix contains only 3 angles. Thus, a 3×3 mixing matrix must have a complex phase which cannot be redefined away (though it could still happen to be zero).

If the mixing matrix is complex, then the Standard Model Hamiltonian would contain complex terms. As mentioned in Section 1.1.2, a complex Hamiltonian necessarily violates T, and thus (assuming CPT invariance, which is necessary for quantum field theories), violates CP as well. Thus, a six-quark model provides a natural mechanism for including CP violation, while a four-quark model does not.

2.2.4 The Discoveries of c, b, and t

These last two proposals were given very little attention until the experimental discoveries of the relevant quarks. In 1975 [5, 6], two independent research teams discovered a new particle, the J/ψ, which was quickly identified as a $c\bar{c}$ meson. In 1977 [7], the Υ was discovered and identified as a $b\bar{b}$ meson. And, in 1995 [8], the top quark was discovered, completing the picture.

---

shifting all their phases by the same amount does not change the matrix. Therefore, it is possible to redefine quark phases and remove $2n - 1$ free parameters from a mixing matrix.

2.3 The Modern Weak Lagrangian

The 3×3 mixing matrix is referred to as the Cabibbo–Kobayashi–Maskawa (CKM) matrix.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.5}$$

As described above, this matrix contains three angles and one complex phase. One commonly used parameterization of this matrix is [32]:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2.6}$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$, $0 \leq \theta_{ij} \leq \pi/2$, and $0 \leq \delta < 2\pi$. The three angles are θ_{12}, θ_{13}, and θ_{23}, and the complex phase is δ. In this parameterization, V_{ud}, V_{us}, V_{cb}, and V_{tb} are all real and have no complex phase.

Experimental measurements indicate that these angles are small, and that $\mathcal{O}(\theta_{12}) \sim \mathcal{O}(\theta_{23}/\theta_{12}) \sim \mathcal{O}(\theta_{13}/\theta_{23})$. Wolfenstein[35] reparameterized these angles as follows: $\sin\theta_{12} = \lambda$, $\sin\theta_{23} = A\lambda^2$, and $\sin\theta_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta)$. This parameterization has A, ρ, & η all of order unity, while λ is small ($\lambda = 0.22$). Using this parameterization and ignoring all terms of $\mathcal{O}(\lambda^4)$ or higher, he arrived at:

$$V \simeq \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{2.7}$$

The complex phases of the elements V_{ub} and V_{td} are $\gamma \equiv -\arg(V_{ub}) = \tan^{-1}(\eta/\rho)$, and $\beta \equiv -\arg(V_{td}) = \tan^{-1}(\eta/1 - \rho)$. The other complex phases are all small: $\arg(V_{cd}) \simeq A^2\lambda^4\eta$, $\arg(V_{cs}) \simeq -A^2\lambda^6\eta$, and $\arg(V_{ts}) \simeq \lambda^2\eta$. All these phases are described by the single parameter η, which will be shown to be the one which determines the magnitude of CP violation in neutral kaons and B mesons.

The following subsections discuss some of the effects of the CKM matrix. The first subsection discusses B^0 mixing, a consequence of the nondiagonal couplings in the CKM matrix. The second subsection discusses implications of the Unitarity constraint on the CKM matrix, and how this constraint can be tested. The third subsection discusses how CP violation in neutral B mesons is predicted if $\eta \neq 0$.

2.3.1 B^0 Mixing

The decays of K_S^0 mesons are predominantly to $\pi\pi$, a CP eigenstate, *i.e.* a final state to which decays from K_L^0 are suppressed. The other decay modes of K^0 mesons all have much smaller partial widths than the $\pi\pi$ decay, so the two physical eigenstates, K_S^0 and K_L^0, have substantially different widths and lifetimes.

The decays of B^0 mesons, in contrast, are mostly not to CP eigenstates, so the widths of the two physical eigenstates are very similar ($\Delta\Gamma/\Gamma < 10^{-2}$). In the relevant literature, the approximation $\Delta\Gamma/\Gamma = 0$ is very common, and we adopt it here, also.

As with the neutral kaon system, the two physical eigenstates of neutral B mesons, B_H^0 (heavy) and B_L^0 (light), are linear combinations of the strong eigenstates, B^0 and $\overline{B}^0$:

$$
\begin{aligned}
|B_H^0\rangle &= p_B|B^0\rangle - q_B|\overline{B}^0\rangle && (2.8)\\
|B_L^0\rangle &= p_B|B^0\rangle + q_B|\overline{B}^0\rangle && (2.9)
\end{aligned}
$$

where p_B and q_B are normalized so that $\sqrt{|p_B|^2 + |q_B|^2} = 1$.

Applying the time-propagation operator to the mass eigenstates gives:

$$
\begin{aligned}
|B_H^0(t)\rangle &= e^{-(\Gamma_H/2+iM_H)t}|B_H^0\rangle = e^{-\Gamma t/2}e^{-i(M+\Delta m/2)t}|B_H^0\rangle && (2.10)\\
|B_L^0(t)\rangle &= e^{-(\Gamma_L/2+iM_L)t}|B_L^0\rangle = e^{-\Gamma t/2}e^{-i(M-\Delta m/2)t}|B_L^0\rangle && (2.11)
\end{aligned}
$$

where M_H and M_L are the masses of the high and low mass eigenstates, M is their average, and $\Delta m \equiv M_H - M_L$ is their difference. $\Delta\Gamma/\Gamma$ is approximated to be zero, so $\Gamma_H = \Gamma_L = \Gamma$. As t increases, the amplitudes of these states decrease (from decays) and the phases change. But, a pure $|B_H^0\rangle$ state remains a pure $|B_H^0\rangle$ state, and a pure $|B_L^0\rangle$ state remains a pure $|B_L^0\rangle$ state.

In contrast, the $|B^0\rangle$ state mixes with the $|\overline{B}^0\rangle$ state:

$$\begin{aligned}
|B^0(t)\rangle &= U(t)\left[\frac{|B_H^0\rangle + |B_L^0\rangle}{2p_B}\right] \\
&= e^{-\Gamma t/2}\left[\frac{e^{-iM_H t}|B_H^0\rangle + e^{-iM_L t}|B_L^0\rangle}{2p_B}\right] \\
&= e^{-\Gamma t/2}e^{-iMt}\left[\frac{e^{-i\Delta m t/2}|B_H^0\rangle + e^{i\Delta m t/2}|B_L^0\rangle}{2p_B}\right] \\
&= e^{-\Gamma t/2}e^{-iMt}\left[\left(\frac{e^{-i\Delta m t/2} + e^{i\Delta m t/2}}{2}\right)|B^0\rangle + \left(\frac{e^{i\Delta m t/2} + e^{-i\Delta m t/2}}{2}\right)\left(\frac{q_B}{p_B}\right)|\overline{B}^0\rangle\right] \\
&= e^{-\Gamma t/2}e^{-iMt}\left[\cos(\Delta m t/2)\,|B^0\rangle + i\,(q_B/p_B)\sin(\Delta m t/2)\,|\overline{B}^0\rangle\right] \qquad (2.12)
\end{aligned}$$

where $U(t)$ is the time-projection operator.

Thus, the probabilities of a particle which was produced as a B^0 being subsequently observed at time t as a B^0 or as a $\overline{B}^0$ are:

$$P(B^0) = \left|\langle B^0|U(t)|B^0\rangle\right|^2 = e^{-\Gamma t}\left(\frac{1+\cos(\Delta m t)}{2}\right) \qquad (2.13)$$

$$P(\overline{B}^0) = \left|\langle \overline{B}^0|U(t)|B^0\rangle\right|^2 = e^{-\Gamma t}\left(\frac{1-\cos(\Delta m t)}{2}\right) \qquad (2.14)$$

The diagrams in figure 2-4 display the main interactions which lead to B^0 mixing. Evaluation of these diagrams leads to the following formula for the mixing frequency [36]:

$$\Delta m = \frac{G_F^2}{6\pi^2} m_B m_t^2 F\left(\frac{m_t^2}{m_W^2}\right)\eta_{QCD} B_{B^0} f_{B^0}^2 \left|V_{tb}^* V_{td}\right|^2 \qquad (2.15)$$

where G_F is the weak coupling constant, m_B is the B^0 mass, m_W the mass of the $W^\pm$ gauge boson, m_t is the top quark mass, $F(x) \simeq 0.784x^{-0.24}$, η_{QCD} is a factor intended to take into account QCD corrections,[5] B_{B^0} is the non-perturbative "bag-factor", and f_{B^0} is the decay constant of the B^0 meson. Similar terms are also present for u and c quarks, but the t quark mass is so much heavier that mixing is dominated by diagrams involving t quarks.

---

[5] "QCD" stands for Quantum Chromodynamics, the theory of strong interactions. Strong interactions cause small adjustments to the above equation.

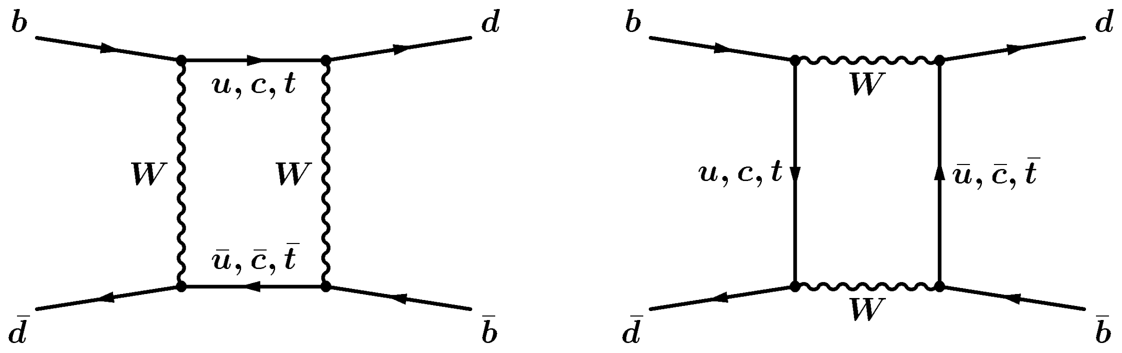

Figure 2-4: Diagrams which lead to $B^0/\overline{B}^0$ mixing.

2.3.2 Unitarity

Because the CKM matrix is simply a change of basis, it must be *unitary* (*i.e.* $V^\dagger = V^{-1}$). The unitarity constraint $V^\dagger V = VV^\dagger = I$ is equivalent to eighteen equations, but six of them are the complex conjugates of six others. The twelve different equations are:

$$V_{ud}V_{ud}^* + V_{cd}V_{cd}^* + V_{td}V_{td}^* = 1 \tag{2.16}$$

$$V_{us}V_{us}^* + V_{cs}V_{cs}^* + V_{ts}V_{ts}^* = 1 \tag{2.17}$$

$$V_{ub}V_{ub}^* + V_{cb}V_{cb}^* + V_{tb}V_{tb}^* = 1 \tag{2.18}$$

$$V_{ud}V_{ud}^* + V_{us}V_{us}^* + V_{ub}V_{ub}^* = 1 \tag{2.19}$$

$$V_{cd}V_{cd}^* + V_{cs}V_{cs}^* + V_{cb}V_{cb}^* = 1 \tag{2.20}$$

$$V_{td}V_{td}^* + V_{ts}V_{ts}^* + V_{tb}V_{tb}^* = 1 \tag{2.21}$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \tag{2.22}$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \tag{2.23}$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \tag{2.24}$$

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{2.25}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.26}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{2.27}$$

The six equations on the left are diagonal; each element is multiplied by its complex conjugate. They are useful for measuring magnitudes, but contain no phase information. The other six equations are complex, and form triangles in the complex plane. As will be shown later, the magnitude of a CP violating phenomenon is proportional to $\sin 2\Phi_{CKM}$ where Φ_{CKM} is the one of the angles of one of these triangles.

The angles of the triangles related to equations 2.23 and 2.26 are particularly interesting, because they can all be large. The terms in these equations are all of order λ^3, so the sides of the triangles should be of similar length. In the other four triangles, one of the sides is much shorter than the other two, so at least one of the angles will be very small. Bjorken [37] took equation 2.26, divided by $V_{cd}V_{cb}^*$, and formed the triangle pictured in figure 2-5.

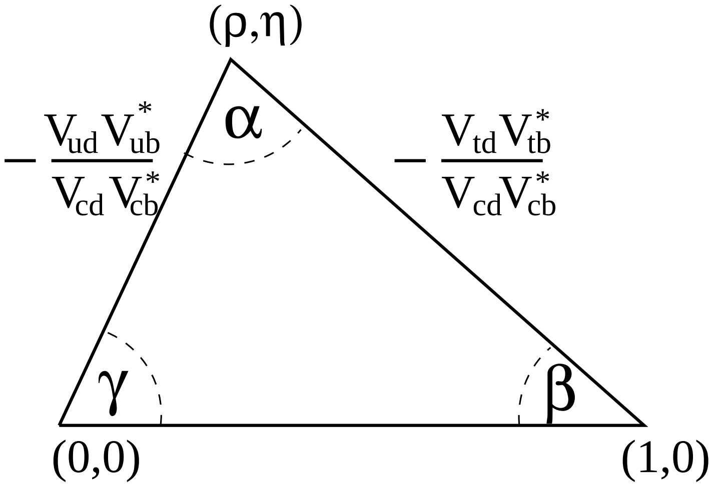

Figure 2-5: The Bjorken triangle.

Precise measurements of the sides and angles of this triangle lead to tests of the unitarity of the CKM matrix. If the lengths and angles are not consistent with a closed triangle, that would indicate a failure of the Standard Model, possibly predicting a fourth generation of quarks. Reference [36] describes how these sides and angles can be measured, and provides analysis of experiments which have measured the quantities.

The magnitude of ϵ, the parameter which describes the magnitude of CP violation in neutral kaons (Section 1.2.3), is shown by [36] to be:

$$|\epsilon| \simeq CA^2\lambda^6\eta \tag{2.28}$$

where A, λ, and η are from the mixing matrix above, and $C \simeq 91$ is a constant.[6] The magnitude of CP violation in neutral kaons is proportional to η, as was mentioned in the chapter introduction. Reference [36] predicts that $\eta = 0.33 \pm 0.05$.

---

[6] The actual formula is considerably more complicated, and many different terms are combined into the single constant C.

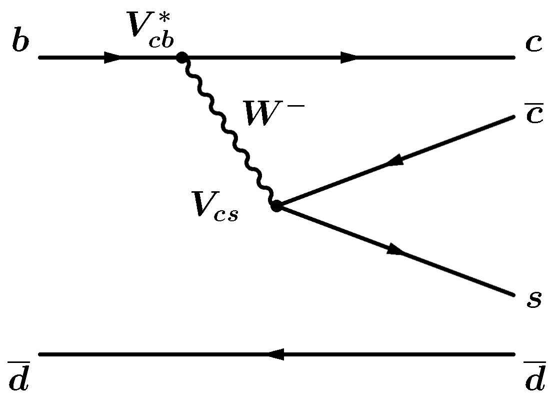

Figure 2-6: The main tree-level diagram for $B^0 \to J/\psi K^0$.

2.3.3 CP Violation

Given that $\eta \neq 0$, the Standard Model also predicts that there should be CP violation in the decays of B^0 mesons. This section derives the relationship between the magnitude of CP violations in the decay $B^0(\overline{B}^0) \to J/\psi K_S^0$ and the angle β, which the phase: $-\arg(V_{td})$.

CP Violation in $B^0(\overline{B}^0) \to J/\psi K_S^0$

The main tree-level diagram in the decay $B^0 \to J/\psi K^0$ is shown in figure 2-6. The weak phase associated with this decay is $\arg(V_{cb}^* V_{cs})$ which is zero up to $\mathcal{O}(\lambda^5)$. This phase is approximated to be zero, which implies:

$$\langle J/\psi K^0 | B^0 \rangle = \langle J/\psi \overline{K}^0 | \overline{B}^0 \rangle \equiv \mathcal{A} \tag{2.29}$$

Combining equations 1.6, 2.12, and 2.29, and approximating $|p_K| = |q_K| = 1$ leads to:

$$\begin{aligned} \langle J/\psi K_S^0 | B^0(t) \rangle &= \langle K_S^0 | K^0 \rangle \langle J/\psi K^0 | B^0 \rangle \langle B^0 | B^0(t) \rangle \\ &\quad + \langle K_S^0 | \overline{K}^0 \rangle \langle J/\psi \overline{K}^0 | \overline{B}^0 \rangle \langle \overline{B}^0 | B^0(t) \rangle \\ &\simeq \mathcal{A} e^{-\Gamma t/2} e^{-iMt} \left[\cos(\Delta m t/2) - i (q_B/p_B) \sin(\Delta m t/2) \right] \end{aligned} \tag{2.30}$$

where the first term indicates the direct decay $B^0 \to J/\psi K_S^0$, and the second term indicates the mixed decay $B^0 \to \overline{B}^0 \to J/\psi K_S^0$.

The phase q_B/p_B is determined [38], using the diagrams in figure 2-4, to be:

$$\frac{q_B}{p_B} \simeq \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} = e^{-2i\beta} \tag{2.31}$$

where diagrams involving t quarks are assumed to dominate.

Therefore the amplitude for $B^0 \to J/\psi K_S^0$ is:

$$\langle J/\psi K_S^0 | B^0(t) \rangle \simeq \mathcal{A} e^{-\Gamma t/2} e^{-iMt} \left[\cos(\Delta m t/2) - i e^{-2i\beta} \sin(\Delta m t/2) \right], \tag{2.32}$$

and the partial widths are:

$$\begin{aligned} \Gamma(B^0(t) \to J/\psi K_S^0) &= |\langle J/\psi K_S^0 | B^0(t) \rangle|^2 \\ &= |\mathcal{A}|^2 e^{-\Gamma t} (1 - \sin 2\beta \sin \Delta m t) \qquad (2.33) \\ \Gamma(\overline{B}^0(t) \to J/\psi K_S^0) &= |\mathcal{A}|^2 e^{-\Gamma t} (1 + \sin 2\beta \sin \Delta m t) \qquad (2.34) \end{aligned}$$

Thus, the time-dependent asymmetry is given by:

$$\begin{aligned} A_{CP}(t) &= \frac{\Gamma(B^0(t) \to J/\psi K_S^0) - \Gamma(\overline{B}^0(t) \to J/\psi K_S^0)}{\Gamma(B^0(t) \to J/\psi K_S^0) + \Gamma(\overline{B}^0(t) \to J/\psi K_S^0)} \\ &= -\sin 2\beta \sin \Delta m t \qquad (2.35) \end{aligned}$$

Therefore, measurement of the time-dependent asymmetry between $B^0 \to J/\psi K_S^0$ and $\overline{B}^0 \to J/\psi K_S^0$ should yield a sinusoidal oscillation, of frequency Δm and amplitude $\sin 2\beta$. As alluded to previously, the amplitude of the effect is related to the parameter η: the angle β is the phase $\beta \equiv -\arg(V_{td}) = \tan^{-1}(\eta/1-\rho)$, so:

$$\sin 2\beta = \frac{\eta(1-\rho)}{\eta^2 + (1-\rho)^2}. \tag{2.36}$$

The parameter η is constrained by the magnitude of CP violation in neutral kaons (among other effects [36]). Since $\eta \neq 0$, the Standard Model indicates that there should be a nonzero asymmetry in the above decays. Reference [36] predicts $\sin 2\beta = 0.68 \pm 0.10$.

If, however, CP violation is due to some mechanism other than a complex phase in the CKM matrix (such as a superweak interaction), then the above prediction would not hold, and no oscillation could be seen. Thus, studying the decays $B^0(\overline{B}^0) \to J/\psi K_S^0$ can provide a powerful test of the Standard Model formalism.

Chapter 3

Studying B Mesons at CDF

The data used in this experiment were collected by the CDF detector. CDF is a multipurpose detector centered around one of the interaction points of the Fermilab Tevatron $p\overline{p}$ collider. The detector and the Tevatron are described in Chapter 4. This chapter provides a summary of some of the practical aspects of the experiment. It describes how B mesons are produced in $p\overline{p}$ collisions, then discusses some issues specific to the study of B mesons, and also discusses issues related to the $p\overline{p}$ environment.

3.1 Production of B Mesons in $p\overline{p}$ Collisions

The B mesons used in this study are produced in proton-antiproton ($p\overline{p}$) collisions. This section discusses how protons and antiprotons interact to produce b and $\overline{b}$ quarks, and how B mesons are formed from these quarks.

3.1.1 Proton Structure

A proton is "made up" of three quarks, two u quarks, and one d quark. An antiproton is likewise made up of three antiquarks: two $\overline{u}$ and one $\overline{d}$. These *valence* quarks are held together by a "cloud" of gluons. The gluons can split into $q\overline{q}$ pairs, which can interact with a quark or gluon from the other hadron. These *sea* quarks are virtual particles and can be of any flavor, even ones with mass greater than that of the proton. All of these particles, the valence quarks, the sea quarks, and the gluons, are collectively referred to as *partons*.

In low energy interactions, protons (antiprotons) act like point particles with electric charge +1 (−1). At higher energies, interactions occur between the partons inside the proton and antiproton. Each parton carries a fraction x of the momentum of the hadron. The *structure function* $F_p^H(x)$ describes the probability for parton of type p in a hadron of type H to have a momentum fraction x. The sea quarks are more numerous at low x (like gluons), so their structure functions all peak at low x.

The transverse (perpendicular to the proton direction) momentum of a parton is limited by the size of the proton, and will therefore not be more than a few hundred MeV/c. The longitudinal (parallel to the proton direction) momentum will be xP, where P is the momentum of the proton. In a high energy collision, P will be much larger than the proton mass ($P/m \approx 1000$ at the Tevatron), so the parton momentum will tend to be nearly parallel to the proton momentum. When a parton with momentum x_1P from a proton interacts with a parton with momentum x_2P from an oncoming antiproton, the center-of-mass frame will have energy $2\sqrt{x_1x_2}P\,c$, and will be boosted by $\beta = (x_1 - x_2)/(x_1 + x_2)$ in the direction of the proton momentum. These fractions will be different for every interaction, so the energies and boosts of interactions will vary.

3.1.2 Production of b Quarks in $p\overline{p}$ Collisions

The cross-section for producing a b quark in a $p\overline{p}$ collision is calculated by [41]:

$$\frac{d^2\sigma}{dp_T(b)dy(b)}(p\overline{p} \to bX) \;=\; \sum_{i,j} \int dx_i dx_j F_i^p(x_i) F_j^{\overline{p}}(x_j) \frac{d^2\hat{\sigma}(ij \to bX)}{dp_T(b)dy(b)} \qquad (3.1)$$

where i and j are partons, and $F_{i,j}^{p,\overline{p}}$ are the relevant proton and antiproton structure functions. The term $\frac{d^2\hat{\sigma}(ij\to bX)}{dp_T(b)dy(b)}$ is the parton-level cross-section formula for the process $ij \to bX$.

These cross-sections are calculated perturbatively, in powers of α_s, the QCD coupling constant. Figure 3-1 shows three diagrams by which a pair of partons can interact to produce a $b\overline{b}$ pair. The amplitudes for these processes are proportional to

α_s^n, where n is the number of gluon vertices. These diagrams are the simplest ones possible, with two vertices each, and are therefore the processes of lowest order in α_s. They are therefore called *leading-order* diagrams. The four diagrams in figure 3-2 each have three vertices, and are called *next-to-leading-order*. Other, higher-order diagrams are also possible.

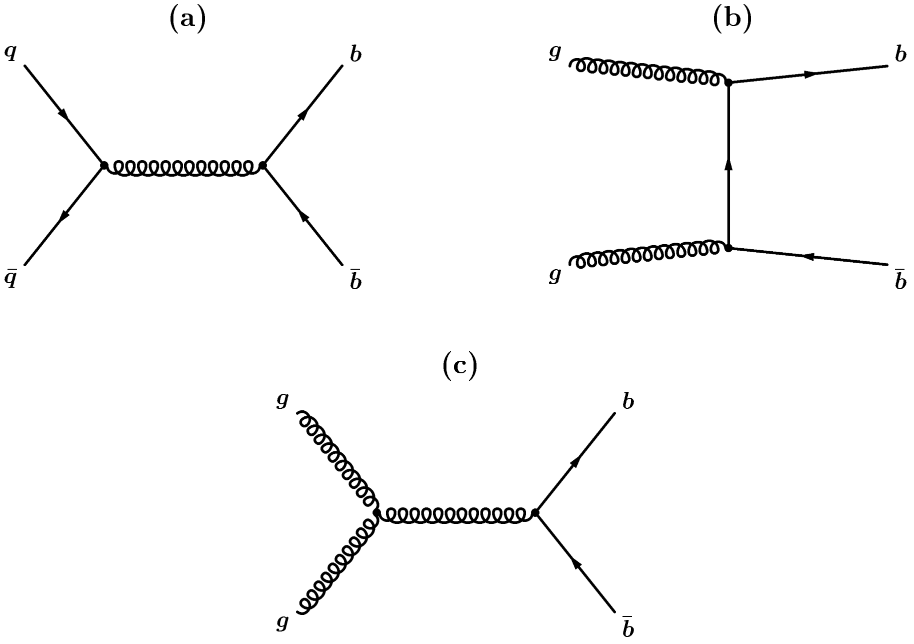

Figure 3-1: Leading-order diagrams for $b\bar{b}$ production.

Measurement of the b Quark Production Cross-section

The cross-section in equation 3.1 is calculated, using the leading-order and next-to-leading-order diagrams [41]. The curve in figure 3-3 shows the integral of this cross-section for all b quark p_T above a minimum value ($p_{T,\mathrm{min}}$). The points in this figure indicate measurements of the cross-section made at CDF [42, 43]. The cross-section measured by CDF is roughly a factor of two higher than the theoretical predictions, which indicates that diagrams of even higher order are needed to properly calculate the cross-section.

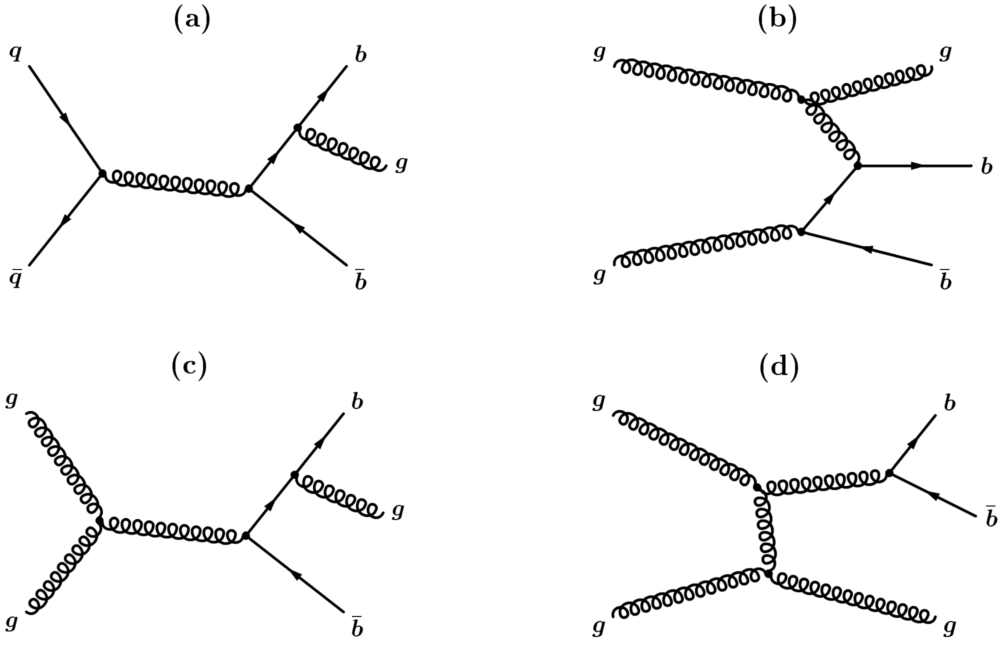

Figure 3-2: Next-to-leading-order diagrams for $b\bar{b}$ production.

3.1.3 Hadronization of b Quarks

One of the complications of calculating the above cross-sections is that the value of α_s depends on the energy (Q^2) of the interaction. For high Q^2, the value of α_s is small (~ 0.1), and perturbative calculations are effective. But for low Q^2, the value of α_s becomes large, and can even be greater than 1. In these cases, the perturbative calculations fail, and phenomenological models are the only way to calculate the expected effects.

The process of b quarks forming into B hadrons (called *hadronization*) is a low Q^2 process, so perturbative calculations are not reliable (or even meaningful when $\alpha_S > 1$). One commonly used model for this process is *string fragmentation* [39]. In this model, the quark-antiquark interaction is modeled with a potential $V(r) \propto kr$, reminiscent of that of a string (the "string" in this model is the a "cloud" of gluons, as described above). As the quark and antiquark separate, the string stretches, until it "breaks," and a new quark-antiquark pair "pops" out of the vacuum to form the new ends of the new strings (see figure 3-4).

These new strings also stretch and break, producing more quark-antiquark pairs.

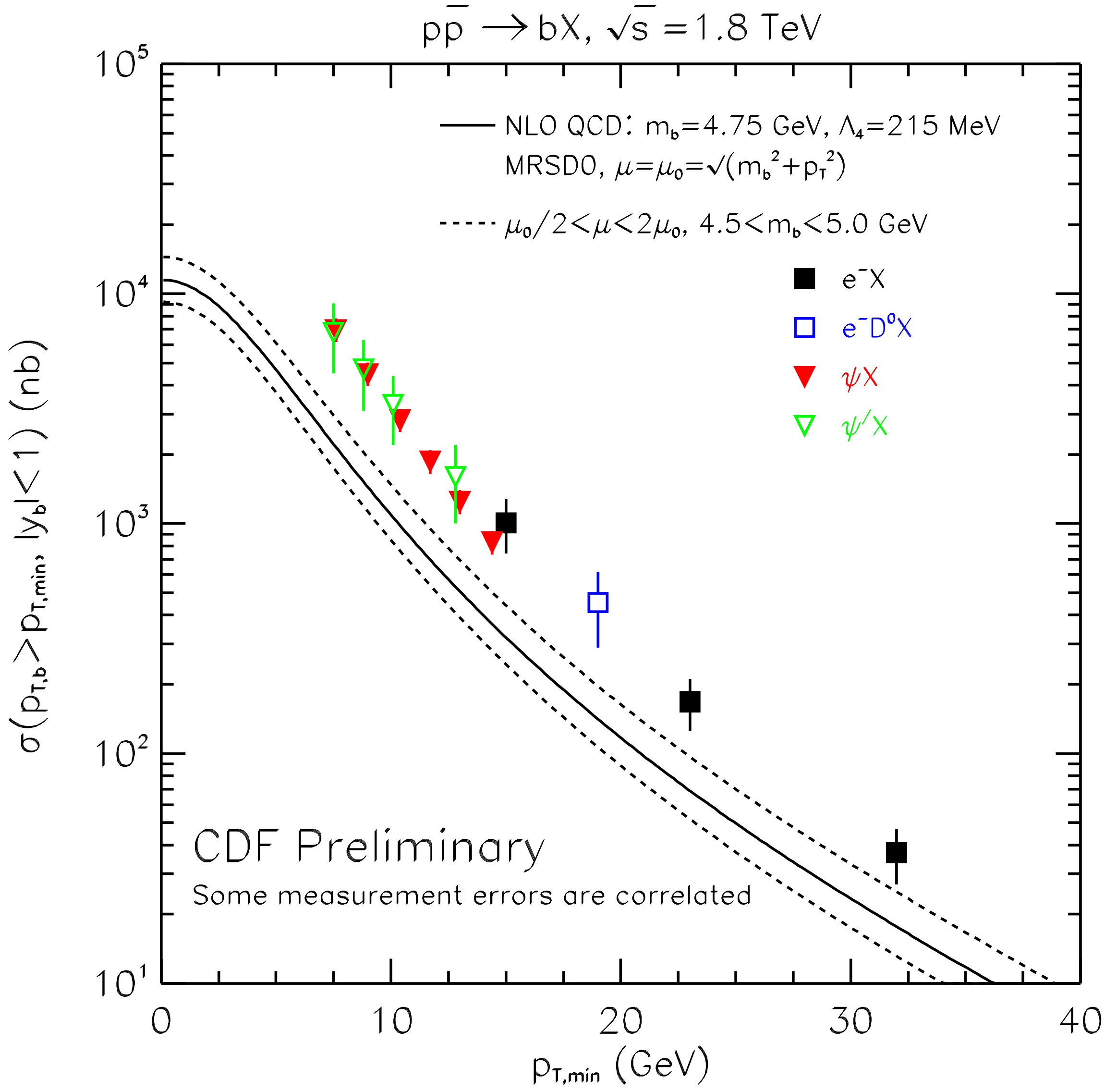

Figure 3-3: The integrated b quark production cross-section. The horizontal scale is the p_T cutoff, and the vertical scale is the integrated cross-section for all b quark p_T above the cutoff. The curve indicates the theoretical calculation, and the points are values measured at CDF.

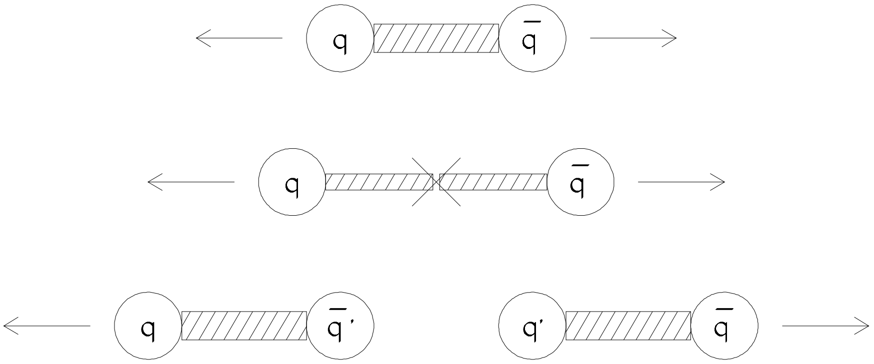

Figure 3-4: A visual representation of string fragmentation. As a $q\overline{q}$ pair separates, the "string" between them breaks, producing a new $q'\overline{q}'$ pair to form the ends of the new strings.

This process continues until each string connects a quark and an antiquark with similar enough momenta there is no longer sufficient energy to generate new $q\overline{q}$ pairs, and no more strings are broken. At this point, *hadronization* is complete. Particles that are sequential in this chain of hadrons will tend to have momenta nearer to one another than particles which are separated in the chain.

Through this process, the b quarks will hadronize into B hadrons. Other particles will also be produced, so the B hadron energy will be lower than that of the original b quark. The particles produced along with the B hadron in the B hadronization process are referred to as *fragmentation* particles.

Flavors of B Hadrons

The probability of a given flavor of quark-antiquark pair popping out of vacuum depends on the mass of the quark and antiquark. The lightest quarks are u and d, so the most common B hadrons are B^+ $(\overline{b}u)$ and B^0 $(\overline{b}d)$.[1] These each comprise roughly 38% of the B hadrons produced [32]. The next lightest quark type is s, and

---

[1] For convenience in this thesis, whenever a particle is referred to, the charge-conjugate particle is also implied. Exceptions to this rule will be made clear by context.

B_s^0 ($\bar{b}s$) is the next most common type of B meson, comprising roughly 11% of B hadrons. The c quark is much more massive than the u, d, and s quarks, so B_c^+ mesons are considerably more rare, comprising only $\approx 1/1000$ of B hadrons produced in $p\bar{p}$ collisions [40].

In addition, if two quark-antiquark pairs "pop" out of vacuum, a baryon can be formed, instead of a meson. Λ_b baryons (udb) comprise roughly 13% of B hadrons formed, slightly more than the fraction of B_s^0 mesons.

Underlying Event

The partons from the proton and antiproton not directly involved in the b quark production are called *beam remnants*. They also undergo hadronization, and the hadrons produced from the beam remnants are called *underlying event* particles. The momenta of the underlying event particles are generally uncorrelated with the final B hadron direction, while the fragmentation particles tend to be clustered near the B hadron direction.

3.2 Physics of B Hadrons

3.2.1 B Hadron Lifetime

B hadrons decay via the weak interaction, and the average lifetime of a B hadron is only 1.5 ps. This is much too fast for them to be detected directly, as even energetic B hadrons only travel a few millimeters before decaying. However, B hadrons can be identified through their decay products: the products of the decay of the B are detected directly, and their paths are extrapolated back to the B decay point. The distance from this point to the point where the B was produced is the decay-length of the B.

Figure 3-5 shows a measurement of the average B hadron lifetime, made at CDF [44] using the inclusive decay $B \to J/\psi X$, with $J/\psi \to \mu^+\mu^-$. As described above, the paths of the μ^+ and μ^- are extrapolated back to where they intersect, which is the B decay vertex.[2] The plot shows the distribution of the number of recorded events versus the reconstructed pseudo-proper decay-length λ, where

$$\lambda = L_{XY} \frac{m}{p_T F(p_T)} \tag{3.2}$$

where L_{XY} is the reconstructed transverse decay-length, m is the J/ψ mass, p_T is the transverse momentum of the J/ψ, and $F(p_T)$ is a correction factor determined from Monte Carlo.

The narrow peak at $\lambda = 0$ represents J/ψ mesons which do not come from B hadrons, the filled histogram represents fake J/ψ candidates, and the diagonal-striped histogram represents J/ψ mesons which are from B hadrons. The CDF detector resolution is represented by the width of the central peak, and is roughly $50\,\mu$m. This corresponds to an uncertainty on the decay-time of a B hadron of 0.17 ps, which is roughly 1/10th of the B hadron lifetime.

3.2.2 B^0 Mixing

As described in Chapter 2, neutral B mesons *mix*; a particle produced as B^0 can decay as a $\overline{B}^0$ or vice-versa. The probability that a particle produced as a B^0 will decay as a $\overline{B}^0$ after a time t is (from equation 2.14):

$$P(B^0 \to \overline{B}^0) = e^{-\Gamma t} \left(\frac{1 - \cos(\Delta m t)}{2} \right) \tag{3.3}$$

where $\Delta m = 0.48\,\hbar \text{ps}^{-1}$ is the difference of the masses of the two mass eigenstates. The asymmetry in the number of B^0 mesons which decay as B^0 versus those that decay as $\overline{B}^0$ is:

$$A(t) = \frac{P(B^0 \to B^0) - P(B^0 \to \overline{B}^0)}{P(B^0 \to B^0) + P(B^0 \to \overline{B}^0)} = \cos(\Delta m t) \tag{3.4}$$

Figure 3-6 shows a published B^0 mixing measurement made at CDF [45]. The asymmetry plotted is the $A(t)$ above, but with the amplitude reduced by a *dilution*

---

[2] The muons come from the J/ψ decay, so they extrapolate back to the J/ψ decay vertex. But J/ψ mesons decay instantaneously ($\tau_{J/\psi}/\tau_B \simeq 10^{-7}$), so the J/ψ decay vertex and the B decay vertex are the same.

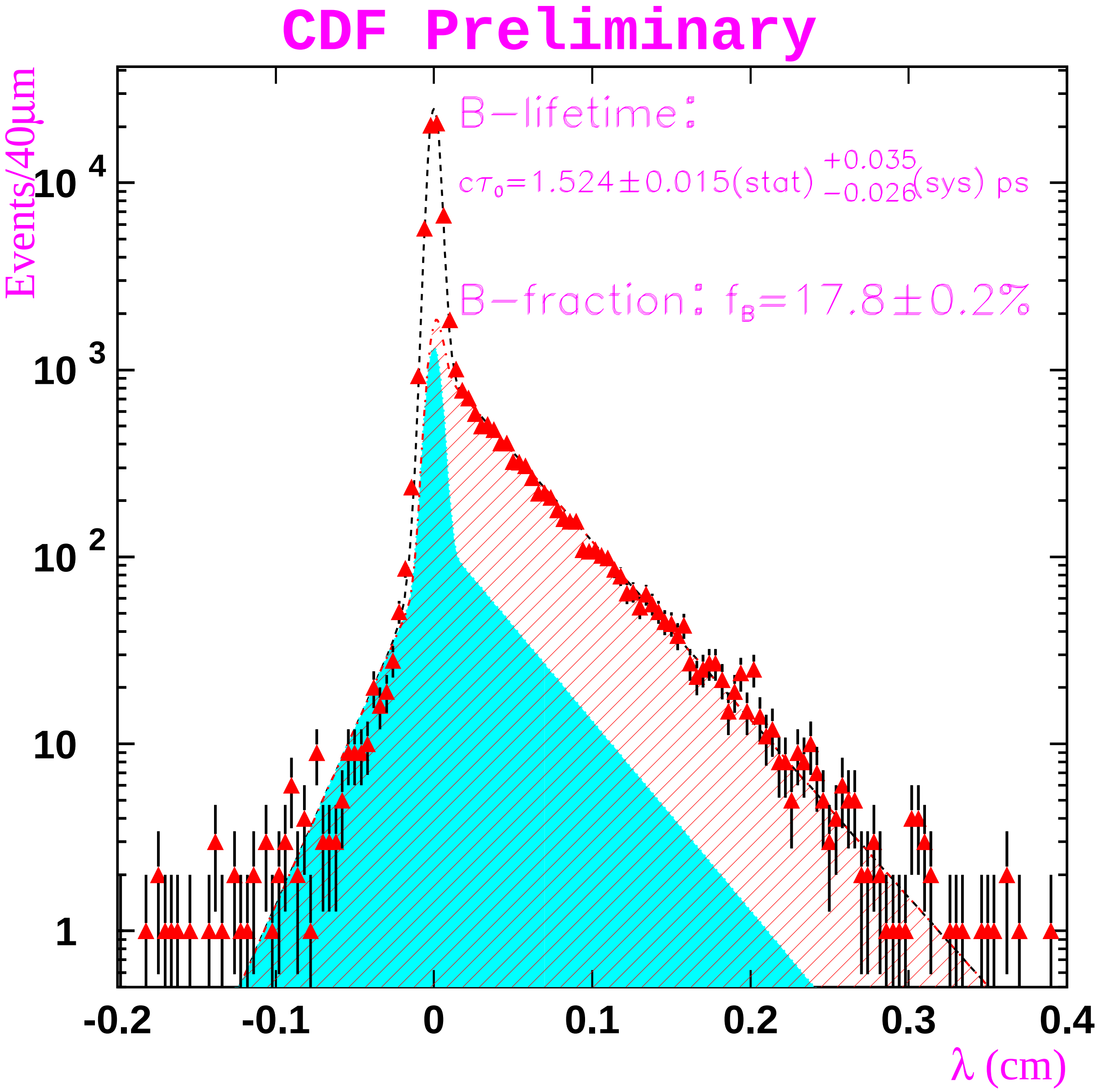

Figure 3-5: The pseudo-proper decay-lengths of J/ψ mesons. The points represent the data measurement at CDF. The solid histogram represents the J/ψ background, the striped one represents the B hadron signal, and the dashed curve represents the sum of these, plus the distribution of prompt J/ψ mesons.

factor $\mathcal{D}$, which will be described in the next section. The period of the oscillation is $2\pi/\Delta m = 13ps$, roughly 9 times the B^0 lifetime, and 80 times the CDF resolution. While these oscillations are very fast, it is well within CDF's capabilities to observe them.

3.2.3 Flavor Tagging

The observation of B^0–$\overline{B}^0$ mixing, and the measurement of the mixing frequency, Δm, requires knowledge of the flavor of the B meson (*i.e.* whether it was a B^0 meson or a $\overline{B}^0$ meson) when it was produced and also when it decayed (so that the two flavors can be compared). Determining the flavor at production is called *flavor tagging*, and several algorithms for flavor tagging have been employed at CDF and elsewhere.

A perfect tagging algorithm will tag every B meson, and will identify all of them correctly. A practical tagging algorithm will tag only a fraction ϵ (called the *tagging efficiency*) of them, and of those tagged, only a fraction P (the *tagging purity*) will be identified correctly. The asymmetry in the number of correct tags (N_{RS}) versus the number of incorrect tags (N_{WS}) is the *dilution factor*:

$$\mathcal{D} \equiv \frac{N_{RS} - N_{WS}}{N_{RS} + N_{WS}} = 2P - 1. \tag{3.5}$$

The uncertainty on an asymmetry measured on a sample of N events, tagged with efficiency ϵ and dilution $\mathcal{D}$ is approximately $1/\sqrt{N\epsilon\mathcal{D}^2}$. Thus, the quantity $\epsilon\mathcal{D}^2$ is called the *effective tagging efficiency*.

Tagging Algorithms

In $p\overline{p}$ collisions, b quarks are produced in pairs, one b and one $\overline{b}$. Identifying the flavor of one of them allows the flavor of the other to be inferred. This is called *opposite-side* tagging, because the opposite B hadron is used for the tag. The two most popular methods of opposite-side tagging are *lepton* tagging [47, 48] and *jet-charge* tagging [49, 50].

B hadrons decay semileptonically to $e\nu X$ or $\mu\nu X$ roughly 20% of the time. By identifying the flavor of the lepton, the flavor of the b quark in the decaying hadron

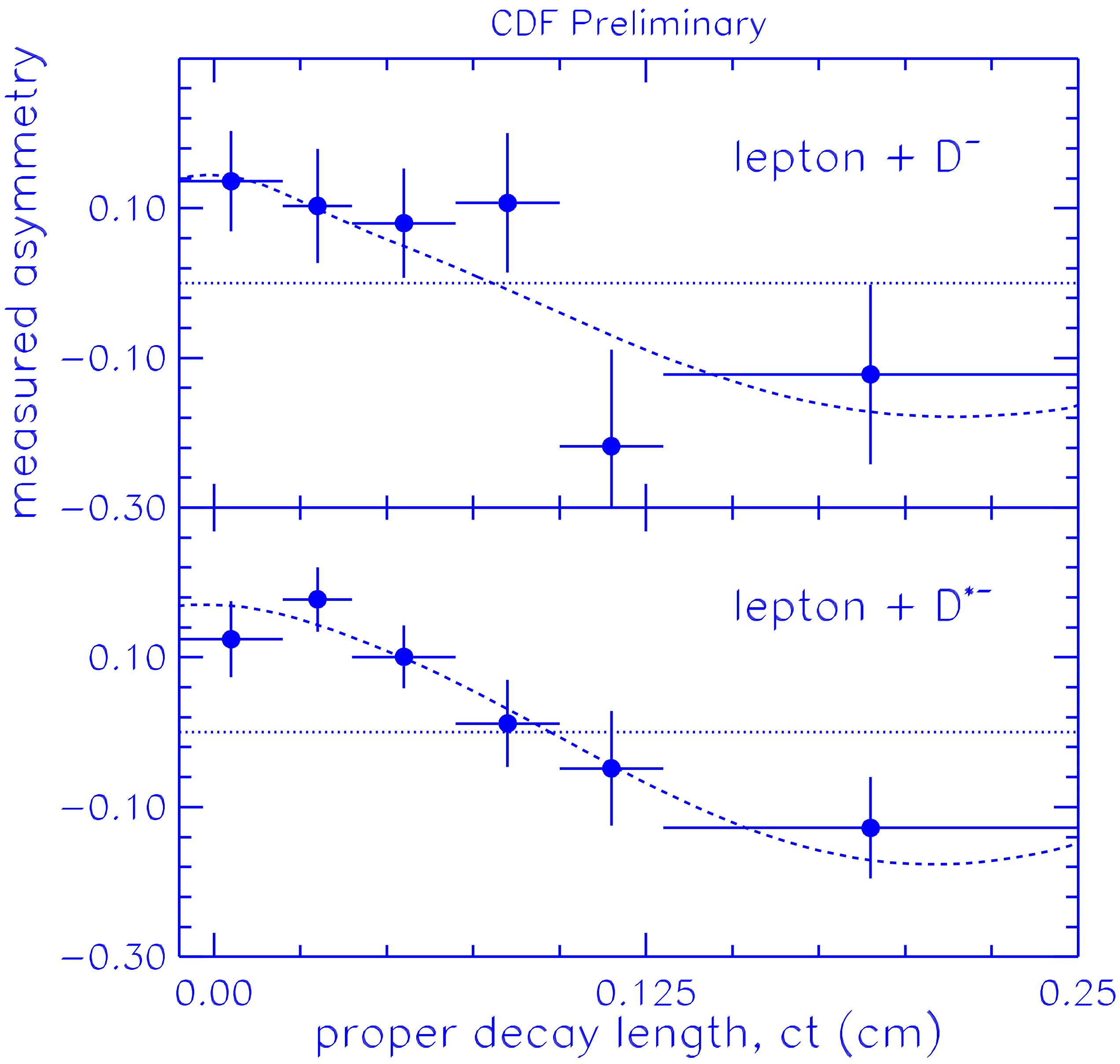

Figure 3-6: Measurement of B^0 mixing. The points are the data, and the curves represent the $\mathcal{D}\cos\Delta mt$ mixing shape.

can be determined: ℓ^+ with $\overline{b}$ and ℓ^- with b. This method can have very high purity, but the efficiency is limited by the low branching fraction to leptons.

Jet-charge tagging attempts to identify the opposite-side B hadron by identifying its displaced point of decay (vertex). Once the displaced vertex is found, statistical methods are used to determine the flavor of the hadron. This method has higher efficiency than lepton tagging, but the purity is much lower.

Both of these tagging algorithms suffer from the complication that sometimes the opposite-side hadron which is being used as a tag will mix before it decays. If this occurs, then even a correctly identified lepton or vertex will yield the incorrect tag. This reduces the dilutions of opposite-side tagging algorithms to roughly 75% of what they would be without opposite-side mixing.

Same-side Tagging

Another approach to tagging is *same-side* tagging, where particles produced near the B meson are used for the tag, and the opposite-side B hadron is ignored. As will be explained in Section 7.1, there should be a correlation between the flavor of the B meson and the flavors of the nearby particles produced by the b quark hadronization. As this correlation is independent of the flavor of the B hadron on the opposite side, same-side tagging does not suffer from the complication of opposite-side mixing, like opposite-side tagging algorithms do.

3.3 Characteristics of the $p\overline{p}$ Environment

3.3.1 Background

The chief advantage of studying B physics at a hadron collider is the high rate; the total cross-section for $b\overline{b}$ production at CDF is roughly $100\,\mu b$. From 1992-1995, $100\,pb^{-1}$ of data were collected, which would correspond to roughly 10 billion $b\overline{b}$ pairs produced. This number is much larger than the combined numbers of all experiments ever performed using e^+e^- collisions to produce $b\overline{b}$ hadrons.

The chief disadvantage of studying B physics at a hadron collider is the background. While the cross-section for $b\bar{b}$ production is roughly $100\,\mu b$, the cross-section for any sort of interaction is $75\,\mathrm{m}b$, three orders of magnitude higher! The key to studying B physics at a hadron collider is finding ways to quickly sort out events containing B mesons.

Identification of B mesons Through J/ψ Decays

One way B hadrons are identified is through their decays to J/ψ mesons. Only about 1% of B hadrons decay to modes containing J/ψ's, but there are few enough other sources of J/ψ's so that B decays still account for 25% of the J/ψ's produced at CDF [43] (see figure 3-7). Thus, requiring an event to contain a J/ψ improves the B hadron purity by a factor of 250.

Furthermore, J/ψ mesons decay instantaneously, and B hadrons are the only particles which decay to J/ψ mesons but live long enough to travel a measurable distance before decaying. Thus, by separating those J/ψ events where the J/ψ decay point is at the primary interaction point (called *prompt* events) from those where the J/ψ decay point is displaced from the primary interaction point (called *displaced* events), it is possible to find a sample with very high B purity (see figure 3-5).

Identification of Leptonic Decays of J/ψ mesons

Roughly 88% of J/ψ decays are to hadrons [32]. Half of the rest decay to $\mu^+\mu^-$ and the other half decay to e^+e^-. While reconstructing hadronic decays of J/ψ's might therefore seem the best way to find the most J/ψ's, it is not. The problem is that the background is too high (and triggering, Section 3.3.2, would be essentially impossible).

Identifying a J/ψ meson requires reconstructing its decay particles and extrapolating back to where they intersect. The major source of background is combinations of particles which are not from a J/ψ decay, but happen to have a combined mass near the J/ψ mass. If these particles all originate in the same place (as they often do), there is no way to distinguish them from the decay products of a real J/ψ mesons.

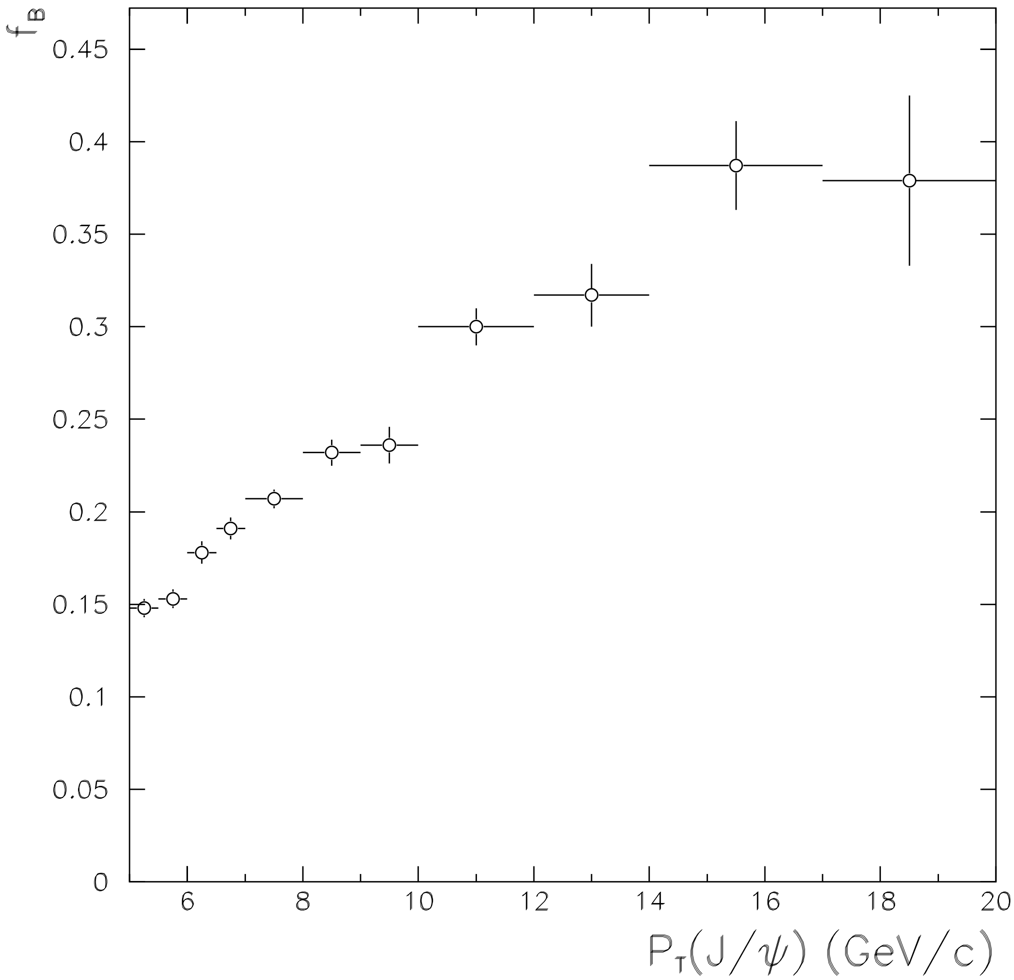

Figure 3-7: The fraction of J/ψ mesons which come from B hadrons (as opposed to other sources), plotted versus the p_T of the J/ψ.

Only about 1 in 20,000 interactions involve J/ψ mesons. Most of the rest involve light quarks and produce many hadrons. This background is so prevalent it is nearly impossible to identify hadronic decays of J/ψ mesons. However, light quark interactions produce very few leptons. Thus, while the leptonic decays of J/ψ mesons are less common than the other modes, they have much less background.

The CDF detector cannot distinguish between pions, kaons, and protons, but it can identify electrons and muons. The electrons are identified because they are so light they undergo much more *bremsstrahlung* radiation than the other charged particles. The muons are identified because they can travel through much more matter without interacting. More details on lepton identification are presented in Chapter 4.

These capabilities allow the CDF detector to take advantage of the fact that leptons are much more likely to come from decays involving the heavier quarks (c, b, t) than from those involving the light quarks.

3.3.2 Triggering

As described above, the total interaction cross-section at CDF is nearly 1000 times higher than that for $b\bar{b}$ production. Thus, most interactions will be discarded as "uninteresting." The decision was made to do so immediately, without even recording the information from the events. One major reason for this decision is that interactions occur at CDF roughly once every 3.5 μs, but it takes 2 ms to read out the detector. Thus, even if every event were considered "interesting," only 1 in 500 could be read out, anyway. This issue is discussed further in Section 4.2.3.

As described above, leptons are relatively rare among interactions involving only light quarks, but are more common in interactions with heavy quarks. The trigger system takes advantage of this by using the presence of one or more leptons as a selection criterion. The triggers used in [43] and [44] above require two muons to be detected, with combined dimuon mass near the J/ψ mass. The triggers used in [42] and [45] require an electron or muon to be present, and that the lepton be highly energetic.

3.3.3 Center-of-Momentum Frame

As referred to above, when a parton with momentum x_1P interacts with an oncoming parton with momentum x_2P, the center-of-momentum frame is not the same as the lab frame, but is boosted by $\beta = (x_1 - x_2)/(x_1 + x_2)$ in the direction of the proton momentum. Since x_1 and x_2 will differ for each event, the center-of-momentum for each interaction will have a different, unknown boost.

Coordinates

The proton and antiproton beams are unpolarized, so the only natural direction to use for coordinates is the direction of the proton beam (which is opposite to that of the antiproton beam). The coordinates used to measure particle momenta are p_T, the momentum transverse to the beamline, ϕ, the azimuthal angle around the beamline, and η, the *pseudo-rapidity.*

The pseudo-rapidity η is closely related to *rapidity* y:

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \tag{3.6}$$

$$\eta = \frac{1}{2}\ln\left(\frac{p+p_z}{p-p_z}\right) = -\ln\tan(\theta/2). \tag{3.7}$$

where $E = \sqrt{m^2 + p^2}$ is the energy of the particle, the z direction is the direction of the proton beam, and θ is the polar angle. A boost of β in the proton-beam direction changes the rapidity by an additive shift: $y' = y + \tanh\beta$. Thus, differences in rapidity are conserved under boosts along the beamline.

When the particle's momentum is much larger than its mass $p \gg mc$, (which is generally the case for particles produced in inelastic $p\overline{p}$ collisions) the pseudorapidity is very close to the rapidity. Using the rapidity is simpler, as it is independent of the particle's mass and it directly corresponds to a direction in the laboratory frame. Since $\Delta\eta$ is (practically) invariant under boosts along the beamline, as is (by definition) $\Delta\phi$, a natural way to measure three-dimensional angles is with $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

Loss of Tags

Since the center-of-momentum varies from event to event, the z-momenta of the B hadrons produced in a $b\bar{b}$ interaction will only be loosely correlated. Even when one B hadron is *central* (has decay products passing through the central, efficient region of the detector) the other B hadron will be central only 40% of the time. This reduces the tagging efficiency for opposite-side tagging algorithms accordingly. Same-side tagging algorithms ignore the opposite-side B hadron and are therefore not affected in this way.

Part II

Experimental Apparatus

Chapter 4

The Experimental Apparatus

This chapter describes the experimental apparatus used in this analysis. As described in Chapter 3, the B mesons used in this analysis are produced in proton-antiproton collisions at $\sqrt{s} = 1.8$ TeV. The proton and antiproton beams were produced by the Fermilab Tevatron, which is described briefly in Section 4.1. The decay products of the B mesons were detected in the Collider Detector at Fermilab (CDF), which is described briefly in Section 4.2, and in more detail in Reference [51].

4.1 The Fermilab Tevatron

Figure 4-1 shows the major components of the accelerator complex. The process leading to the proton-antiproton collisions begins with electrical discharges into a hydrogen gas. The gas is in a Cockroft-Walton chamber, across which there is a large electrostatic field. The discharges produce H^- ions, which are accelerated by the electric field to 750 keV. The ions are then sent down a linear accelerator (Linac), which accelerates them to 400 MeV. On exiting the Linac, the ions pass through a carbon foil, which strips off the electrons, turning the H^- beam into an H^+ (proton) beam.

The beam then passes into the *Booster*, a 75m radius synchrotron, where it is accelerated up to 8 GeV. The proton beam in the Booster is not a continuous beam, but is collected into *bunches*. This allows an *RF* cavity to accelerate the beam.

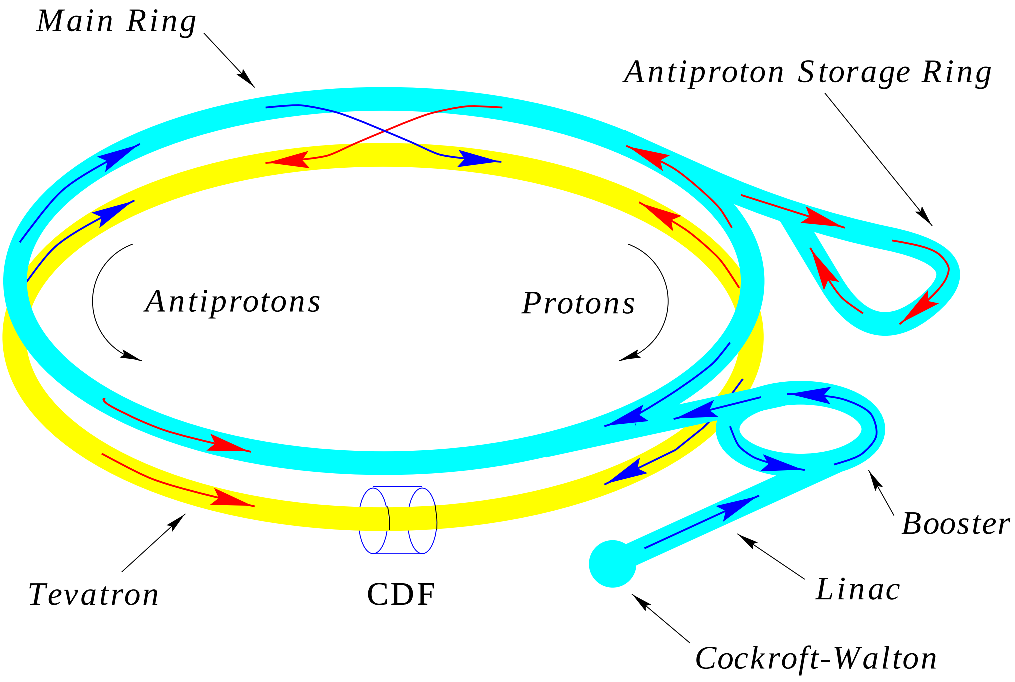

Figure 4-1: The Fermilab accelerator complex (not to scale).

When the protons reach 8 GeV, the beam is passed to the 1 km radius Main Ring synchrotron, which accelerates the protons up to 150 GeV.

Part of the proton beam is then stripped off and sent into a tungsten target, producing antiprotons through the interaction $p + p \rightarrow p + p + p + \overline{p}$. For every million protons fired into the target, roughly 20 antiprotons are produced. These antiprotons are then collected, cooled stochastically, and stored in the antiproton storage ring.

When enough antiprotons are collected, the proton beam is moved to the *Tevatron*, a superconducting synchrotron in the same tunnel as the Main Ring. The antiprotons are then injected into the Main Ring, traveling counter-clockwise (opposite to the protons, see figure 4-1). The antiprotons are then accelerated to 150 GeV and inserted into the Tevatron, still traveling counter to the proton beam. Because the proton and antiproton beams have opposite charge and are traveling in opposite directions, one set of bending magnets can be used to keep both beams traveling in circular paths.

The proton and antiproton beams are then accelerated together to 900 GeV. Finally the beams are focused to give maximum luminosity. The focused beams have cross-sectional widths of about 40μm, and each bunch is about 30 cm long. A typical store has 2×10^{11} protons per bunch, and 6×10^{10} antiprotons per bunch. There are six bunches of each, so bunch crossings occur roughly once every 3.5μs. The beams are set to intersect at two places (*interaction regions*), labeled by their geographical locations as "B0" and "D0". A detector facility is placed around each interaction region; the one at B0 is named "CDF", and the one at D0 is named "D0". The data used in this analysis were collected at the CDF detector.

4.2 The CDF Detector

The Collider Detector at Fermilab (CDF) was built around the Tevatron interaction region at B0. It was designed primarily for the discovery/study of high p_T phenomena (*eg.* the top quark, Z^0 and $W^{\pm}$ bosons), but has proven to be very effective for studying bottom-mesons as well. Figure 4-2 shows a schematic diagram of the CDF detector. This section discusses the components of the CDF detector used in this analysis.

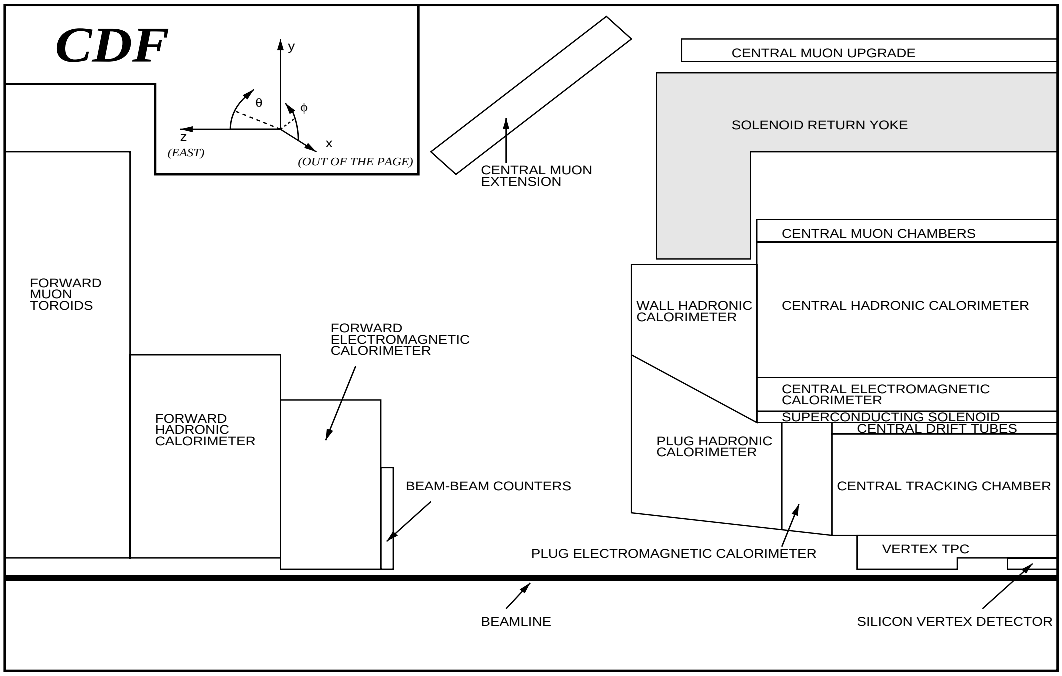

Figure 4-2: A side-view cross-section of the CDF detector. The detector is forward-backward symmetric, and has radial symmetry.

4.2.1 CDF Tracking Chambers

A charged particle passing through matter interacts with the matter electromagnetically, leaving behind a *track* which indicates the path the particle followed. By detecting these electromagnetic interactions, the CDF *tracking chambers* are able to trace the paths of the charged particles which pass through the detector.

Surrounding the central region of the detector is a superconducting solenoidal magnet which is 1.5m in radius and 5m long. Current running through 1164 turns of NbTi/Cu wire provides a fairly uniform magnetic field of 1.4 T along the beamline. Charged particles passing through the region follow helical paths with axes parallel to the beamline. Each of these helices can be described by five parameters: d_0, the distance of closest approach to the beamline, ϕ_0, the ϕ direction of the particle path at the point of closest approach, z_0, the z position at point of closest approach, C, the inverse of the radius of curvature of the track (also called the track "curvature"), and $\cot\theta$, the cotangent of the polar angle θ.

The curvature is inversely proportional to the transverse-momentum of the particle: $C = qB/p_T$, where q is the electric charge of the particle and B is the strength of the magnetic field. The angle θ is related to the pseudorapidity: $\eta = -\ln\tan(\theta/2)$. Thus, the three canonical track coordinates (see Section 3.3.3), p_T, η, and ϕ, are measured directly by three of the helix parameters, C, $\cot\theta$, and ϕ_0. The other two helix parameters locate the point of origin of the track. All five parameters are measured by the Central Tracking Chamber (CTC). The Silicon Vertex Detector (SVX) provides information which improves the accuracy of the d_0 measurement, and the Vertex Time-Projection Chamber (VTX) provides information which is used to improve the accuracy of the z_0 and $\cot\theta$ measurements.

Central Tracking Chamber (CTC)

The CTC is a cylindrical chamber, 2.30 m long, extending from an inner radius of 27.4 cm to an outer radius of 138 cm, and filled with Argon/Ethane/Ethanol gas (49.6/49.6/0.8%). When a charged particle passes through the CTC, it ionizes gas molecules, releasing electrons. To collect these electrons and detect the passage of the particle, the CTC has has 36,504 wires strung between its two endplates. A combined tension of ≈ 25 tons on the wires keeps them taut. Figure 4-3 shows a diagram of a CTC endplate, showing the locations of the slots in which the wire planes are strung.

The slots in figure 4-3 indicate the locations of the *wire-planes*. These planes are placed at high voltage, alternating between positive and negative around the detector. This produces a nearly uniform electric field throughout the rings of wire-planes (called *superlayers*). The magnitude of the electric field is $E_0 = 1350$ V/cm and the RMS of the field variation is $dE_0/E_0 \approx 1.5\%$. The drift velocity of free electrons in the tracking volume is roughly 6×10^4m/s.[1]

Most of the wires in these planes are "field" wires, and serve only to shape the electric field. The other wires are the "sense" wires, and are used directly to collect the electrons released by passing charged particles. The field wires vary from 150 to 300 μm in radius, and the sense wires are 40μm in radius. The sense wires are small enough for *gas amplification*,[2] but the field wires have radii too large to initiate gas amplification. In the CTC, gas amplification produces 3×10^4 electrons for each one originally liberated by a passing charged particle.

When the electrons reach the wire, they produce a pulse of current. The time between the primary proton-antiproton interaction and the detection of this pulse is dominated by the drift-time of the electrons in the gas. Thus, by measuring the timing of the pulse, it is possible to measure how far the original charged particle path was from the wire. The accuracy for these position measurements is better than 200μm.

---

[1] Free electrons in a gas in an electric field will be accelerated by the electric field, but slowed by collisions with gas molecules. The average speed of drifting electrons depends on the type of gas and magnitude of the electric field. However, there is a *saturation* field above which the drift velocity is roughly constant. The electric field throughout most of the CTC is above this saturation level.

[2] If the electric field is high enough, the electrons will occasionally gain enough energy between collisions with the gas to ionize the gas molecules. This releases more electrons which in turn are accelerated and ionize even more gas molecules. This continues until the electrons reach the wires producing the field and are collected. This process is called *gas amplification*. Since the electric field near individual wires increases as $1/r$, the field near the sense wires gets high enough to start gas amplification.

Two charged particles passing through the detector can be resolved separately, even when they are as close as 5mm apart.

While the magnetic field is important for momentum measurement, it presents a minor complication in the CTC. With crossed electric and magnetic fields, the electrons no longer drift along the electric field direction, but at an angle β with respect to it, where β is given by:

$$\beta = \frac{v(E, B = 0)B}{kE} \tag{4.1}$$

where $v(E, B = 0)$ is the drift velocity without the magnetic field, B and E are the magnetic and electric field magnitudes, and k is a parameter which depends on the gas chosen (in this case, $k \simeq 0.4$). For the electric and magnetic fields in the CTC, $\beta \simeq 45^\circ$.

It is best if the drift direction is perpendicular to the path of the charged track being measured. If the angle between the two directions is θ, the drift distance will be $D/\sin(\theta)$, where D is the actual distance to the wire. The uncertainty on the drift distance is constant, so the uncertainty on the distance from the wire increases as $1/\sin(\theta)$. Since each particle passing through the detector will be traveling in a different direction, maximizing $\sin(\theta)$ for some directions will make it smaller for others. The CTC was designed to be most effective for high-energy particles, which have little curvature and nearly radial tracks. This required the drift direction to be azimuthal, so the wire places are tilted by 45° with respect to the radial direction (see figure 4-3.)

There are a few other advantages to this tilt. First, the tilt allows the wire planes to overlap, which means that high p_T tracks must pass near at least one wire in each superlayer. This is taken advantage of by the Central Fast Tracker in the level-2 trigger, as described in Section 4.2.3.

While it is possible to determine how close a particle has passed to a given wire, it is not possible to determine on which side it passed. A particle passing by a plane of wires leaves several "hits," which are grouped into a track "stub." Unfortunately the left-right ambiguity means there are two stubs, one real stub and one "ghost" stub. The angle between these stubs for high p_T tracks is $\Delta = \tan^{-1}(2\tan(\beta)) \simeq 70°$. Having large β and tilted wire planes simplifies the resolution of the left-right ambiguity for high p_T tracks.

Figure 4-4 shows the hits left by a high p_T track in a real CDF event. Both the real and ghost hits are visible. The real track stubs are differentiated from the ghost stubs by lining them up with stubs from other layers. As the figure indicates, it is very easy to differentiate between the real and ghost stubs for high p_T tracks. These stubs are collected into tracks, which provide the d_0, ϕ_0 and C of the track helix. The transverse momentum accuracy is $\delta p_T/p_t^2 < 0.002\,(\mathrm{GeV/c})^{-1}$, and the impact-parameter accuracy is $\delta d_0 \simeq 200\mu\mathrm{m}$. The azimuthal angle accuracy is $\delta\phi \simeq 0.02$. Measurements of z_0 and $\cot\theta$ are taken from *stereo* information.

The superlayers are labeled $\{0..8\}$, from the innermost out. The even superlayers are the *axial* layers, and have 12 sense wires each. The odd ones are the *stereo* layers, and have only 6 sense wires each. Superlayers 1 and 5 are tilted by $+3°$ $(d(r\phi)/dz = 1/120)$, and superlayers 3 and 7 are tilted by $-3°$. Hits in these superlayers are offset by an amount which depends on where the hit is in z. The hits on the odd superlayers in figure 4-4 do not line up with the track because of this offset. Combining this stereo information with the axial track, z can be calculated for each stereo hit (with accuracy $\delta z = 200\mu\mathrm{m}/\sin 3° = 4\mathrm{mm}$). These several z measurements are combined to give z_0 and $\cot\theta$. The z_0 accuracy is $\delta z_0 \simeq 1\mathrm{cm}$, and the $\cot\theta$ accuracy is roughly $\delta\cot\theta \simeq 0.01$.

Silicon Vertex Detector

B mesons have lifetime of $\approx 1.5ps$, so a typical B meson travels about 1mm before decaying. D mesons have lifetimes of roughly 1/3 this value. While it is possible to resolve such long-lived particles with an individual track impact-parameter resolution of $\approx 200\mu\mathrm{m}$, having more accuracy would greatly help background rejection and improve the accuracy of lifetime measurements. The SVX was included at CDF specifically for this purpose.

As such, the SVX is the detector component closest to the beam. It has four layers,

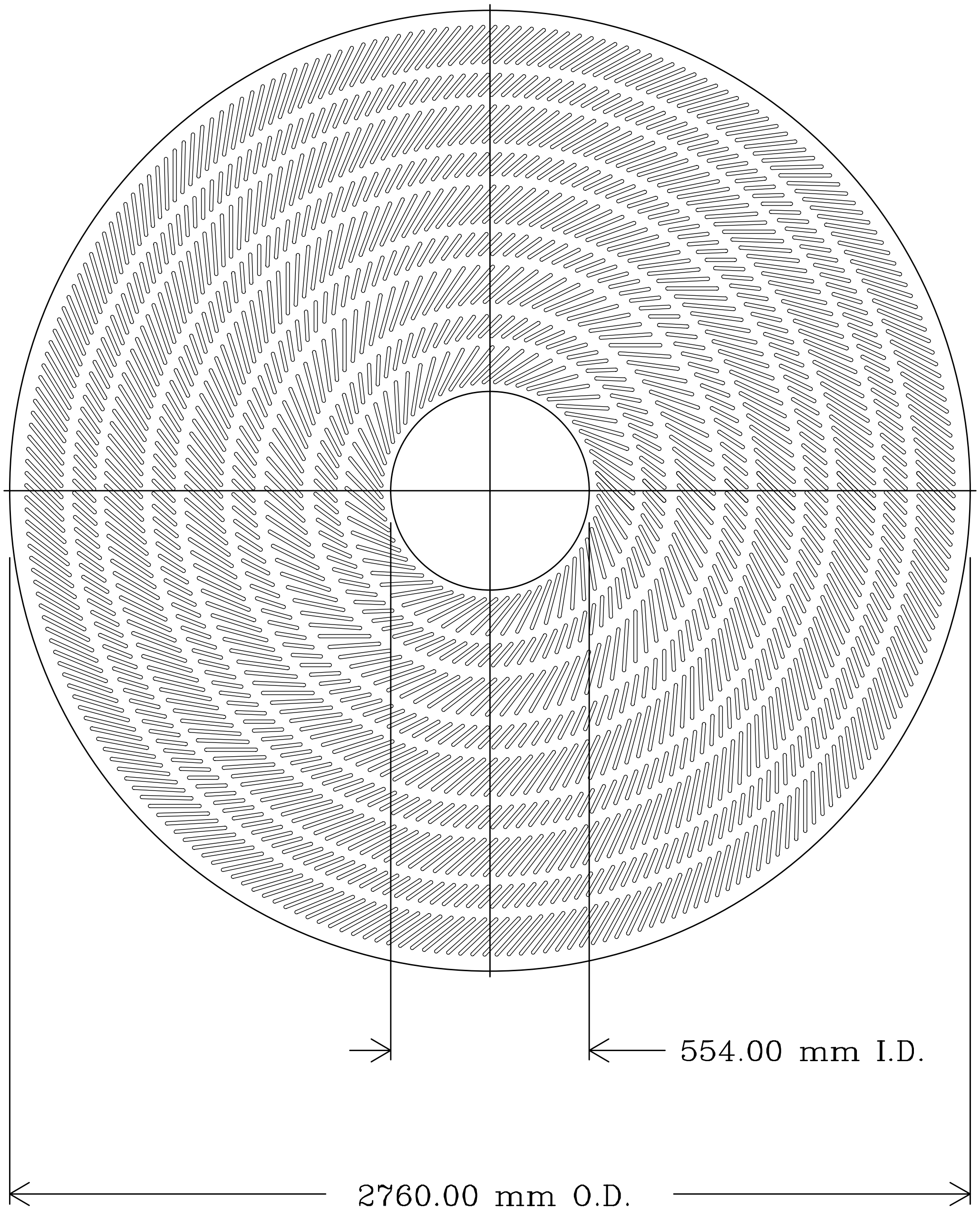

Figure 4-3: View of the CTC endplate, with wire-plane slots indicated.

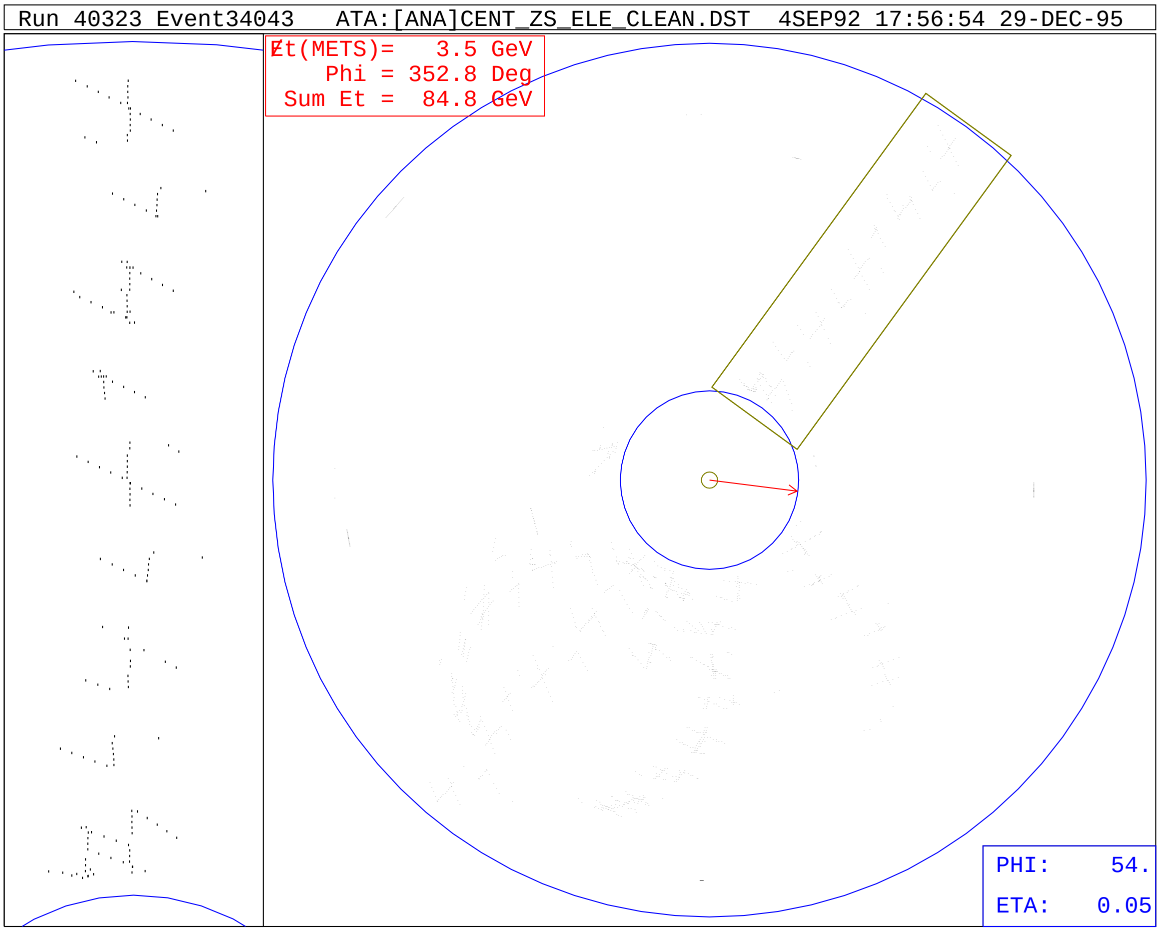

Figure 4-4: Display of CTC information for a $Z^0 \rightarrow e^+e^-$ candidate from 1992. The display on the left is an expansion of the indicated region in the display on the right. Dots indicate the locations of the recorded hits; both real and ghost hits are shown. The large angle between the real and ghost stubs for the high p_T tracks makes it simple to differentiate between them. The hit information assumes $z = 0$, so the points on the stereo layers do not line up exactly with those on the axial layers.

the closest of which is only 3 cm from the beam, and the farthest is about 8 cm from the beam. Each layer has 12 “ladders”, arranged in a dodecagon. Figure 4-5 shows a diagram of an SVX ladder. Each ladder has three crystals, 8.5cm in length, which are connected together to form one large piece, 25.5cm long. The ladders have metal strips along them; the inner three layers have a strip pitch of 60 μm, and the outer layer has a pitch of 55 μm. This is nearly 200 times more dense than the spacing of sense wires in the CTC.

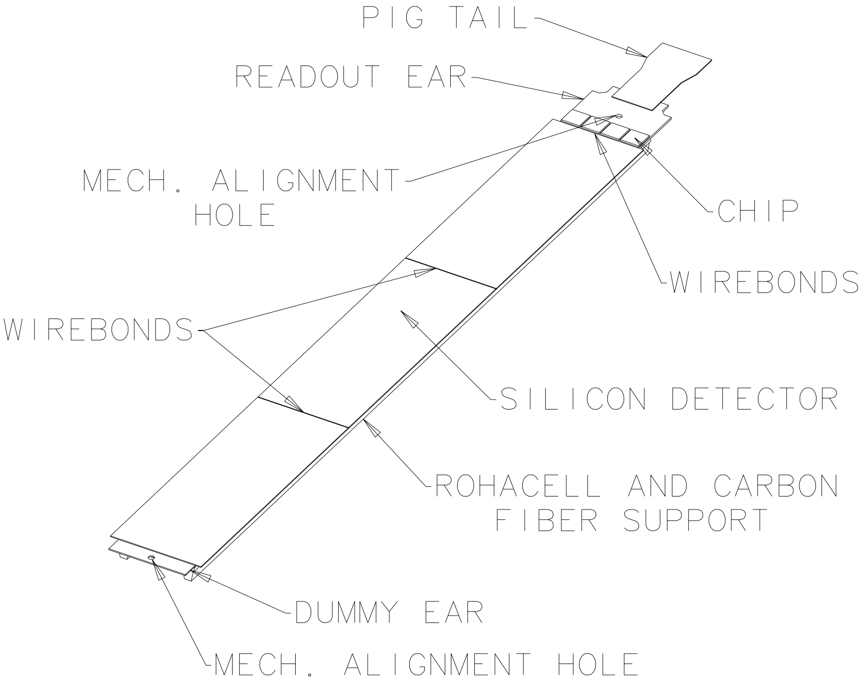

Figure 4-5: An SVX “ladder.”

Figure 4-6 shows an SVX “barrel”. The four dodecagons are aligned, creating 12 “wedges” in ϕ, each covered by four ladders. There are two barrels, one at positive η, and one at negative. They are separated by 2.15 cm at $z = 0$. Roughly 60% of the $p\overline{p}$ interactions at CDF fall within the SVX fiducial region.

The ladders are built of n-doped semiconductor, and under each metal strip is a p-doped layer. This forms an array of diodes, across which is held a steady voltage. A passing charged particle excites electrons into conduction energy bands, allowing current to flow. The position of a charged particle as it passes a given layer can be measured with accuracy of roughly 13 μm.

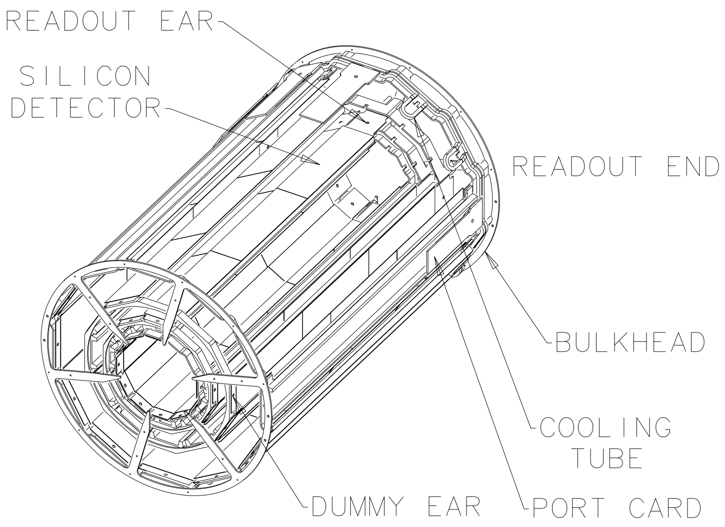

Figure 4-6: A schematic diagram of one of the two SVX barrels.

SVX information is combined with CTC information by taking a CTC track, extrapolating it through the SVX volume, and looking for nearby hits. This combined SVX-CTC track has an impact parameter resolution of $\delta d_0 \approx 13 + 40/p_T\,\mu\mathrm{m}$ (where p_T is the transverse momentum of the track, measured in GeV/c). This is over 4 times more accurate than the measurement with CTC information, alone.

Vertex Time Projection Chamber (VTX)

Since the proton and antiproton beams have significant physical length, not all interactions happen at $z = 0$; they are spread out, following a Gaussian distribution with width $\sigma_z \approx 27$ cm. The VTX detects the location of primary interactions, to improve the z_0 and $\cot\theta$ information for tracks found in the CTC.

The VTX is another gas-filled drift chamber like the CTC, except that its wires are oriented transverse to the beamline. It fills the volume between the SVX and the CTC, extending to $|z| = 1.5$m. Since it is closer to the beam than the CTC, it covers a much larger region of pseudorapidity: $|\eta| < 3.25$.

The VTX is segmented into 28 modules in z. The inner 16 surround the SVX and have 16 sense wires each. The rest extend inward to 7cm, and have 24 sense wires each. Each module is divided into two drift chambers by a high-voltage grid. Electrons released in the drift region drift away from the central grid, through a cathode grid, and into a proportional chamber, where they are detected. The sense wires are made of eight straight sections, forming an octagon, giving r–z information for the particles passing through the VTX. Consecutive modules are canted by 15° in ϕ, allowing some ϕ information to be collected.

Combining VTX information with CTC information in a way similar to the combination of CTC and SVX information was considered, but abandoned. The ambiguity is too high to resolve which VTX stubs belong with which CTC tracks. Partly, this is because the VTX can "see" very soft tracks which curve too much in the magnetic field to be reconstructed in the CTC.

Instead, the VTX is used to find the z locations of the proton-antiproton interactions (called *primary vertices*). The track stubs in the VTX are collected and extrapolated back to the beamline, and fits are performed to find the primary vertices. This is very effective, and finds primary vertices with accuracy $\delta z =$1–2 mm, depending on the number of tracks used. In this analysis, the B-mesons are constrained to have come from a found primary vertex, and this constraint improves the accuracy of the z_0 and $\cot\theta$ information for the tracks.

4.2.2 Other CDF Components

In addition to the tracking chambers which measure the momenta of charged particles, CDF has several other significant components. The electromagnetic calorimeters allow for the detection of electrons and photons. The hadronic calorimeters measure the energy of hadrons (both charged and neutral) produced in the interactions. The muon chambers identify tracks left by muons. The beam-beam counters identify which beam-crossings contain interactions and provide timing information for those events.

Calorimeters

Calorimeters are designed to measure the energy of particles. CDF has two types of calorimeters, *electromagnetic*, which measure electrons and photons, and *hadronic*, which measure hadrons. The *central* calorimeters are placed around the CTC and cover the region $|\eta| < 0.9$. The *plug*, *wall* and *forward* calorimeters cover regions of larger pseudorapidity, $1.1 < |\eta| < 4.2$ (see figure 4-2).

When electrons pass through matter, they lose energy very quickly through *bremsstrahlung* radiation.[3] Photons lose energy through *pair production*, which means they interact with the electric field of an atom and split into an e^-e^+ pair. These processes interrelate: photons split into e^+e^- pairs, which undergo *bremsstrahlung* radiation, emitting more photons, *etc.*. An incoming electron or photon will thus produce an *electromagnetic shower.* The electromagnetic calorimeters measure the energy of an electron or photon by counting how many electrons and positrons are produced in the electromagnetic shower.

Hadrons do not undergo significant *bremsstrahlung* radiation (especially the neutral ones), and thus generally pass through the electromagnetic calorimeters without depositing much energy. But they do interact strongly with the protons and neutrons in the nuclei of the atoms making up the detector. These interactions break up the atomic nuclei, throwing out more hadrons which also interact, snowballing into a shower of hadrons. The hadronic calorimeters contain 80 cm of steel and 30 cm of scintillator to collect and measure hadronic showers. This is roughly 4.5 interaction lengths, so most hadronic showers are entirely contained within the calorimeters.

For moderate-energy ($<\approx$ 50 GeV) charged hadrons, the CTC momentum measurement will be more accurate than the energy measurement from the hadron calorimeters. Since the hadrons used in this analysis all have energies far below 50 GeV, the

---

[3] As a charged particle passes near a nucleus, the electric field of the nucleus accelerates the particle, causing it to radiate. This process is called *bremsstrahlung* radiation, and the rate of energy loss is proportional to $1/m^2$, where m is the mass of the charged particle. This process is the dominant source of energy loss for electrons, but is not very important for other particles, which have $m^2 > 40,000\, m_e^2$.

calorimeter information is not used to improve the momentum measurement from the CTC. The muons used in this analysis will pass through calorimeters before reaching the muon chambers, so they are required to deposit some energy in the calorimeters. Except for this muon requirement, the calorimeter information is not used at all by this analysis.

Muon Chambers

Muons are roughly 200 times as massive as electrons, and therefore undergo $\approx 40,000$ times less *bremsstrahlung* radiation. They are not hadrons, and therefore do not interact strongly with atomic nuclei. Thus, muons can penetrate much more material than any other type of charged particle. This fact is often exploited by designers of particle detectors, who place chambers behind the calorimeters (or some other sort of shielding) to detect the muons which pass through them.

CDF has several chambers designed to detect muons. The Central Muon Chamber (CMU), the Central Muon Upgrade (CMP), and the Central Muons Extension (CMX) are all situated to detect muons which pass through the CTC and central calorimeters. The η-ϕ regions covered by these detectors are shown in figure 4-7. The CMU and CMP each consist of four layers of drift chambers (operating in limited-streamer mode), covering the region $|\eta| < 0.6$. To extend the coverage, four free-standing arches were constructed (the CMX), each with drift chambers and scintillation counters. The CMX covers much of the region $0.6 < |\eta| < 1.0$.

There are also Forward Muon Chambers (FMU), which detect muons produced with $|\eta| > 1$ and which pass through the plug and forward calorimeters. Because accurate momentum measurement is necessary for this analysis, and this information is only available for particles which pass through the CTC, muons detected in the FMU are not used.

Beam-Beam Counters (BBC)

The BBC consists of two sets of crossed scintillator planes, located 5.8m before and behind the center of the interaction region; they cover the region $3.2 < |\eta| < 5.9$.

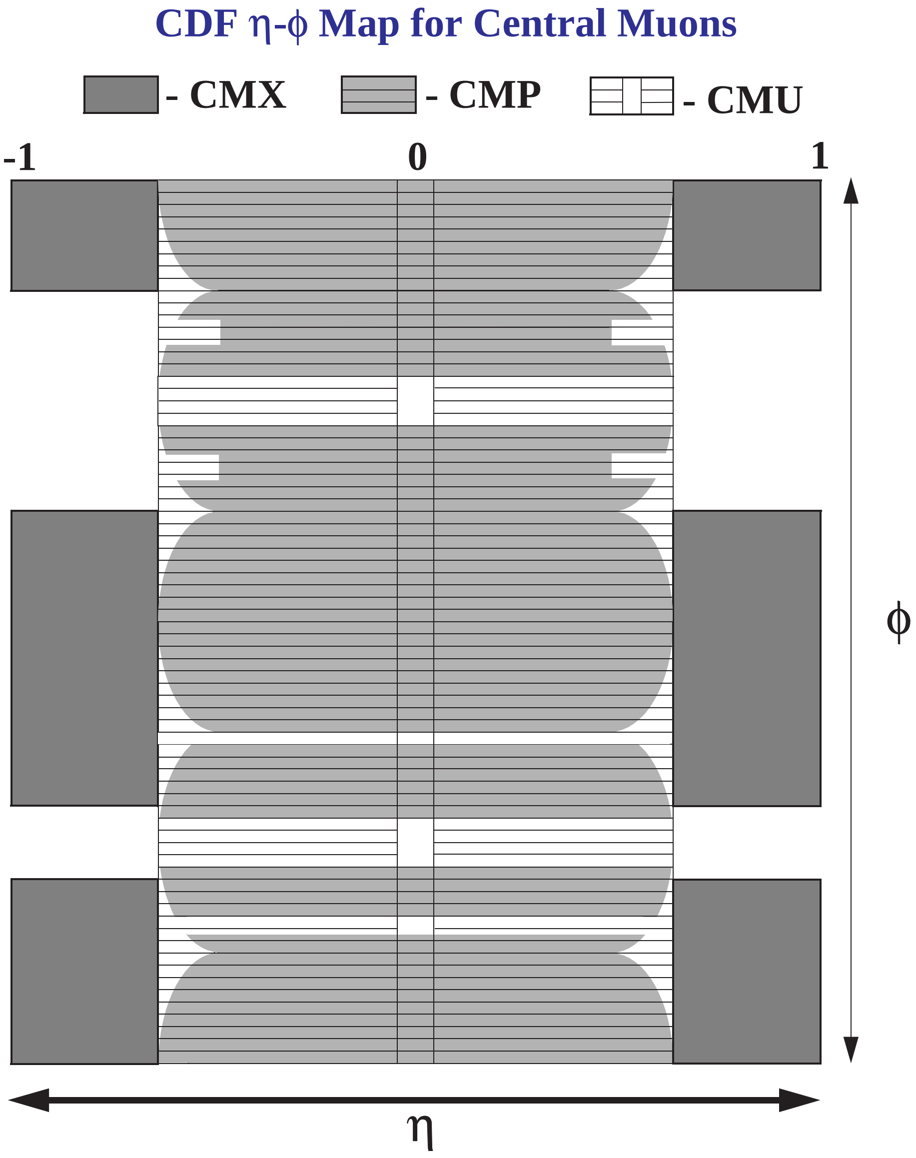

Figure 4-7: Regions of η–ϕ space covered by the CMU, CMP, and CMX.

When there is an interaction there will almost always be particles (often remnants from the original proton and antiproton) which pass through the beam-beam counters.

The BBC was originally designed to act as a first-level of triggering. But the luminosities during the Run I data-taking period (which provided the data used in this analysis) was high enough that the BBC has become almost saturated, decreasing its usefulness as a trigger. However, the BBC also provides a very accurate measurement of the time of the interaction (within 200 ps). Since the tracking chambers measure particle positions by measuring drift-times, the timing information from the BBC is very important.

The probability of there being no interactions in a beam-crossing, when μ is the average number of interactions per beam-crossing, is given by Poisson statistics:

$$P(0;\mu) \quad = \quad e^{-\mu} \tag{4.2}$$

Thus, by counting how often the BBC does not fire, one can use it to measure the luminosity, even when $\mu > 1$. Using the BBC, the integrated luminosity is measured to within an accuracy of $\approx 3.6\%$.

4.2.3 Triggers

CDF has a multilevel triggering system for deciding which events to keep and which to discard. This is for two reasons. The total $p\overline{p}$ cross-section at $\sqrt{s} = 1.8$ TeV is 75 mb. The cross-section of "interesting" physics is much lower. Most of the events would therefore be discarded as "uninteresting." The decision was made to do this immediately, without even recording the information from the events.

Another vital reason for having a trigger system is that it takes roughly 2 ms to read out the detector, while beam crossings occur every 3.5 μs. Thus, it would not be possible to read out the detector more often than once every $\approx$ 500 beamcrossings, no matter how interesting the events were. The trigger system allows the more interesting events to be kept while the less interesting ones are discarded.

Level 0

The level-0 trigger requires coincidence of the BBC. The luminosity was often so high that more than one interaction was expected per crossing. The BBC level-0 trigger is still used to filter out the events where absolutely nothing happens.

Level 1

The first real level of triggering is required to make its decision in less than the 3.5 μs between beam crossings. Analog electronics connected directly to the detector readouts perform simple algorithms using the raw data collected by each detector component. There are many different level-1 triggers, one looking for two muons, one looking or a single high-energy muon, one looking for high missing E_T, etc. The one used in this analysis requires there to be 2 or more muon stubs. If any level-1 trigger passes, the event is passed on to level-2. The combined rate of level-1 accepts is about 1–2 kHz, two orders of magnitude less than the input rate of 300 kHz.

Level 2

The level-2 trigger performs more complicated procedures and takes more time, typically 25–35 μs. This level is also implemented with hardwired processors, which is why it can work so quickly. When processing the level-2 trigger, the detector is blind to the 7-10 beam crossings which occur during that time. This loss of events is termed *dead-time*, and the dead-time for level 2 is the product of the level-1 accept rate by the level 2 processing time: about 2–5%.

The level-2 trigger uses the *Central Fast Tracker* (CFT). One of the advantages of the 45° tilt to the CTC wire-planes is that a high energy track will pass very close to at least one sense wire in each superlayer. The hits left on these wires are collected very quickly (because they have small drift distance), and are thus termed "prompt" hits. The CFT scans through the axial superlayers, finding all the prompt hits. It then examines the wires adjacent to the ones which have prompt hits, looking for "delayed" hits. Once the prompt and delayed hits are located, the CFT starts with the prompt hits on the outer superlayer (8), and, using a simple lookup table, extrapolates tracks inward to the other superlayers.

This algorithm is very fast. It also has the advantage that while several nearby particles might leave hits on CTC wires that might confuse detailed tracking, the prompt hit information will be unaffected, so tracks should not be lost due to nearby occupancy. It has the disadvantage that particles with transverse momentum below 2 GeV/c have too much curvature and the algorithm will not be able to find them. The CFT works within the 25 μs time available to the level-2 trigger, producing a list of tracks, with very rough momentum measurement.

The level-2 triggers used in this analysis require two muons stubs, at least one of which matches a CFT track. Due to the limited accuracy of the CFT, this amounts to approximately a 5° matching requirement.

Level 3

When level-2 accepts an event, the detector is read out, which makes the detector blind for 2 ms. But the level-2 accept-rate is only about 20 Hz, so this deadtime is only about 6%. The event data is read out by Frontend Readout Cards (FRCs), which pass their information on to Scanner CPUs (SCPUs).

There are six SCPUs, which each collect data from several FRCs and combine this data into one piece (*event fragment*), to be passed on to the level-3 event builders. From the SCPUs, the data passes through an Ultranet hub, which collects the six event fragments from the six SCPUs, and forwards them all to a single destination. The hub can handle several events simultaneously, forwarding separate events to separate destinations. The *event-builder* then takes the six fragments and formats them into one complete event, which it passes on to the level-3 processor.

The level-3 processors scan the events in more detail than is possible at level-2. All of the hits in the CTC are used for track reconstruction, but only tracks with $p_T > 2$ GeV/c are reconstructed, because the only tracks required are those used for the trigger decisions. The other detector components are similarly studied in more detail, to collect more information for the level-3 trigger. The level-3 trigger for this analysis requires that there be 2 muon stubs, and that each one has a fully reconstructed track which points to it (to within 5 standard deviations).

The level-3 rejection is only about a factor of 3, and level-3 processing takes roughly 1s. However, while level-3 is processing an event, the detector is live again, looking for more events. Level-3 deadtime only occurs when events are passed to it more quickly than it can process them, and the level-2 trigger is designed to prevent this from happening. Events are rejected at level-3 primarily to save storage space and CPU time in later reprocessing.

Level-3 also performs another important task: it sorts the events into *streams*. There are three streams, and the limited offline computers process events in order, according to which stream they are in. Events in stream A have the highest priority, followed by those in stream B, then those in stream C. This allows events of particular importance and interest to be studied quickly, while those of less interest are made available more slowly. Events which pass a two-muon level-3 trigger and have $\mu\mu$ invariant mass near the J/ψ mass are placed in stream A. These are the events used in this analysis.

4.3 Offline Reconstruction

After the data is collected and stored to tape, it is reprocessed using offline reconstruction code, which examines events more thoroughly than does the level-3 online code. Tracks are found in the CTC, with no minimum p_T requirement. SVX information is matched up with CTC tracks to form combined SVX-CTC tracks. Muon stubs are matched up with CTC tracks to form muon candidates. The VTX information is scanned for the locations of primary vertices. Other objects, like electron candidates and calorimeter jets, are also found, but they not used in this analysis.

4.3.1 Track Reconstruction

The track-finding algorithm starts by looking for hits in the outer layers of the CTC. These hits are combined into track stubs, which are then extrapolated inward toward the beamline. Hits are sought along these extrapolations, and any hits found are added to the track. This continues through to the inner layer of the CTC. The track-reconstruction efficiency depends on the p_T of the charged particle that left the track; tracks with $p_T > 400$ MeV/c have high reconstruction efficiency, and the efficiency drops with track p_T.

Tracks used in this analysis are required to have at least 2 hits in each of two axial and two stereo superlayers, to allow for accurate momentum measurement. They are also required to pass through the side barrel of the CTC, rather than through the endplate.[4] Tracks which do not reach the outer layers of the CTC have poorly understood reconstruction efficiency, which is why they are not used.

4.3.2 Muon Reconstruction

A muon must leave 3 or 4 hits in a muon chamber to form a muon stub. The χ^2 of fitting these hits to a straight line must be less than 10. The stub is then "matched" to a CTC track: the CTC track is extrapolated out to the radius of the muon chambers, and at that point, it must agree with both the location and direction of the muon stub in both r–ϕ and r–z. Each of these four quantities must match within 3 standard deviations. The muon purity and detection efficiency are both high for tracks with $p_T > 2$ GeV/c. The efficiency drops off below this because even muons with $p_T < 2$ GeV/c will often *range-out* in the calorimeters.

4.3.3 SVX Track Reconstruction

Each CTC track is examined to see if it extrapolates back through the fiducial region of the SVX. If it does, then the SVX is searched for hits consistent with being from that track. If hits are found on three of the four SVX layers, the SVX information is combined with the CTC information to form a combined SVX-CTC track. Otherwise,

---

[4] Particles with $p_T <\approx 250$ MeV/c will have too small a radius of curvature to reach the outer layer of the CTC, and will loop back through the CTC. Tracks from these particles (when found) are not used in this analysis.

any SVX information is discarded, and only the CTC information for that track is used.

4.3.4 Primary Vertex Location

Primary vertices are taken to lie at the center of the proton-antiproton beams. The beam-location is determined by combining the SVX and CTC information from tracks collected from many events over a several-hour running period. This "run-averaged" beam-position is stored in a database and is used later in offline processing.

The VTX is used to locate the z positions of the individual primary vertices. Several tracks from each vertex will pass through the VTX, which measures their r and z positions at several points. This information is then combined into stubs, which are extrapolated back to the beamline. A fit is then performed, using information from several stubs, to find the location of the primary vertex. In many events, there are more than one primary vertex, and the VTX information will usually separate them, locating both vertices.

Part III

Experimental Approach

Chapter 5

Method for Measuring $\sin 2\beta$

As explained in Chapter 2, $\sin 2\beta$ is the amplitude of the oscillation of the time-dependent asymmetry in the decays of B^0 and $\overline{B}^0$ mesons to the final state $J/\psi K^0_S$:

$$\frac{\Gamma(B^0(t) \to J/\psi K^0_S) - \Gamma(\overline{B}^0(t) \to J/\psi K^0_S)}{\Gamma(B^0(t) \to J/\psi K^0_S) + \Gamma(\overline{B}^0(t) \to J/\psi K^0_S)} = -\sin 2\beta \sin \Delta m t \qquad (5.1)$$

This chapter provides an overview of the experimental approach used to study this asymmetry and measure $\sin 2\beta$. It also serves as an introduction to the following chapters which describe the experimental method in more detail.

5.1 Identification of B mesons

The first step towards measuring $\sin 2\beta$ is the reconstruction and identification of B mesons decaying through the chain $B^0 \to J/\psi K^0_S$, with $J/\psi \to \mu^+\mu^-$ and $K^0_S \to \pi^+\pi^-$. As was explained in Chapter 3, requiring the J/ψ to decay to two muons reduces the backgrounds to manageable levels. The K^0_S is required to decay to two charged pions because charged pions are much more easily identified at CDF than are neutral pions.

All four particles are required to leave hits in the CTC, so that their momenta can be measured accurately. The two muons are also required to leave hits in the muon chambers (to allow them to be identified as muons) and in the SVX (to allow for accurate determination of the point of the B decay). Chapter 6 provides details of the event selection procedure.

5.2 Measurement of ct

Since the asymmetry to be measured depends on the proper decay time, this quantity must be accurately measured. A particle with momentum p and mass m which decays after a proper time t travels a distance $L = pt/m$ before decaying. The distance traveled transverse to the beamline, L_{XY}, is then

$$L_{XY} = L\sin\theta = \frac{p\,t}{m}\sin\theta = \frac{p_T\,t}{m}. \tag{5.2}$$

where θ is the polar angle of the particle's momentum, and p_T is its transverse momentum. Thus, by measuring a particle's L_{XY} and p_T (and knowing its mass), it is possible to calculate the proper time of its decay.[1]

To calculate L_{XY}, it is necessary to know the point of production and the point of decay of the particle. For the B mesons used in this study, the point of origin is taken to lie on the beamline, and the point of decay is calculated by extrapolating the tracks left by the muons back to their point of intersection (the tracks left by the pions from the K^0_S originate at the K^0_S decay point, which is displaced from the B decay point). The B momentum is determined by finding the total momentum of all the products from the B decay.

For convenience, the term *proper decay length* (ct) is used, and it refers to the proper decay time multiplied by the speed of light.

5.2.1 Uncertainty on ct

The uncertainty on the proper decay length (σ_{ct}) comes from three sources, namely the measurements of p_T, m, and L_{XY}. The uncertainties on p_T and m are typically

---

[1] Recall from Section 4.2.1 that the SVX only measures track information in the transverse plane. Thus, the measurement of L_{XY} will be considerably more accurate than the measurement of the distance traveled along the beamline, L_Z. This does not present a problem, as p_T is measured directly in the CTC, so L_Z is not needed.

less than 1%, so σ_{ct} is dominated by the uncertainty on L_{XY}. The typical uncertainty on L_{XY} is $\sim 100\mu\text{m}$, which translates into a typical decay length uncertainty of $\sigma_{ct} \simeq 50\,\mu\text{m}$ ($p_t/m \simeq 2$, on average). This corresponds to an uncertainty of ~ 0.17 ps on the decay time t, which is much smaller than either the B meson lifetime (~ 1.5 ps) or the period of the B^0–$\overline{B}^0$ mixing oscillation (~ 13 ps).

5.3 Flavor Tagging

Once the $B^0(\overline{B}^0) \rightarrow J/\psi K^0_S$ candidates are identified and their decay times are measured, it is then necessary to determine which ones originated as B^0 and which originated as $\overline{B}^0$. This requires *flavor tagging*, which is the subject of Chapter 7. Flavor tagging has been successfully used by many experiments to measure the B^0 mixing frequency, Δm. The method used by this analysis is referred to as *same-side tagging*, and it works by identifying certain pions which are produced along with the B meson. These pions exhibit a charge correlation with the flavor of B meson produced, and this correlation is used for the tagging.

5.4 Fitting for CP Asymmetry

To combine all this data and determine the value of $\sin 2\beta$ which is most consistent with it, an *unbinned maximum-likelihood fit* is used. Chapter 8 describes this fit, and Chapter 9 describes how the input parameters of the fit (and their uncertainties) are determined.

A simple *binned* fit would separate the events into bins in some variable (like ct), count how many events fall into each bin, and fit this distribution to a shape which depends on the value of of various parameters, like $\sin 2\beta$. One disadvantage of this approach is that it necessarily discards information about where the events are distributed within the bins. Another disadvantage is that it is difficult to generalize the fit to more than one or two variables, as the number of bins would grow too large.

An *unbinned* fit, however, represents the data as a collection of points in an n-

dimensional vector space, where n is the number of variables used. The fit then generates an n-dimensional distribution function based on its set of parameters, and calculates the combined probability of occurrence ("likelihood") of the collection of the set of points. It then varies the parameters to see what set of parameters has the maximum likelihood. This method has the advantages that it does not lose information from binning and that it easily generalizes to large n (for the fit in this analysis, $n = 7$). The main disadvantages are that unbinned fits are slow and their results are more difficult to visualize.

Visualization of the results of a fit is an important step in the interpretation of its results. If one were to fit a set of points to a straight line, it would be simple to plot the points, overlay the line, and interpret visually how well the fit worked. This process is not so simple with a multi-dimensional unbinned fit. The compromise made is that of making binned histograms over single variables and overlaying the curves that would have resulted from binned fits which had the same fit parameters as the unbinned fit. The plots in Chapter 10 are all of this type.

While this compromise does aid the reader in the interpretation of the results, it does not provide for detailed tests of the validity of the fitter. Those tests are performed using *Toy Monte Carlo* simulations, described in Appendix B.

5.4.1 Systematic Biases

As will be discussed more in Chapter 7, there is a systematic effect which biases this measurement of $\sin 2\beta$. The CDF detector has different efficiencies for reconstructing low-momentum positive and negative tracks. This affects the ratio of positive to negative charged particle tags, which in turn biases the measured value of $\sin 2\beta$. A correction for this bias has been incorporated into the likelihood fit (Section 8.2.3). The uncertainty on this correction leads to a systematic uncertainty on $\sin 2\beta$, as described in Chapter 11.

5.4.2 The value of $\sin 2\beta$

As will be described in Chapter 8, the amplitude of the experimentally measured asymmetry is not $\sin 2\beta$, but $\mathcal{D}\sin 2\beta$, where $\mathcal{D}$ is the *dilution factor* (described in Section 3.2.3). The dilution factor must be measured and divided out of the amplitude of the oscillation in order for $\sin 2\beta$ to be measured. This process is described in Chapter 12.

5.5 Control Samples

In order to help evaluate certain possible systematic biases, two other decays are considered: $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow J/\psi K^{*0}$ with $K^{*0} \rightarrow K^+\pi^-$. These two modes are kinematically very similar to $B^0 \rightarrow J/\psi K^0_S$, and should be equally vulnerable to any unforeseen effects which might bias the $\sin 2\beta$ measurement.[2]

One main difference between these two modes and $J/\psi K^0_S$ is that the flavor of the B meson at the time of decay can be determined by the flavor of the charged kaon: a K^+ would come from a B^+ or B^0, and a K^- would come from a B^- or $\overline{B}^0$. Since the final states are different for B^0 and $\overline{B}^0$, there can be no quantum interference as there is in $J/\psi K^0_S$, and therefore no CP violation.[3]

While there is no CP asymmetry to study in these modes, there is another asymmetry of interest. By comparing the decay flavor with the flavor predicted by the tagging, it is possible to use these modes to test the tagging algorithm. The time-dependence of this asymmetry will differ from that of $J/\psi K^0_S$. The $J/\psi K^+$ mode should exhibit no time-dependence (since B^+ mesons do not mix), and the $J/\psi K^{*0}$ asymmetry should vary as $\cos \Delta m t$ (as a result of B^0–$\overline{B}^0$ mixing). This is discussed (and verified) in Chapter 10.

---

[2] For simplicity, we will use the decay products of the B meson to identify the decay mode: $B^0 \rightarrow J/\psi K^0_S$ will be referred to as "$J/\psi K^0_S$", and the other two modes will be referred to as "$J/\psi K^+$" and "$J/\psi K^{*0}$".

[3] While it is possible that these modes could be susceptible to *direct* CP violation, such effects are expected to be negligible.

Chapter 6

Identification of B Mesons

This chapter describes the method used to identify the B mesons used in this analysis. It begins with a brief description of the topologies of the three decay modes reconstructed. Then it describes the reconstruction procedure and the selection criteria used to reject background. At the end are plots of the mass distributions for accepted $J/\psi K^+$, $J/\psi K^{*0}$, and $J/\psi K^0_S$ candidates.

6.1 B Meson Decay Topology

As described previously, B mesons decay via the weak interaction, with an average lifetime of ~ 1.5 ps. The B mesons used in this analysis typically have momentum of ~ 10 GeV/c, so they travel an average of ~ 1 mm before decaying. Figure 6-1 shows a diagram of the topology of the decay $B^+ \rightarrow J/\psi K^+$. The B^+ is produced at the *primary interaction point* (P.I.),[1] travels a small distance, and decays to $J/\psi K^+$. The J/ψ then decays immediately to $\mu^+\mu^-$ (J/ψ mesons, unlike B mesons, can decay via strong and electromagnetic interactions, so the J/ψ lifetime is much shorter than the B lifetime: $\tau_{J/\psi}/\tau_B \simeq 10^{-7}$).

Figure 6-2 shows a similar diagram for the decay $B^0 \rightarrow J/\psi K^{*0}$. The B^0 is

---

[1] In addition to B mesons, ***fragmentation particles*** are produced at the primary interaction point (Section 3.1.3). Some of these particles will be considered for ***same-side tagging*** candidates (Chapter 7).

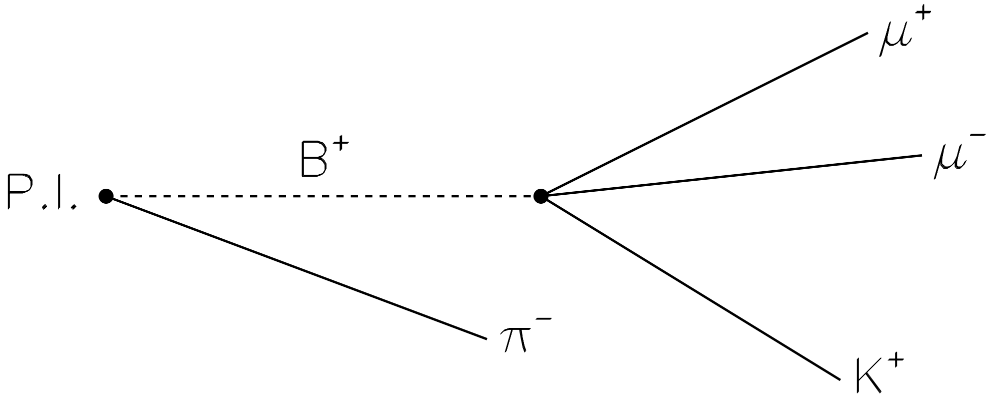

Figure 6-1: Topology of the decay $B^+ \rightarrow J/\psi K^+$. The additional track is from the primary vertex, and might be used for tagging.

produced at the primary interaction point, travels, and decays to $J/\psi K^{*0}$. The J/ψ then immediately decays to $\mu^+\mu^-$, and the K^{*0} immediately decays to $K^+\pi^-$ (K^{*0} mesons also decay via strong interactions, with an average lifetime roughly 600 times shorter than that of J/ψ mesons).

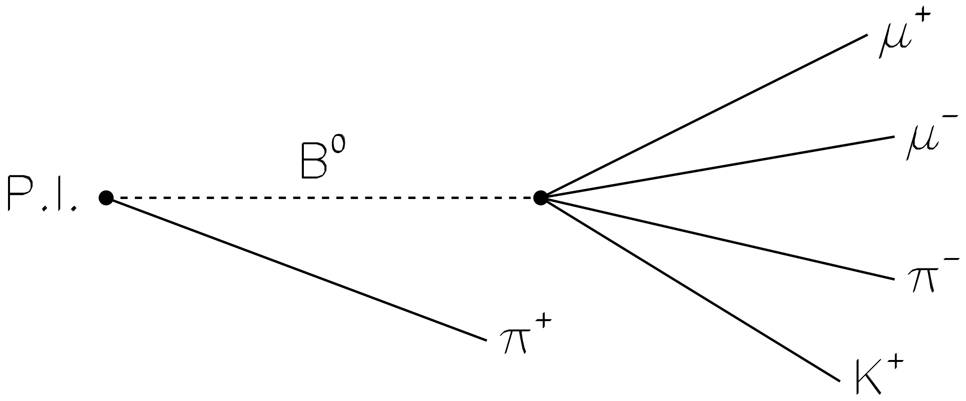

Figure 6-2: Topology of the decay $B^0 \rightarrow J/\psi K^{*0}$. The additional track is from the primary vertex, and might be used for tagging.

Figure 6-3 shows a similar diagram for $B^0 \rightarrow J/\psi K^0_S$. The K^0_S meson, unlike the J/ψ or K^{*0}, has a lifetime that is *longer* than that of the B meson ($\tau_{K^0_S}/\tau_B \simeq 60$). Thus, the K^0_S decay point will be well displaced from the B decay point.

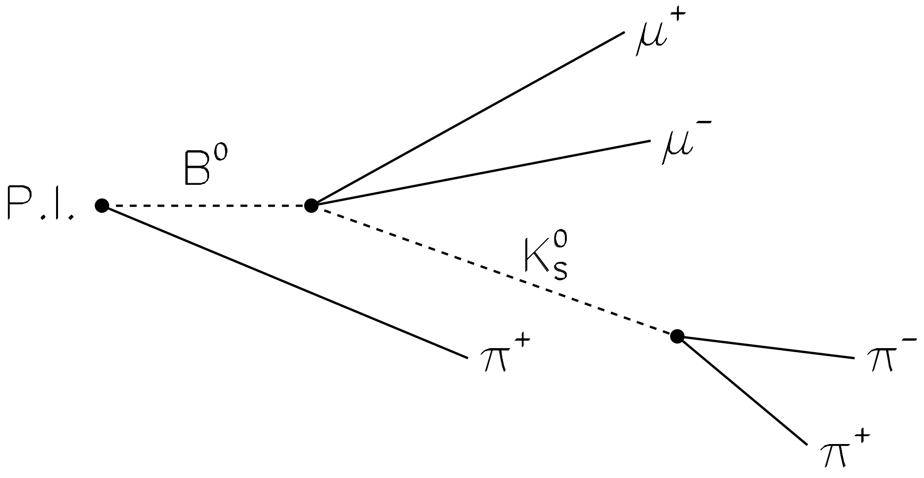

Figure 6-3: Topology of the decay $B^0 \rightarrow J/\psi K_S^0$. The additional track is from the primary vertex, and might be used for tagging. The diagram is not to scale: the K_S^0 lifetime is nearly 60 times larger than the B^0 lifetime.

6.2 *B* Meson Reconstruction

Reconstruction of these decays begins with the selection of the decay products: the muons, the pions, and the kaons. All these particles are required to pass through the CTC, so that their momenta can be measured accurately. The tracks from these particles are required to pass the selection criteria described in Section 4.3.1.

The muons are also required to leave hits in the muon chambers, so that they may be identified as muons. Muon candidates which do not pass the selection criteria described in Section 4.3.2 are rejected.

In order for the proper decay length (described in Section 5.2) to be measured accurately, at least two of the tracks from the B decay point are required to have valid SVX information (*i.e.* pass the selection criteria in Section 4.3.3). For the $J/\psi K_S^0$ decay, these two tracks are the two muon tracks, as the pion tracks from the K_S^0 do not pass through the B decay point (see figure 6-3). For the other two decays, any two of the tracks from the B decay can be used. Without this requirement, the average uncertainty on the B decay time would be a few hundred microns, rather than the 40 – 50 μm seen with this requirement.

6.3 Background Rejection

In addition to the properly reconstructed (*signal*) candidates, there are also *background candidates*. These are events which do not contain one of the decay modes sought, but instead contain tracks which *mimic* the decay. One method for rejecting background candidates is *track refitting*, which is described in Section 6.3.1. Another way that background is reduced is through the use of *kinematic selection criteria*, which is described in Section 6.4. Ultimately, some background will remain; these will be handled statistically, via sideband subtraction in the likelihood fit, as described in Chapter 8.

6.3.1 Track Refitting

The tracks in the signal candidates should intersect, as shown in figures 6-1 to 6-3. However, because the CDF detector has finite resolution, the reconstructed tracks will not always intersect. In addition, the reconstructed J/ψ, K^{*0}, K^0_S, and B meson masses will generally not be exactly correct, and the reconstructed B and K^0_S momenta will not always point directly from their respective points of origin to their points of decay.

The tracks in background candidates will often not intersect, but merely pass close to one another. In addition, the reconstructed J/ψ, K^{*0}, or K^0_S mass will be wrong for background candidates where those mesons are not actually present. In all the background events, at least one of the tracks from the reconstructed B decay will not be from the B meson,[2] so the B meson mass for the background candidates will be independent of the world-average value. By *refitting* the tracks, subject to the constraints that they intersect, that the reconstructed masses correspond to world-average values, and that the momenta of long-lived particles point directly from their

---

[2] The one exception to this is the "satellite-peak" background, discussed in Section 11.2.

points of origin to their points of decay, the amount of background can be reduced.

The offline track-reconstruction returns the 5 helix parameters for each track and also a 5×5 covariance matrix which expresses the uncertainties on and correlations between these parameters. These covariance matrices are used in a multi-track fit to determine a new set of track parameters where the above constraints are satisfied.

Track refitting is done according to the topologies displayed in figures 6-1 to 6-3. In all three modes, the two muon tracks are required to intersect (at the B decay point) and have combined mass equal to the world-average J/ψ mass (taken from [32]). For $J/\psi K^+$ and $J/\psi K^{*0}$, the other tracks are also required to pass through the B decay point, while, for $J/\psi K^0_S$, the two pion tracks are required to intersect at a separate point (the K^0_S decay point) and have combined mass equal to the world-average K^0_S mass.[3] The B momentum for each decay mode is required to point directly from the primary interaction point (found as described in Section 4.3.4) to the B decay point, and the K^0_S momentum is required to point directly from the B decay point to the K^0_S decay point.

The B meson mass is *not* fixed to the world average value in any of these fits. While doing so would reduce the amount of background , it would not eliminate all the background, and it would make *background subtraction* impossible. The B mass is left unconstrained in the fit, and it is used in Chapter 8 to differentiate between signal and background, statistically.

Track refitting helps reduce the background in two ways. First, the resolution of the raw B mass (without refitting) is roughly 60 MeV/c^2, while the resolution after fitting is roughly 14 MeV/c^2. Thus, refitting the tracks effectively removes 3/4 of the background under the $J/\psi K^0_S$ signal peak by making the peak 4 times narrower.

2-d χ^2 Cut

The second way that track refitting reduces the background is by providing a χ^2, which can be cut on. The χ^2 cut imposed uses only the r–ϕ information (C, d_0,

---

[3] The combined $K^+\pi^-$ mass is not required to match the world-average K^{*0} mass in the $J/\psi K^{*0}$ decay, as the K^{*0} has a natural width larger than the detector resolution (see Section 9.2).

and ϕ_0). The r–z information (z_0 and $\cot\theta$) is not used, as its ability to distinguish background from signal is very weak. This is because the typical resolution on d_0 is ~ 1 mm with CTC information alone, and $\sim 100\mu$m using combined SVX+CTC information, whereas the resolution on z_0 is ~ 1 cm either way. Since the B decay point is typically only displaced by ~ 1 mm from the primary interaction point, the z_0 information does not help distinguish signal candidates from background candidates.

This quantity is referred to as "$\chi^2_{2d\,\mathrm{tr}}$" as it is calculated using only the r–ϕ adjustments to the tracks. Figure 6-4 shows the value of $S^2/(S+B)$ as a function of the $\chi^2_{2d\,\mathrm{tr}}$ cutoff for each decay mode.[4] For all three modes, this increases for low $\chi^2_{2d\,\mathrm{tr}}$ (as more signal events are included), up to $\chi^2_{2d\,\mathrm{tr}} \simeq 20$, and then levels off. For larger values of $\chi^2_{2d\,\mathrm{tr}}$, this would drop down again, as more background is added with no appreciable increase in the amount of signal. Therefore 20 is chosen as the cutoff, and candidates with $\chi^2_{2d\,\mathrm{tr}} > 20$ are rejected. The value of $S^2/(S+B)$ is not strongly dependent on the exact value of the cutoff, so, for simplicity, the same cutoff is chosen for all three modes.

In addition to the tracks (and their covariance matrices), the location of the primary interaction point is used in the track refitting, to allow the B meson momentum to be constrained to point from the primary interaction point to the B decay point. The primary interaction point can also be adjusted in the fit (it has an associated covariance matrix). The quantity $\chi^2_{2d\,\mathrm{pi}}$ is the χ^2 of the displacement of the primary interaction point, in the r–ϕ plane. Figure 6-5 shows the dependence of $S^2/(S+B)$ on $\chi^2_{2d\,\mathrm{pi}}$ for the three decay modes. The three distributions level off after $\chi^2_{2d\,\mathrm{pi}} \simeq 4$, so candidates with $\chi^2_{2d\,\mathrm{pi}} > 4$ are rejected.

---

[4] $S^2/(S+B)$ is the figure of merit, as the statistical uncertainty on an asymmetry measurement is proportional to $1/\sqrt{S^2/(S+B)}$; maximizing $S^2/(S+B)$ minimizes the uncertainties on the asymmetry measurements.

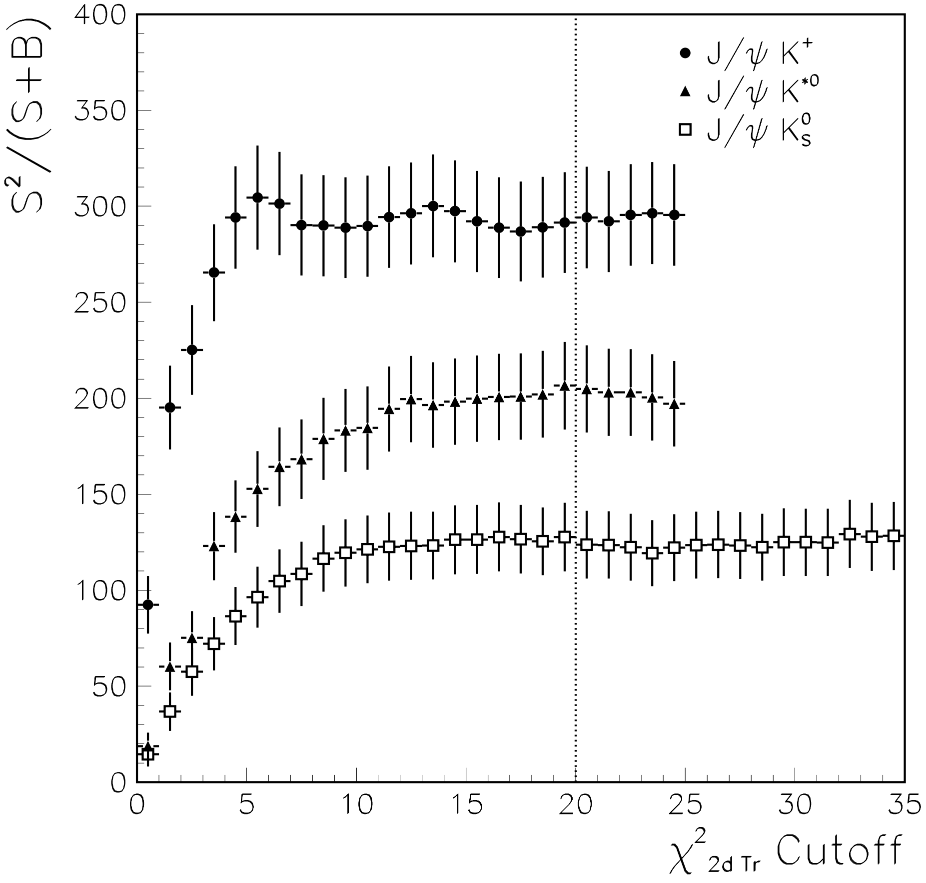

Figure 6-4: Dependence of $S^2/(S+B)$ on the $\chi^2_{2d\,\mathrm{tr}}$ cutoff. The vertical line shows the cutoff used in this analysis: candidates with $\chi^2_{2d\,\mathrm{tr}} > 20$ are rejected. This information was not collected for $J/\psi K^+$ or $J/\psi K^{*0}$ candidates for cutoffs above 25.

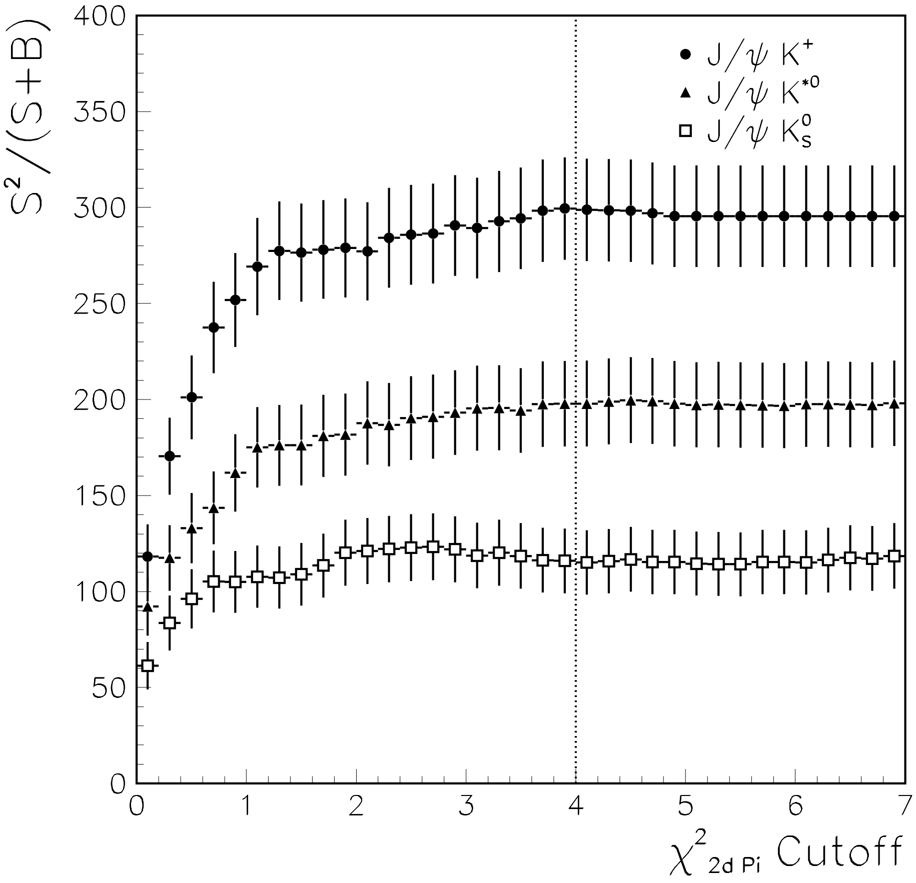

Figure 6-5: Dependence of $S^2/(S+B)$ on the $\chi^2_{2d\,\mathrm{pi}}$ cutoff. The vertical line shows the cutoff used in this analysis: candidates with $\chi^2_{2d\,\mathrm{tr}} > 4$ are rejected.

6.4 Kinematic Selection Criteria

The kinematic distributions of the signal and background events are different, so kinematic selection requirements are used to further reduce the background levels. Requirements are placed separately on the J/ψ, the K (K^*), and the B mesons.

6.4.1 J/ψ Selection

The first requirement for J/ψ selection is requiring that the event have passed the level-3 dimuon trigger (see Section 4.2.3). This reduces the amount of data to be scanned by a factor of roughly 3000, and the trigger is designed to be highly efficient for the decay $J/\psi \to \mu^+\mu^-$. This sample contains slightly more than 1 million events.

Roughly 40% of the events in this sample contain real J/ψ mesons; the rest are accidental combinations of real or fake muons. The muon selection criteria described above (Section 4.3.2) reduces the background somewhat, but 95% of these events contain two real muons. Not all these combinations of muons are consistent with being the decay products of J/ψ's, though. Requiring that the two tracks have z_0 within 5 cm of one another and have combined mass in the range 2.8 – 3.4 GeV/c^2 reduces the sample to roughly 750,000 events, with practically no loss of J/ψ signal.

Figure 6-6 shows two mass distributions for J/ψ candidates: the *raw* mass distribution calculated directly from the muon momenta (without refitting), and the *refit* mass distribution, where the two muon tracks are required to intersect. The raw mass resolution is typically 37 MeV/c^2, and the refit resolution is typically 31 MeV/c^2, about 20% smaller.

6.4.2 K and K^* Selection

The background events tend to have less energy than the signal events, so minimum values of p_T are required for the K and K^* mesons. The K^+ in the $J/\psi K^+$ decay is required to have $p_T > 1.75$GeV/c. The K^+ and the π^- in the $J/\psi K^{*0}$ decay are required to each have $p_T > 0.5$ GeV/c, and the K^{*0} is required to have $p_T > 3$ GeV/c. The K^0_S in the $J/\psi K^0_S$ decay is required to have $p_T > 0.7$ GeV/c. The p_T cutoffs are

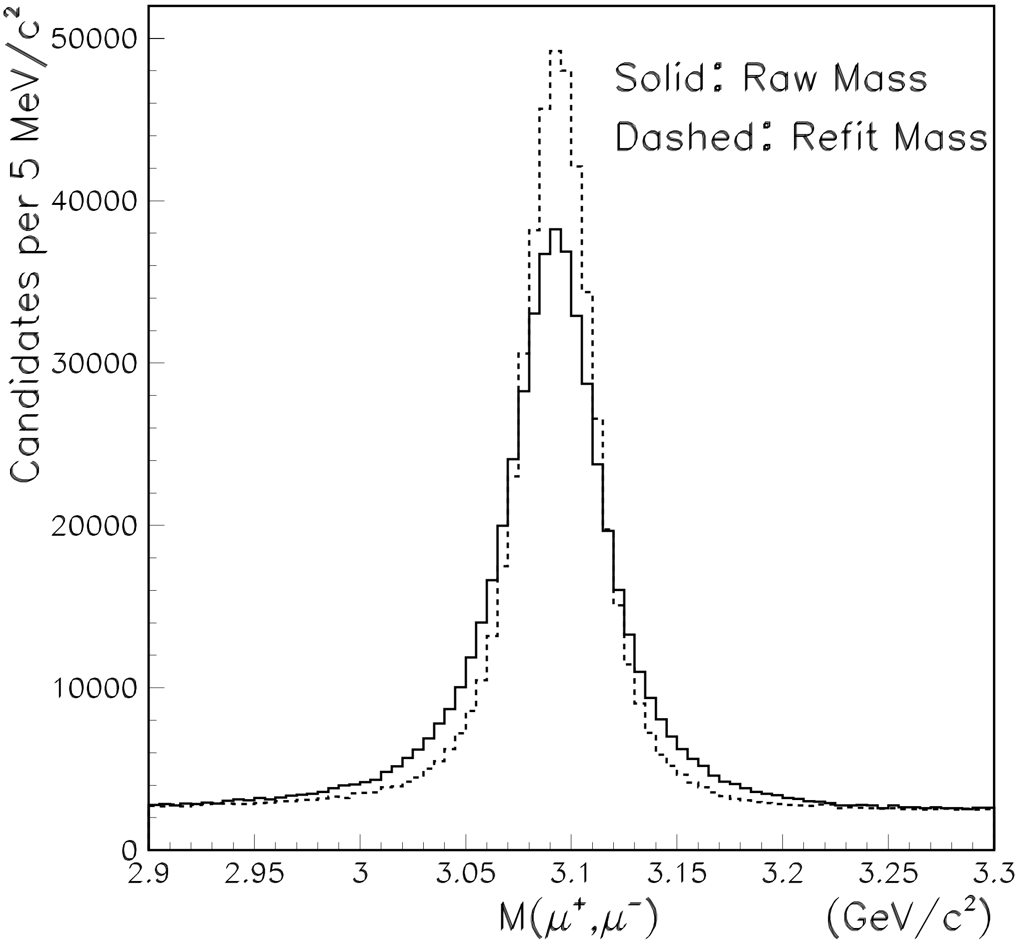

Figure 6-6: Raw and refit mass distributions for J/ψ candidates. There are roughly 430,000 signal events in each peak. Average mass resolution is 37 MeV$/c^2$ for the raw mass and 31 MeV$/c^2$ for the refit mass.

different for the three modes, as the amount of background differs between the modes; for $J/\psi K_S^0$, there is relatively little background, while for $J/\psi K^{*0}$, the background is much more prevalent.

In addition, the pion and kaon tracks for $J/\psi K^+$ and $J/\psi K^{*0}$ are required to have z_0 within 5 cm of the average z_0 of the two muons. This removes candidates where the tracks are from different primary interactions or are badly mis-measured. For the $J/\psi K_S^0$ candidates, this requirement is relaxed to 15 cm, as the K_S^0 mesons can travel 10 cm, or more, before decaying.

For $J/\psi K^{*0}$ candidates, the raw $K^+\pi^-$ mass is constrained to be in the range 700 – 1100 MeV/c^2, as the K^{*0} mass width is so large (see Section 9.2). The refit K^{*0} mass is required to be within 80 MeV/c^2 of the world-average K^{*0} mass (896 MeV/c^2).

The raw $\pi^+\pi^-$ mass for $J/\psi K_S^0$ candidates is required to be in the range 300 – 850 MeV/c^2. As the K_S^0 mass is constrained to the world-average value in the four-track $\mu^-\mu^+\pi^-\pi^+$ fit, it cannot be used for candidate selection. Instead, a two-track fit is performed, using only the two pion tracks, requiring that they intersect. The refit mass is required to be within 5 standard-deviations of the world-average value (498 MeV/c^2). The χ^2 of this fit is required to have probability of at least 0.1%. In addition, the K_S^0 decay length (the distance from the B decay point to the K_S^0 decay point, measured in the four-track fit) is required to be 5 standard-deviations positive. This removes the possibility of having background candidates where all four tracks originate from the same point.

6.4.3 B Selection

As mentioned above, the background events tend to have lower energy than the signal events, so candidates with B meson p_T below 4.5 GeV/c are rejected. Also, events with raw B mass outside the range 4.5 – 6.0 GeV/c^2 are rejected, before the tracks are even refit. After the fit is performed, only B meson candidates with reconstructed mass within 20 standard-deviations of the average value[5] are kept. This window is

---

[5] The average mass of the B mesons used in this sample is 5.277 GeV/c^2, which is slightly different from the world-average value, 5.279 GeV/c. As this could be due to a slight underestimation of the

very wide, to provide wide background sidebands, for use in sideband-subtraction.

To avoid correlations that could occur between multiple candidates in a single event, only one candidate is allowed per event. If more than one candidate is reconstructed, the one with the lowest $\chi^2_{2d\,\mathrm{tr}}$ is selected. Finally, any candidates with reconstructed proper decay length $ct < -2$ mm are considered badly mis-measured and discarded. This cut rejects 4 $J/\psi K^+$ candidates and 1 $J/\psi K^0_S$ candidate.

6.5 Mass Distributions for Accepted B Candidates

6.5.1 $J/\psi K^+$ Mass Distributions

Figure 6-7 shows the mass distributions for accepted $J/\psi K^+$ candidates, with $ct > 0$ and with $ct < 0$. The horizontal scale is the normalized mass $M_N = (M_{FIT} - M_0)/\sigma_{FIT}$, where M_0 is the central fit value, 5.277 GeV/c^2, M_{FIT} is the reconstructed mass, and σ_{FIT} is the uncertainty on the reconstructed mass. The value of M_N for the signal events is expected to follow a Gaussian distribution with mean 0 and width 1, while it should follow a linear distribution for the background. In fact, due to inaccuracies in the calculation of the covariance matrix, the width is slightly wider than 1 (this effect is taken into account via the fit parameter X, in Chapter 8). Superimposed over the histograms are curves which indicate the results of the fit (described in Chapter 10). There are 12,564 candidates in the sample, and the fit indicates that 846 ± 41 are signal.

6.5.2 $J/\psi K^{*0}$ Mass Distributions

The normalized mass distributions for the accepted $J/\psi K^{*0}$ candidates are shown in figure 6-8. As with $J/\psi K^+$, candidates with $ct > 0$ and with $ct < 0$ are shown separately, the horizontal scale is normalized mass, and curves which indicate the results of the fit are superimposed over the data histograms. There are 2339 candidates in

---

reconstructed track momentum, the value 5.277 is used throughout this analysis as the average B meson mass, instead of the world-average value.

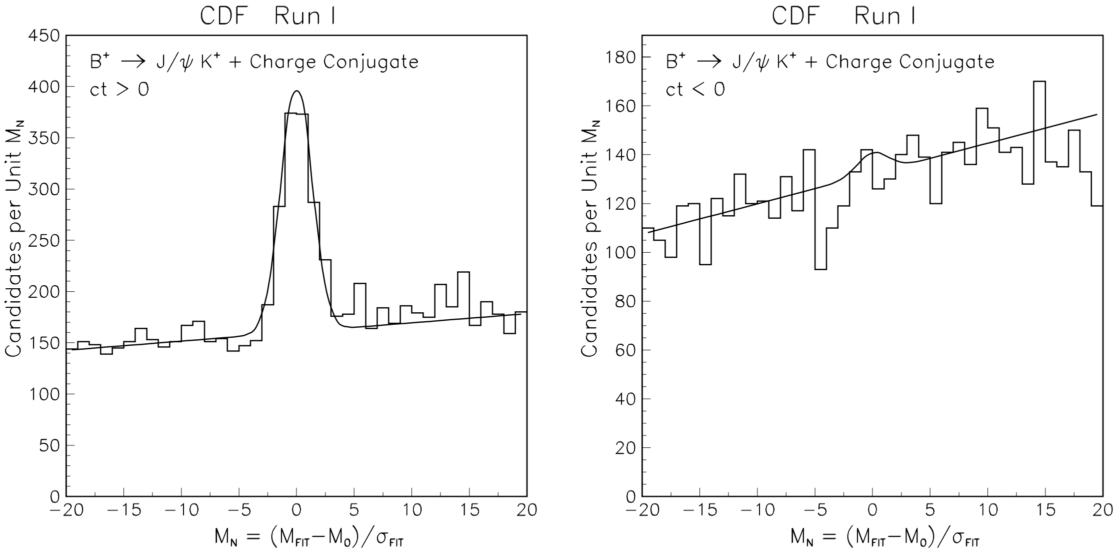

Figure 6-7: Normalized $J/\psi K^+$ mass distribution for selected events with $ct > 0$ (left), and $ct < 0$ (right). Superimposed over the histograms are the mass distributions from the likelihood fit. The fit indicates that 846 of the 12564 events in the sample are $J/\psi K^+$ signal.

the sample, and the fit indicates that 365 ± 22 are signal.

6.5.3 $J/\psi K^0_S$ Mass Distributions

The normalized mass distributions for the accepted $J/\psi K^0_S$ candidates are shown in figure 6-9. As with $J/\psi K^+$ and $J/\psi K^{*0}$, candidates with $ct > 0$ and with $ct < 0$ are shown separately, the horizontal scale is normalized mass, and curves which indicate the results of the fit are superimposed over the data histograms. There are 1696 candidates in the sample, and the fit indicates that 198 ± 17 are signal.

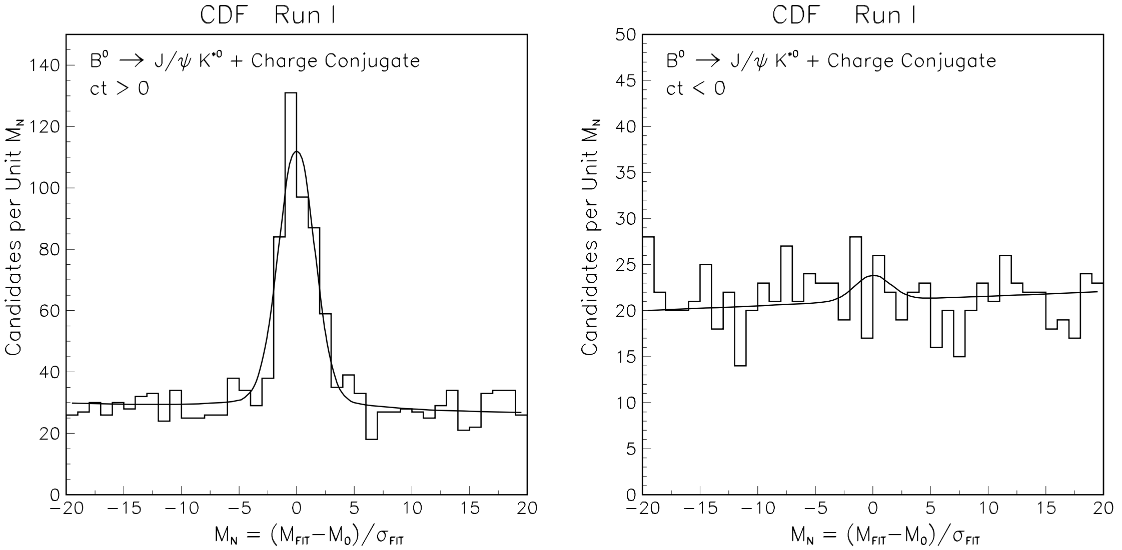

Figure 6-8: Normalized $J/\psi K^{*0}$ mass distribution for selected candidates with $ct > 0$ (left), and $ct < 0$ (right). Superimposed over the histograms are the predicted mass distributions from the fit. The fit indicates that 365 of the 2339 events in the sample are $J/\psi K^{*0}$ signal.

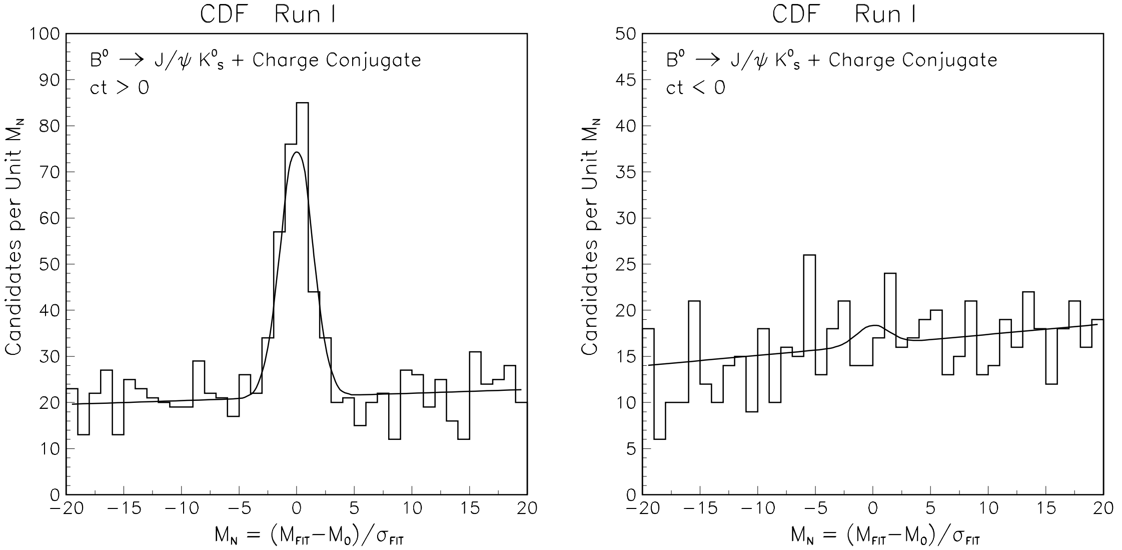

Figure 6-9: Normalized $J/\psi K^0_S$ mass distribution for selected candidates with $ct > 0$ (left) and $ct < 0$ (right). Superimposed over the histograms are the mass distributions from the likelihood fit. The fit indicates that 198 of the 1696 events in the samples are $J/\psi K^0_S$ signal.

Chapter 7

Flavor Tagging

Once the B mesons have been identified, the next step is to identify their flavors (*i.e.* whether they contain a b or a $\overline{b}$ quark) at production, a process called *flavor tagging.* The three most popular flavor tagging algorithms, *lepton tagging*, *jet-charge tagging*, and *same-side tagging* are described in Section 3.2.3. Same-side tagging (SST) is the algorithm used in this analysis.

This chapter describes the principle of same-side tagging, the specifics of the algorithm used, and the expected differences between tagging B^+ versus B^0 mesons.

7.1 Principle of Same-Side Tagging

In Ref. [52], Gronau *et al* explain how the decay products of an orbitally-excited B meson (B^{**}) can be used to tag the flavor of a B^0 meson: a B^{**+} can decay to $B^0\pi^+$, but not to $\overline{B}^0\pi^+$. If the B^{**} resonances are narrow, these decays can be cleanly identified (*eg.* by reconstructing a peak in the $J/\psi K^0_S\pi^+$ invariant mass distribution), and the charge of the pion can be used to tag the flavor of the B^0.

In [53], Gronau *et al.* further argued that it would still be possible to tag B^0 mesons with same-side tagging, even if the B^{**} resonances could not be cleanly identified. Particles produced in fragmentation should also show a correlation, and this correlation would be the same as that of the pions from the decays of B^{**}.

As a $\overline{b}$ quark *hadronizes* into a B meson (see Section 3.1.3), other *fragmentation* particles are produced along with the B meson. The closer two particles are in the "fragmentation chain," the closer their final momenta are expected to be. Figure 7-1 shows a depiction of one way a $\overline{b}$ quark might hadronize. At the final "string breaking," a $d\overline{d}$ pair "pops" out of the vacuum. The d quark combines with the $\overline{b}$ quark to form a B^0 meson, and the remaining $\overline{d}$ quark combines with a u quark from a $u\overline{u}$ pair to form a π^+. Thus, in this fragmentation chain, the particle nearest to the B^0 is a π^+.

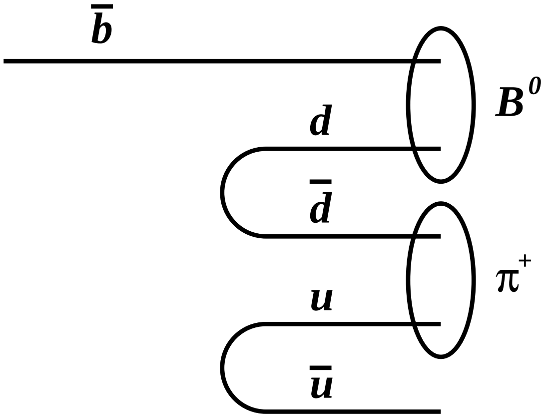

Figure 7-1: Correlation between the flavor of a B meson and pions produced in fragmentation.

In similar diagrams, the nearest particle could be a π^0, a ρ^+, a ρ^0, *etc.*, but never a π^-, as it must contain the $\overline{d}$ quark leftover from the $d\overline{d}$ pair. Thus, the charge of the nearest charged particle in the fragmentation chain will be correlated with the flavor of the B meson that was produced: a positive particle with a B^0 and a negative one with a $\overline{B}^0$. This correlation is the same as for the decay products of B^{**} mesons. Thus, a correlation should be seen in the particles produced near the B meson, whether the B^{**} decays are specifically reconstructed or not.

If the nearest particle in the fragmentation chain is a ρ^0, then it will decay to $\pi^+\pi^-$. Similarly, ω's, η's, and some other excited-state neutral mesons can produce charged pions in its decay. The above mesons decay via strong interactions, so it is not possible to distinguish the pions from these decay from the pions that give rise to the same-side tagging correlations, so any practical tagging algorithm will necessarily have a purity (fraction of tags with the correct correlation) of less than 100%.

7.2 The SST Algorithm

Same-side tagging has been used at CDF to measure the B^0 mixing frequency Δm [45] (shown in figure 3-6). For consistency, this analysis uses the identical tagging algorithm to that used in [45].

The events are scanned for tracks with momenta that point near the B meson momentum direction; any tracks with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.7$ are considered for use as tags. Because the tags are expected to come from the primary interaction point,[1] the tracks are required to be consistent with it: z_0 is required to be within 5 cm, and d_0 is required to be within 3 standard deviations of 0. The tracks are required to have valid SVX information (as in Section 4.3.3), to improve the discriminatory power of the impact-parameter requirement. Finally, the tracks are required to have $p_T > 0.4$ GeV/c, to reduce bias introduced by asymmetries inherent in the CDF detector performance (described below, in Section 7.2.1).

Roughly 60% of the B meson candidates have at least one track satisfying these criteria. If there is more than one track, the one "nearest" the B meson is chosen. The "nearest" track is defined to be the one with the smallest p_T^{rel}, which is the momentum of the track, transverse to the sum of the B momentum and the track momentum, as shown in figure 7-2.

7.2.1 Detector Bias

The efficiency of the CDF track-reconstruction is not the same for positive and negative tracks, because of the wire-plane tilt in the CTC. Low-momentum positive tracks travel nearly parallel to the wire planes near the outside of the CTC, leaving hits on more wires than an average track would. Low-momentum negative tracks travel nearly perpendicular to these wire-planes, leaving hits on fewer wires than

---

[1] B^{**} mesons decay via strong interactions, so their decay products originate at the primary interaction point. Fragmentation tracks obviously come from the primary interaction point.

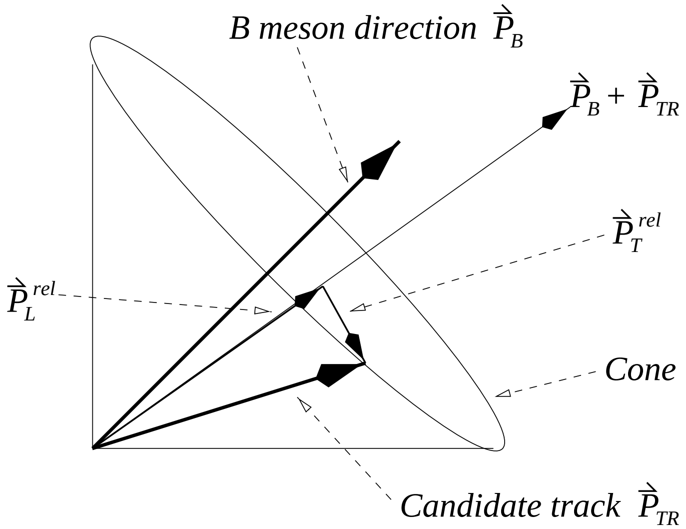

Figure 7-2: Schematic drawing of an SST tag and determination of p_T^{rel}.

average. This gives rise to an asymmetry in the track reconstruction efficiency for low-momentum positive and negative tracks.[2]

This asymmetry *increases* the dilution for B^0 mesons (where the correlation is with a π^+) and *decreases* that of $\overline{B}^0$ mesons (where the correlation is with a π^-). It also decreases the efficiency of tagging on $\overline{B}^0$ mesons, relative to that of tagging on B^0 mesons. Both of these effects directly bias the measurement of $\sin 2\beta$. The magnitude of the asymmetry is measured in an inclusive $B \rightarrow J/\psi$ sample, as will be described in Chapter 9. To correct for this bias, the likelihood fit incorporates separate dilutions and efficiencies for B^0 and $\overline{B}^0$ mesons, as will be described in Section 8.2.3.

---

[2] Another detector bias is present: *spallation.* A pion will occasionally collide with a nucleus in the detector, ejecting a proton. As the detector is made of *matter* which contains protons but not antiprotons, the particles ejected from these interactions will all be positive. The selection cuts used in this analysis (most particularly the impact-parameter requirement) reduce spallation to negligible levels.

7.3 Tagging Charged B Mesons

The SST dilution is not expected to be the same for B^+ mesons as it is for B^0 mesons. Figure 7-3 shows a diagram similar to that of figure 7-1, but for B^+ mesons. The pion produced with a B^+ is a π^-, opposite to that produced with a B^0, even though both B^+ and B^0 mesons contain $\overline{b}$ quarks.

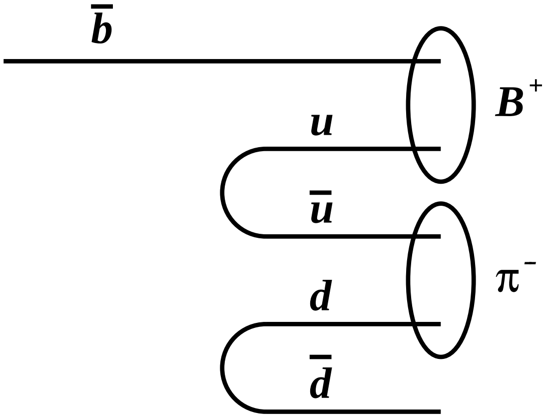

Figure 7-3: Correlation between the flavor of a B meson and pions produced in fragmentation.

But the tagging correlation is not exactly opposite, because not all tags are on pions; kaons and protons can also be used as tags. Figure 7-4 shows what would happen if, instead of a lighter pair of quarks, an $s\overline{s}$ pair "popped" out of the vacuum.[3] A B^+ would have a K^- produced along with it, which could be used as a tag as if it were a π^-. But a B^0 would have a $\overline{K}^0$, which could not be used as a tag.

Tagging on kaons increases both the efficiency and the dilution of tags for B^+ mesons relative to that of B^0 mesons. Similar diagrams with excited state mesons (K^*) even out the efficiency difference somewhat, but do not bring the dilutions closer together. Similar diagrams with protons and neutrons indicate that the effects of tagging on protons should be similar to those of tagging on kaons.

---

[3] Recall from Section 3.1.3 that the ratio of B_s^0 production to that of B^+ or B^0 production is roughly $0.11/0.38 \simeq 0.3$. The ratio of the rate of charged kaon production to that of charged pion production is expected to be roughly the same.

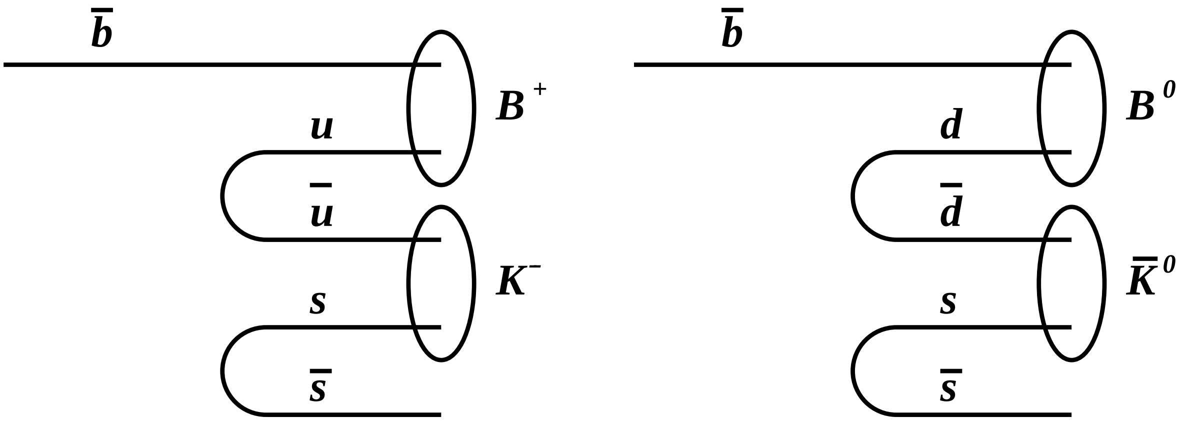

Figure 7-4: Correlation between the flavor of a B meson and kaons produced in fragmentation.

7.4 Dilution of Measured Asymmetries

As discussed in Section 3.2.2, the asymmetry in the number of direct versus mixed decays of B^0 mesons is $A(t) = \cos \Delta mt$ (equation 3.4).

If a sample of B^0 mesons is tagged with a tagging algorithm with dilution $\mathcal{D}_0$, then the fractions of correct (f_{RS}) and incorrect (f_{WS}) tags will be:

$$f_{RS} = \left(\frac{1+\mathcal{D}_0}{2}\right) \tag{7.1}$$

$$f_{WS} = \left(\frac{1-\mathcal{D}_0}{2}\right) \tag{7.2}$$

These equations come directly from equation 3.5 and the requirement that $f_{RS} + f_{WS} = 1$.

The collection of events that appear to be correctly tagged at time t will contain both correctly tagged unmixed events and incorrectly tagged mixed events. Thus, the probability of a tag at time t will appear to be correct is:

$$\begin{aligned} P_{RS}(t) &= \left[P(B^0 \rightarrow B^0)(t) f_{RS} + P(B^0 \rightarrow \overline{B}^0) f_{WS}\right] \\ &= \left[\left(\frac{1+A(t)}{2}\right)\left(\frac{1+\mathcal{D}_0}{2}\right) + \left(\frac{1-A(t)}{2}\right)\left(\frac{1-\mathcal{D}_0}{2}\right)\right] \\ &= \left(\frac{1+\mathcal{D}_0 \cos \Delta mt}{2}\right) \end{aligned} \tag{7.3}$$

and, similarly:

$$P_{WS}(t) = \left(\frac{1 - \mathcal{D}_0 \cos \Delta m t}{2}\right). \tag{7.4}$$

Thus the asymmetry in correct versus incorrect tags will be:

$$A'(t) = \left(\frac{P_{RW}(t) - P_{WS}(t)}{P_{RS}(t) + P_{WS}(t)}\right) = \mathcal{D}_0 \cos \Delta m t. \tag{7.5}$$

Thus, the original asymmetry is reduced by the asymmetry factor $\mathcal{D}_0$. This obviously does not depend on the type of the asymmetry inherent in $A(t)$, so the same effect would be expected in $J/\psi K^0_S$ in the measurement of $\sin 2\beta$. In this sample, the measured asymmetry should be:

$$A'(t) = \mathcal{D}_0 \sin 2\beta \sin \Delta m t. \tag{7.6}$$

7.5 Summary

Gronau *et al.* [52, 53] predict that there should be correlations between the produced flavors of neutral B mesons and the charges of nearby pions. These correlations can be exploited in a process called *same-side tagging* to tag the flavor of a neutral B meson when it was produced. This technique has been employed at CDF [45] to measure the B^0 mixing frequency, Δm.

This same tagging algorithm is used here, in an attempt to measure the CP asymmetry $\sin 2\beta$ in the decays of $B^0 \to J/\psi K^0_S$, where it is expected that (from equation 2.35):

$$\begin{aligned} A_{CP}(t) &= \frac{\Gamma(B^0(t) \to J/\psi K^0_S) - \Gamma(\overline{B}^0(t) \to J/\psi K^0_S)}{\Gamma(B^0(t) \to J/\psi K^0_S) + \Gamma(\overline{B}^0(t) \to J/\psi K^0_S)} \\ &= -\sin 2\beta \sin \Delta m t \end{aligned} \tag{7.7}$$

The amplitude of the measured asymmetry, in the number of events tagged as $\overline{B}^0$ versus the number tagged as B^0, will be reduced by the tagging dilution factor $\mathcal{D}_0$ (Section 7.4), so the measured asymmetry will be (equation 7.6)

$$A'(t) = \mathcal{D}_0 \sin 2\beta \sin \Delta m t. \tag{7.8}$$

In addition, two control samples, $B^+ \to J/\psi K^+$ and $B^+ \to J/\psi K^{*0}$, will be studied, as they samples should be equally vulnerable to any unforeseen effects which might bias the $\sin 2\beta$ measurement. In these samples, the asymmetries measured will be between the number of B mesons which are correctly tagged and the number which are incorrectly tagged. In the $J/\psi K^{*0}$ sample, the asymmetry should be proportional to $\cos \Delta m t$, due to B^0–$\overline{B}^0$ mixing, while in $J/\psi K^+$, the asymmetry should not depend on t, as B^+ mesons do not mix. The amplitudes of these asymmetries, $\mathcal{D}_+$ for $J/\psi K^+$ and $\mathcal{D}_0$ for $J/\psi K^{*0}$, reflect the dilution factors for tagging charged and neutral B mesons, and it is expected that we will have $\mathcal{D}_+ > \mathcal{D}_0$ (Section 7.3).

Once the B mesons are reconstructed and tagged, the next step is to fit them to determine $\mathcal{D}_0 \sin 2\beta$ in $J/\psi K^0_S$, $\mathcal{D}_+$ in $J/\psi K^+$, and $\mathcal{D}_0$ in $J/\psi K^{*0}$. Chapters 8 and 9 describe the fitting algorithm used in this analysis, and Chapters 10, 11, and 12 describe the results of these fits and the conclusions drawn from these results.

Chapter 8

The Unbinned Likelihood Fit

Once the B mesons are reconstructed and their flavor at $t = 0$ is determined, the next step is to determine what value of $\sin 2\beta$ is most consistent with the data. In order to make maximal use of the available statistics, an "unbinned maximum-likelihood" fit is used. This chapter describes the fit, the next chapter describes the inputs assumed for the fit, and the following chapter describes the results of the fit.

8.1 The Likelihood Fit

In this paper, the word *variable* is used to refer to a value measured for each event (*eg.:* mass), and the word *parameter* is used to refer to some quantity which applies to the whole sample (*eg.:* $\sin 2\beta$). The variables are listed in table 8.1 and the fit parameters are summarized in Section 8.6. Two of the variables, M_{FIT} and σ_{FIT}, are combined to form a new one: the *normalized mass* $M_N = (M_{FIT} - M_0)/\sigma_{FIT}$. As mentioned in the previous chapter, the signal events are expected to have M_N distributed as a Gaussian, with mean 0 and width 1.[1]

Each event represents a point in a seven-dimensional vector space:

$$(M_N, t, \sigma_t, r, s, p_T(tag), n_{PI}).$$

where table 8.1 presents the definitions of these variables.

---

[1] The actual width is slightly larger than 1, as will be discussed in Section 8.2.1.

Variable	Name
Reconstructed (fit) B meson mass	M_{FIT}
Uncertainty on B meson mass	σ_{FIT}
Reconstructed proper decay time	t
Uncertainty on proper decay time	σ_t
Reconstructed flavor of B meson	r
Sign of charge of tag	s
p_T of track used for tag	$p_T(tag)$
Number of primary interactions	n_{PI}

Table 8.1: Variables used by the maximum-likelihood fit.

A seven-dimensional distribution function $\mathcal{L}$ is used to model the data. The value of $\mathcal{L}$ is calculated for each event, and these values are combined to form the total likelihood:

$$\mathcal{L}_T = \prod^{N_{evt}} \mathcal{L}(M_N, t, \sigma_t, r, s, p_T(tag), n_{PI}) \qquad (8.1)$$

where N_{evt} is the number of events (including background events) in the sample. The shape of the likelihood function $\mathcal{L}$ is determined by the *fit parameters*. These parameters are varied until the set which maximizes $\mathcal{L}_T$ is found.

If there are n_c independent continuous variables used in the fit, then $\mathcal{L}_T$ will be of the order $\exp(-n_c\, N_{evt})$, which is a very small number ($\sim 10^{-2000}$ for $J/\psi K^0_S$). To avoid limitations inherent in computer technology, the quantity used is the logarithm of $\mathcal{L}_T$:

$$\ln \mathcal{L}_T = \sum^{N_{evt}} \ln \mathcal{L}(M_N, t, \sigma_t, r, s, p_T(tag), n_{PI}). \qquad (8.2)$$

This quantity is of the order $n_c\, N_{evt}$, and easily handled by a computer.

Using the logarithm has the additional advantage that the quantity $-2\ln\mathcal{L}_T$ behaves like a χ^2. In other words, if varying the fit parameter x from the best value x_0 to another value $x_0 + \delta_x$ increases $-2\ln\mathcal{L}_T$ by 1, then the *statistical uncertainty* on x

is equal to δ_x. The MINUIT fitting package [55] is used to find the set of parameters which minimize the quantity $-2\ln\mathcal{L}_T$ and to calculate the statistical uncertainties on these parameters.

8.1.1 Notation

For simplicity and clarity, the following notation is used:

- $E(t;\tau)$ is a normalized exponential distribution with "lifetime" τ,

$$E(t;\tau) \quad = \quad \frac{1}{\tau}\exp\{-t/\tau\}, \tag{8.3}$$

- $G(t;\mu,\sigma)$ is a normalized Gaussian with mean μ and standard deviation σ,

$$G(t;\mu,\sigma) \quad = \quad \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\}, \tag{8.4}$$

- $\otimes$ denotes convolution:

$$X(t,t')\otimes Y(t') \quad = \quad \int_0^\infty dt' X(t,t')Y(t'). \tag{8.5}$$

8.2 The Likelihood Function

For the purpose of the fit, the background events are separated into two types, prompt and long-lived. The prompt background events are those where the J/ψ candidate (real or fake) was produced directly in the primary interaction. These events should all have decay times consistent with 0. The rest of the background events are considered to be long-lived.[2]

The likelihood function $\mathcal{L}$ contains one term each for the signal, prompt background, and long-lived background:

$$\mathcal{L} = \sum^{\phi} f'_\phi \mathcal{L}_\phi \quad = \quad f_B\mathcal{L}_B + f'_P\mathcal{L}_P + f'_L\mathcal{L}_L, \tag{8.6}$$

---

[2] Some of the long-lived background will have decay times consistent with 0. They cannot be differentiated from the prompt background based on their decay times, alone. In some cases, they can be differentiated by their dependencies on other variables, but these events are generally only differentiated statistically.

where f_B is the fraction of events which are signal, f'_P is the fraction which are prompt background, and f'_L is the fraction which are long-lived background. These three parameters must sum to 1, and are therefore correlated. The two parameters f_B and $f_L \equiv f'_L/(f'_P + f'_L)$ are used because they can each vary independently over the range $[0, 1]$:

$$\mathcal{L} \;=\; f_B\mathcal{L}_B + (1 - f_B)(1 - f_L)\mathcal{L}_P + (1 - f_B)f_L\mathcal{L}_L. \tag{8.7}$$

Each term $\mathcal{L}_\phi$ can be decomposed into three factors:

$$\mathcal{L}_\phi \;=\; \mathcal{M}_\phi(M_N)\mathcal{T}_\phi(t, \sigma_t)\mathcal{F}_\phi(r, s, t, p_T(tag), n_{PI}), \tag{8.8}$$

where $\phi = B$, P, or L (*i.e.* B signal, prompt background, or long-lived background, respectively). The first factor $\mathcal{M}$ describes the mass dependence, the second factor $\mathcal{T}$ describes the decay-time dependence, and the third factor $\mathcal{F}$ describes the flavor of the meson. These factors are different for each type of event.

8.2.1 The Mass Factor, $\mathcal{M}$

Signal events all have true B mass equal to the correct value, so the distribution of M_N for these events should be a Gaussian of width 1 and mean 0. The actual width is somewhat larger than 1, because the covariance matrix calculated for the tracks is only an approximation (see figure 6-7). Therefore, the signal is modeled as a Gaussian of width X and mean 0, where X is a parameter, left free to float in the fit. The typical value of X is ~ 1.3, as will be discussed in Chapter 10.

Events with $|M_N| < W = 20$ are included in the fit to give wide sidebands to help characterize the backgrounds. The background events are expected not to show any mass dependence except for phase-space considerations. They are therefore modeled as having a linear distribution in M_N with fractional slope ζ. Because the two types of background come from different sources (*i.e.* the track combinatorics are different), their mass-dependencies are not assumed to have the same slope.

Thus, the mass-dependencies are:

$$\mathcal{M}_P(M, \sigma_M) \;=\; \left(\frac{1 + \zeta_P M_N}{2W}\right) \tag{8.9}$$

$$\mathcal{M}_L(M,\sigma_M) = \left(\frac{1+\zeta_L M_N}{2W}\right) \tag{8.10}$$

$$\mathcal{M}_B(M,\sigma_M) = G(M_N;0,X). \tag{8.11}$$

8.2.2 The Decay-Time Factor, $\mathcal{T}$

The probability for a B meson to decay at proper time t' is $E(t';\tau_B)$, where τ_B is the B meson lifetime. The reconstructed decay time t will differ from t' by a Gaussian smearing, caused by the detector resolution. The width of this smearing should be σ_t, the calculated uncertainty on t (Section 5.2.1). A scale factor Y is included to allow for the possibility that the uncertainty could be systematically under/over estimated.

Reference [44] presents a measurement of the B meson lifetimes using fully-reconstructed $B \to J/\psi K$ events at CDF. The events used in that analysis have large overlap with the ones used here, so the background lifetime distributions are modeled here in the same way as in [44]. The prompt background events all have $t' = 0$ (by definition), but t is smeared by the lifetime uncertainty, as for the signal, above. The long-lived background events are modeled with the following distribution for t':

$$T_L(t';f_N,f_{\tau 2},\tau_1,\tau_2) = f_N E(-t';\tau_2) + (1-f_N)\left\{f_{\tau 2}E(t';\tau_2) + (1-f_{\tau 2})E(t';\tau_1)\right\} \tag{8.12}$$

This distribution is modeled with three exponentials ("tails"), two with positive "lifetimes," and one with a negative "lifetime." The "lifetime" of the negative tail (τ_2) is fixed to that of the shorter-lived positive tail. Two fractions, f_N and $f_{\tau 2}$ define the relative fractions of long-lived background events in the three tails. The reconstructed decay time t is expected to be smeared by the detector resolution, as for the signal and the prompt background.

The three decay-time distributions are:

$$\mathcal{T}_P(t,\sigma_t) = G(t;0,Y\sigma_t) \tag{8.13}$$

$$\mathcal{T}_L(t,\sigma_t) = G(t;t',Y\sigma_t) \otimes T_L(t';f_N,f_{\tau 2},\tau_1,\tau_2) \tag{8.14}$$

$$\mathcal{T}_B(t,\sigma_t) = G(t;t',Y\sigma_t) \otimes E(t';\tau_B). \tag{8.15}$$

The background "lifetimes" τ_1 and τ_2 depend upon the specifics of reconstruction and are therefore allowed to be different for the three decay modes ($J/\psi K^+$, $J/\psi K^{*0}$, and $J/\psi K^0_S$) used.

Both the $J/\psi K^0_S$ and $J/\psi K^{*0}$ signal events have a time-dependent asymmetry. The convolution above includes the time-dependence of the asymmetry factor, as well.

8.2.3 The Asymmetry Factor, $\mathcal{F}$

The asymmetry factor is a combination of three factors:

$$\mathcal{F}(r,s,t') = \sum_p \mathcal{P}_1(r)\mathcal{P}_2(p|r,t')\mathcal{P}_3(s|p,p_T(tag),n_{PI}) \tag{8.16}$$

$\mathcal{P}_1$ is the *reconstruction* factor, and is included to handle asymmetries in the number of reconstructed B or $\overline{B}$ mesons. $\mathcal{P}_2$ is the *matching* factor, and is included to model B^0 mixing in $J/\psi K^{*0}$ and CP violation in $J/\psi K^0_S$. $\mathcal{P}_3$ is the *tagging* factor, and is included to model the tagging dilution factor and any inherent biases in the tagging algorithm. These three terms are described in more detail, below.

Asymmetries

Each of these factors expresses the probability that a variable a will take on a certain value (usually ±1), given an asymmetry A. If the number of events with $a = +1$ is $N(+1)$ and the number with $a = -1$ is $N(-1)$, then the asymmetry between these numbers is:

$$A \equiv \left(\frac{N(+1) - N(-1)}{N(+1) + N(-1)}\right). \tag{8.17}$$

The probability that an event will have $a = \pm 1$ is then:[3]

$$P(a = \pm 1) \equiv \left(\frac{N(\pm 1)}{N(+1) + N(-1)}\right) = \left(\frac{1 \pm A}{2}\right) = \left(\frac{1 + aA}{2}\right). \tag{8.18}$$

Equation 8.17 is equivalent to equation 8.18. This equivalence will be used implicitly throughout this chapter.

---

[3] This equation assumes that $+1$ and -1 are the only possible values for a. For the tagging factor, 0 is also a possible value, and the probabilities therefore include the efficiency, $\epsilon \equiv 1 - P(0)$, in addition to the asymmetry A.

The Reconstruction Factor, $\mathcal{P}_1$

The variable r represents the *reconstructed flavor*. It follows the charge of the b quark in the reconstructed B meson: $r = +1$ for B^0 and B^+, and $r = -1$ for $\overline{B}^0$ and B^-. The decay flavor has no meaning for the $J/\psi K^0_S$ sample, so all events in that sample are defined to have $r = +1$.[4]

The fraction of events reconstructed with flavor r is:

$$\mathcal{P}_1(r) \quad = \quad \frac{1 + rR}{2} \tag{8.19}$$

where R is the *reconstruction asymmetry*. R may represent an actual detector bias or merely a statistical fluctuation in the numbers of each flavor reconstructed. In either case, R is the asymmetry in the number of B^0 and $\overline{B}^0$ (or B^+ and B^-) that are reconstructed. Including this term allows other asymmetries to be calculated without the bias due to having different numbers of B and $\overline{B}$ mesons in the sample.

For $J/\psi K^0_S$, it is defined that $r = +1$, so $\mathcal{P}_1(+1) = 1$ and $\mathcal{P}_1(-1) = 0$, which implies that $R = 1$. In the other two samples, R_B, R_P, and R_L are fit parameters.

The Matching Factor, $\mathcal{P}_2$

The next variable is p, the *produced flavor*, *i.e.* the flavor of B meson that was originally produced. The convention is the same as for r: $p = +1$ for B^+ and B^0, and $p = -1$ for B^- and $\overline{B}^0$. The produced flavor is not actually known, so the asymmetry factor is the sum of the $p = +1$ and $p = -1$ terms.

The probability that the reconstructed flavor r matches the produced flavor p is:

$$\mathcal{P}_2(p|r,t) \quad = \quad \frac{1 + rpP(t')}{2} \tag{8.20}$$

where $P(t')$ is the *matching asymmetry*, and may depend on the decay time.

B^+ mesons do not mix, so $P = 1$ for that signal. Similarly, the "produced flavor" is not clearly defined for the background, so $P = 1$ for all backgrounds. B^0 mesons do mix, so $P(t') = \cos\Delta m t'$ for the $J/\psi K^{*0}$ signal. The partial width asymmetry

---

[4] This is not an arbitrary choice. Setting $r = +1$ for this entire sample allows the CP asymmetry to be included without the addition of special terms, as will be explained below.

in $J/\psi K_S^0$ can, due to the definition of $r = +1$ for this sample, be parameterized by setting $P(t') = -\sin 2\beta \sin \Delta m t'$.

Tagging Factor, $\mathcal{P}_3$

The *tagging factor* models the asymmetries related to the tagging. The variable s is used to denote the "sign" of the tag: $s = +1$ for tags on positive tracks, $s = -1$ for tags on negative tracks, and $s = 0$ when no track is available. In the absence of detector bias, the probability that, for a B meson of flavor p, the tag returned will be s is:

$$\begin{aligned} \mathcal{P}_3(s|p) &= \epsilon\left(\frac{1+sp\mathcal{D}}{2}\right) && (s \neq 0) \\ &= (1-\epsilon) && (s = 0) \end{aligned} \tag{8.21}$$

The parameters ϵ and $\mathcal{D}$ are the efficiency and dilution of the tagging algorithm. For B^+ mesons, the correlation is expected to be opposite, so the parameter $\mathcal{D}$ is negative for that mode.

More generally, the probabilities $\mathcal{P}_3(s|p)$ are independent, except that $\sum_s \mathcal{P}_3(s|p) \equiv 1$ for each p. This leaves four free parameters. The following four parameters are the ones used in this analysis:

$$\epsilon = \frac{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}{2} \tag{8.22}$$

$$\mathcal{D} = \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)} \tag{8.23}$$

$$\alpha = \frac{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)} \tag{8.24}$$

$$\gamma = \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)} \tag{8.25}$$

With no detector bias, α and γ will be zero, and ϵ and $\mathcal{D}$ will be the tagging efficiency and dilution, as above. If there is any charge bias, ϵ will represent the average tagging efficiency, $\mathcal{D}$ the average dilution, α the charge bias in favor of positive tags, and γ the efficiency bias in favor of B^0 or B^+ mesons.

These four quantities are chosen because they can be measured independently. The average tagging efficiency ϵ is measured directly in each $J/\psi K$ sample. The average dilution $\mathcal{D}$ is measured directly in $J/\psi K^+$ and $J/\psi K^{*0}$; for $J/\psi K^0_S$, the value of $\mathcal{D}$ used is extrapolated from that in other decay modes. The tagging charge bias α is measured in an inclusive $B \to J/\psi X$ sample, as will be described in Section 9.3.1. The parameter γ is not actually measured, but is constrained, as will be described in Section 9.3.2. The uncertainties on the values of α and γ are included in the systematic uncertainty, as will be described in Chapter 11.

The separate B and $\overline{B}$ dilutions and efficiencies are related to these four quantities as follows:

$$\begin{aligned}
\epsilon(B) &\equiv \mathcal{P}_3(+|+) + \mathcal{P}_3(-|+) = \epsilon(1+\mathcal{D}\gamma) && (8.26)\\
\mathcal{D}(B) &\equiv \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(-|+)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(-|+)} = \frac{(\mathcal{D}+\alpha)}{(1+\mathcal{D}\gamma)} && (8.27)\\
\epsilon(\overline{B}) &\equiv \mathcal{P}_3(-|-) + \mathcal{P}_3(+|-) = \epsilon(1-\mathcal{D}\gamma) && (8.28)\\
\mathcal{D}(\overline{B}) &\equiv \frac{\mathcal{P}_3(-|-) - \mathcal{P}_3(+|-)}{\mathcal{P}_3(-|-) + \mathcal{P}_3(+|-)} = \frac{(\mathcal{D}-\alpha)}{(1-\mathcal{D}\gamma)} && (8.29)
\end{aligned}$$

As expected, if $\gamma = \alpha = 0$, then $\epsilon(B) = \epsilon(\overline{B}) = \epsilon$ and $\mathcal{D}(B) = \mathcal{D}(\overline{B}) = \mathcal{D}$.

Equations 8.22 – 8.25 can be inverted to arrive at the $\mathcal{P}_3(s|p)$ in terms of ϵ, $\mathcal{D}$, α, and γ:

$$\begin{aligned}
\mathcal{P}_3(s|p) &= \frac{\epsilon}{2}\left(1 + s\alpha + sp\mathcal{D}(1+s\gamma)\right) && (s \neq 0)\\
&= 1 - \epsilon\left(1 + p\mathcal{D}\gamma\right) && (s = 0) \qquad (8.30)
\end{aligned}$$

Two additional definitions are introduced:

$$\begin{aligned}
\kappa(s;\alpha,\gamma,\epsilon) &= s\left(\frac{1+s\gamma}{1+s\alpha}\right) && (s \neq 0)\\
&= -\frac{\gamma\epsilon}{1-\epsilon} && (s = 0) \qquad (8.31)\\
\mathcal{E}(s;\alpha,\epsilon) &= \epsilon(1+s\alpha) && (s \neq 0)\\
&= 2(1-\epsilon) && (s = 0) \qquad (8.32)
\end{aligned}$$

which allow $\mathcal{P}_3$ to be cast in the more intuitive form (similar to equation 8.21):

$$\mathcal{P}_3(s \mid p;\alpha,\mathcal{D},\gamma,\epsilon) = \mathcal{E}(s)\left(\frac{1+\kappa(s)p\mathcal{D}}{2}\right). \qquad (8.33)$$

The factor $\kappa(s)$ is called the *asymmetry-corrected tag*, and $\mathcal{E}(s)$ is called the *efficiency factor*.

Tagging Charge Bias

As will be discussed in Section 9.3.1, the tagging charge bias α is measured in inclusive $B \to J/\psi X$ decays. It is found to exhibit significant dependence on only two variables: $p_T(tag)$ (the p_T of the track used for the tag) and n_{PI} (the number of reconstructed primary interaction points). The dependence is:

$$\alpha(p_T(tag), n_{PI}) \;=\; [a_1(n_{PI} - n_0) + b_1](p_T(tag)^{-4} - p_{T0}^{-4}) + [a_2(n_{PI} - n_0) + b_2] \quad (8.34)$$

The quantities a_1, b_1, a_2, b_2, n_0, and p_{T0} are all provided in Section 9.3.1.

This value of α is used for the signal events in all three modes. Because the background events are not fully-reconstructed B mesons, it is not necessarily expected that the asymmetries will be the same for them as for the signal events. Therefore, *the parameter α for the background is allowed to float freely in the fit.*

Tagging Efficiency Bias

The determination of the tagging efficiency bias is discussed in Section 9.3.2, where it is shown that this value (for signal events) should be in the range $0 < \gamma < 2.5$. For the background, this parameter would be allowed to float freely in the fit, except that it is too highly correlated with the dilution factor, $\mathcal{D}$. Instead, the parameter $\delta \equiv \mathcal{D}\gamma$ is used, as it is not highly correlated with $\mathcal{D}$. Thus, for the signal events, γ is constrained, and for the background, δ is allowed to float freely.[5]

The Combined Tagging Factor, $\mathcal{F}$

Combining the three tagging factors yields:

$$\begin{aligned}
\mathcal{F}(r, s, t') &= \sum_p \mathcal{P}_1(r)\mathcal{P}_2(p|r, t')\mathcal{P}_3(s|p) \\
&= \sum_p \left(\frac{1 + rR}{2}\right)\left(\frac{1 + prP(t')}{2}\right)\left(\frac{1 + \kappa(s)p\mathcal{D}}{2}\right)\mathcal{E}(s) \\
&= \left(\frac{1 + \kappa(s)rP(t')\mathcal{D}}{2}\right)\left(\frac{1 + rR}{2}\right)\mathcal{E}(s) \qquad (8.35)
\end{aligned}$$

---

[5] Except in the $J/\psi K_S^0$ backgrounds. As will be explained in Section 8.3, D and γ are both fixed to zero for these backgrounds.

8.3 $\mathcal{L}$ for $J/\psi K_S^0$

For the $J/\psi K_S^0$ sample, $r = +1$ for all events. For the signal events, $P(t') = -\sin 2\beta \sin \Delta m t'$, and $R = 1$. The dilution is $\mathcal{D} = \mathcal{D}_0$, where the subscript "0" denotes that this is the dilution for neutral B mesons. For the background $P = 1$ and $R = 1$, which means that all the background events are defined to be fake B^0. Lack of fake $\overline{B}^0$ events makes γ and $\mathcal{D}$ meaningless for the background, so both parameters are fixed at zero. Thus, the likelihood functions for $J/\psi K_S^0$ are:

$$\mathcal{L}_P = \left(\frac{1+\zeta_P M_N}{2W}\right) G(t;0,Y\sigma_t) \left(\frac{1}{2}\right) \mathcal{E}(s) \tag{8.36}$$

$$\mathcal{L}_L = \left(\frac{1+\zeta_L M_N}{2W}\right) G(t;t',Y\sigma_t) \otimes T_L(t';f_N,f_{\tau 2},\tau_1,\tau_2) \left(\frac{1}{2}\right) \mathcal{E}(s) \tag{8.37}$$

$$\begin{aligned}\mathcal{L}_B = {} & G(M_N;0,X)\ G(t;t',Y\sigma_t) \otimes \\ & \left[E(t';\tau_{B^0}) \left(\frac{1-\kappa(s)\mathcal{D}_0 \sin 2\beta \sin(\Delta m t')}{2}\right)\right] \mathcal{E}(s)\end{aligned} \tag{8.38}$$

8.4 $\mathcal{L}$ for $J/\psi K^+$

For the $J/\psi K^+$ signal, $P = 1$, because B^+ mesons do not mix. The dilution is $\mathcal{D} = -\mathcal{D}_+$, where $\mathcal{D}_+$ is the absolute value of the dilution for tagging charged B mesons. The dilution is expected to be negative because the correlation for B^+ mesons is opposite to that for B^0 mesons. As with all the other backgrounds, $P = 1$ for the $J/\psi K^+$ backgrounds. Thus, the likelihood functions for $J/\psi K^+$ are:

$$\mathcal{L}_P = \left(\frac{1+\zeta_P M_N}{2W}\right) G(t;0,Y\sigma_t) \left(\frac{1+r\kappa(s)\mathcal{D}_P}{2}\right) \left(\frac{1+rR_P}{2}\right) \mathcal{E}(s) \tag{8.39}$$

$$\begin{aligned}\mathcal{L}_L = {} & \left(\frac{1+\zeta_L M_N}{2W}\right) G(t;t',Y\sigma_t) \otimes T_L(t';f_N,f_{\tau 2},\tau_1,\tau_2) \\ & \times \left(\frac{1+r\kappa(s)\mathcal{D}_L}{2}\right) \left(\frac{1+rR_L}{2}\right) \mathcal{E}(s)\end{aligned} \tag{8.40}$$

$$\mathcal{L}_B = G(M_N;0,X)\ G(t;t',Y\sigma_t) \otimes E(t';\tau_B) \left(\frac{1-r\kappa(s)\mathcal{D}_+}{2}\right) \left(\frac{1+rR_B}{2}\right) \mathcal{E}(s) \tag{8.41}$$

8.5 $\mathcal{L}$ for $J/\psi K^{*0}$

There is an additional complication in the $J/\psi K^{*0}$ channel, arising from the possibility that the B flavor at decay will be determined incorrectly, as will be discussed in Section 9.2. The K and the π in the $J/\psi K^{*0}$ decay will sometimes be *swapped*, *i.e.* the kaon has been misidentified as a pion and the pion has been misidentified as a kaon. In these cases, the reconstructed B decay flavor and the reconstructed B mass will both be incorrect.

This swapping is incorporated into the fit through the addition of a second signal term, with the decay flavor switched. This term has a fraction P_S of the signal events (and the original signal term now has only $1 - P_S$ of the signal events). The swapped term has a different mass distribution, as well: $G(M_N, \mu_S, X_S)$. The parameters P_S, μ_S, and X_S are determined from Monte Carlo simulations, as will be explained in Section 9.2.

For the $J/\psi K^{*0}$ signal events, $P = \cos(\Delta mt)$. The dilution is $\mathcal{D} = \mathcal{D}_0$, the dilution for tagging neutral B mesons. As with the other modes, $P = 1$ for the backgrounds. Thus, the likelihood functions for $J/\psi K^{*0}$ are:

$$\mathcal{L}_P = \left(\frac{1+\zeta_P M_N}{2W}\right) G(t;0,Y\sigma_t) \left(\frac{1+r\kappa(s)\mathcal{D}_P}{2}\right) \left(\frac{1+rR_P}{2}\right) \mathcal{E}(s) \tag{8.42}$$

$$\begin{aligned}\mathcal{L}_L = & \left(\frac{1+\zeta_L M_N}{2W}\right) G(t;t',Y\sigma_t) \otimes T_L(t';f_N,f_{\tau 2},\tau_1,\tau_2) \\ & \times \left(\frac{1+r\kappa(s)\mathcal{D}_L}{2}\right) \left(\frac{1+rR_L}{2}\right) \mathcal{E}(s)\end{aligned} \tag{8.43}$$

$$\begin{aligned}\mathcal{L}_B = & (1-P_S)G(M_N;0,X)\ G(t;t',Y\sigma_t) \\ & \qquad \otimes \left[E(t';\tau_B)\left(\frac{1+r\kappa(s)\mathcal{D}_0\cos(\Delta mt')}{2}\right)\right] \left(\frac{1+rR_B}{2}\right)\mathcal{E}(s) \\ + & P_S G(M_N;\mu_S,X_S)\ G(t;t',Y\sigma_t) \\ & \qquad \otimes \left[E(t';\tau_B)\left(\frac{1-r\kappa(s)\mathcal{D}_0\cos(\Delta mt')}{2}\right)\right] \left(\frac{1-rR_B}{2}\right)\mathcal{E}(s)\end{aligned} \tag{8.44}$$

8.6 Summary of Fit Parameters

Table 8.2 summarizes the allowed ranges for the fit parameters. The fit parameters are all allowed to vary freely, within the range specified.[6] These limits are necessary to keep $\mathcal{L}$ positive for every event.

The parameter $\sin 2\beta$ is not fit for directly; instead, the combined quantity $\mathcal{D}_0 \sin 2\beta$ is measured. This quantity is then divided by the $\mathcal{D}_0$ to determine $\sin 2\beta$, as will be described in Chapter 12.

Parameters	Description	Allowed Range
ζ_P, ζ_L	Background mass slopes	$[-1/W, 1/W]$
f_B, f_L, f_N, $f_{\tau 2}$	Event fractions	$[0.0, 1.0]$
τ_1, τ_2	Background "lifetimes"	$(0.0, +\infty)$
X, Y	Error scale-factors	$(0.0, +\infty)$
R_B, R_P, R_L	Reconstruction asymmetries	$[-1.0, 1.0]$
ϵ_B, ϵ_P, ϵ_L	Tagging efficiencies	$[0.0, 1.0]$
α_P, α_L	Tagging charge biases	$[-1.0, 1.0]$
δ_P, δ_L	Tagging efficiency asymmetries	$[-1.0, 1.0]$
$\mathcal{D}_P$, $\mathcal{D}_L$, $\mathcal{D}_0$, $\mathcal{D}_+$	Dilutions	$[-1.0, 1.0]$
$\mathcal{D}_0 \sin 2\beta$	Raw CP asymmetry	$[-1.0, 1.0]$

Table 8.2: Allowed ranges for the fit parameters. For $J/\psi K^0_S$, the following parameters are fixed: $R_B = R_P = R_L = 1$, and $\mathcal{D}_P = \mathcal{D}_L = \delta_P = \delta_L = 0$. The parameter $\mathcal{D}_0 \sin 2\beta$ is fit as a single quantity.

In addition to these *fit parameters*, there are a number of *input parameters*, such as τ_{B^0}, the B^0 meson lifetime. These quantities are determined separately from the

---

[6] Except in $J/\psi K^0_S$, where the reconstruction asymmetries R_B, R_P, and R_L, are all fixed to 1, because there is no information available on the flavor of B meson at decay. The background dilutions and efficiency asymmetries $\mathcal{D}_P$, $\mathcal{D}_L$, δ_P, and δ_L for this sample are all fixed to 0 because the background events are all defined to be fake B^0.

likelihood fit, and are therefore fixed in the fit. The determination of these parameters and their uncertainties is described in Chapter 9. The measured values for the fit parameters will be discussed in Chapter 10. The input parameters are varied within their uncertainties, to calculate systematic uncertainties on the measured asymmetries, as will be described in Chapter 11.

Chapter 9

Input Parameters and Constraints for Likelihood Fit

In addition to the *fit parameters*, the maximum likelihood fit contains several *input parameters*. These are quantities measured independently of the likelihood fit. Section 9.1 discusses the B meson decay parameters τ_{B^0}, τ_{B^+}, and Δm. Section 9.2 discusses the parameters which describe the swapping effect in the $J/\psi K^{*0}$ sample: P_S, μ_S, and X_S. Section 9.3 describes the rest of the input parameters, which are used to describe the tagging charge bias and tagging efficiency asymmetry.

9.1 B Decay Parameters

The B meson lifetimes, τ_{B^+} and τ_{B^0}, and the B^0 mixing frequency, Δm, have been measured much more accurately elsewhere than can be done in this analysis. Therefore these parameters are fixed to the following values, which come from the Particle Data Group [32]:

$$\Delta m = 0.474 \pm 0.031 \text{ ps}^{-1} \tag{9.1}$$

$$\tau_{B^0} = 1.56 \pm 0.06 \text{ ps} \tag{9.2}$$

$$\tau_{B^+} = 1.62 \pm 0.06 \text{ ps.} \tag{9.3}$$

9.2 $K\pi$ Swapping in K^{*0} Reconstruction

At CDF, it is not possible to differentiate between tracks from pions and tracks from kaons, so the only way to determine whether a given track pair is $K^+\pi^-$ from a K^{*0} or $K^-\pi^+$ from a $\overline{K}^{*0}$ is to form the combined mass for each combination, and see which one is nearer the world-average K^{*0} mass (896 MeV/c). This is usually the correct choice, but the K^{*0} meson has a natural width so large that sometimes the incorrect assignment will be chosen.

To study this, a Monte Carlo sample of $B^0 \rightarrow J/\psi K^{*0}$ decays is generated and reconstructed with the same selection criteria as for the data.[1] Figure 9-1 shows the combined $K\pi$ masses for the correct ("unswapped") and incorrect ("swapped") assignments. The width of the swapped peak ($\approx$ 30 MeV/c^2) is due to the natural width of the K^{*0} resonance; the detector resolution is only $\approx$ 10 MeV/c^2 for these decays. Roughly 10% of the time, the swapped mass is closer to the world-average K^{*0} mass than the unswapped mass is, and the incorrect assignment is chosen.

Because the flavor of the B^0 meson is taken from the charge of the kaon, the swapped events will have the incorrect B^0 flavor at decay. Because the K and π have different masses, the reconstructed B meson mass (M_{FIT}) will be incorrect. These effects bias the dilution measurement in $J/\psi K^{*0}$ and must therefore be corrected for. As discussed in Chapter 8, this correction is incorporated into the fit through the introduction of a new signal term into the likelihood fit. The correction requires three input parameters, P_S, μ_S, and X_S, which are taken from the Monte Carlo simulation.

Figure 9-2 shows the normalized mass distributions for unswapped and swapped Monte Carlo candidates. The distribution for unswapped events behaves as expected, *i.e.* it follows a Gaussian distribution with mean $\mu = 0$ and RMS $\sigma \simeq 1$. The swapped distribution is much wider and shifted slightly downward in M_N. Roughly 10% percent of the reconstructed candidates are swapped.

The three parameters which describe swapping are taken from this plot: $P_S = 0.1$, $\mu_S = -0.5$, and $X_S = 5$. Figure 9-3 shows the normalized mass distribution for the

---

[1] The generation of this Monte Carlo sample is described in Appendix A.

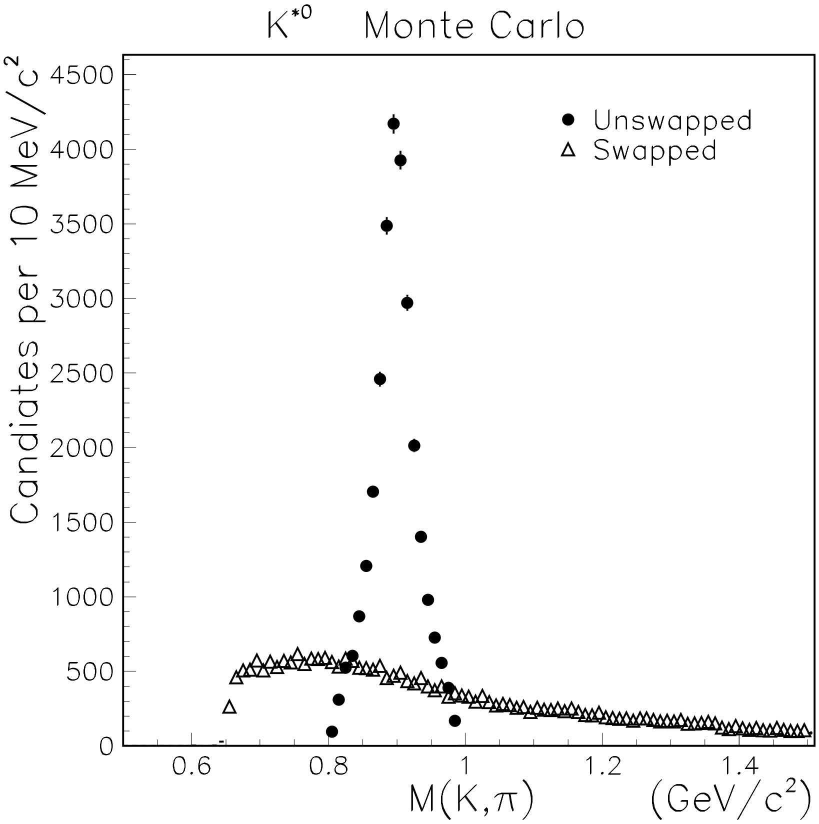

Figure 9-1: The combined $K\pi$ mass distributions for Monte Carlo K^{*0} mesons. The circles represent the correct $(K^{+}\pi^{-})$ mass assignment, and the triangles represent the incorrect $(K^{-}\pi^{+})$ mass assignment. The width of the "unswapped" peak (circles) is roughly 30 MeV$/c^2$, much larger than the detector resolution of 10 MeV$/c^2$ (the size of one bin on the plot).

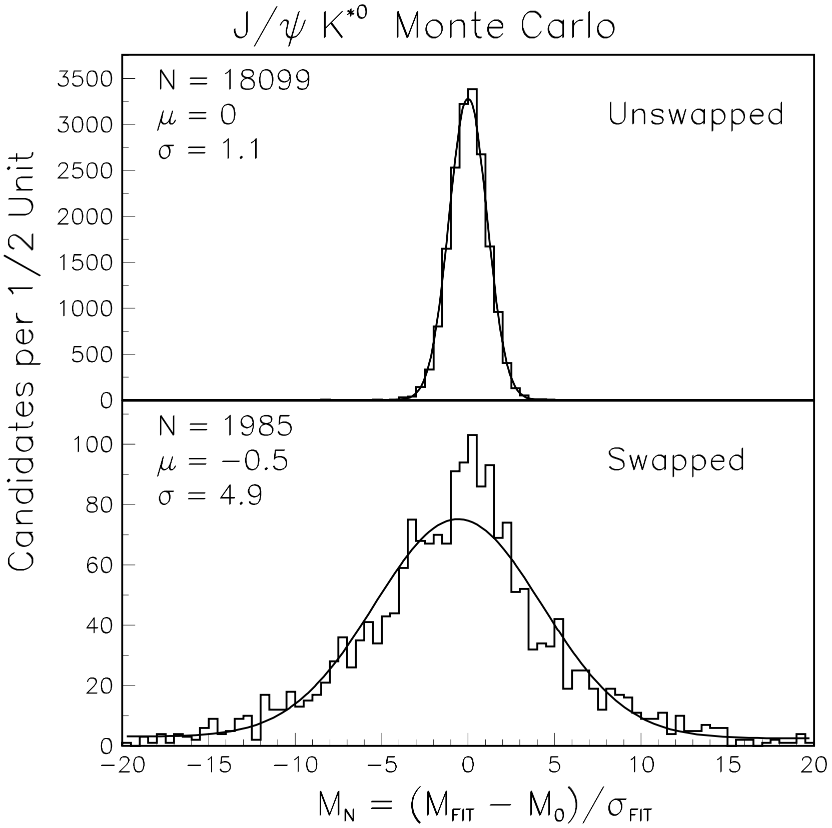

Figure 9-2: The normalized B mass distribution of $J/\psi K^{*0}$ Monte Carlo for those events reconstructed with the correct K-π assignment (top), and those which were swapped (bottom). There are 18099 unswapped candidates and 1985 swapped candidates. The distributions are each fit to a Gaussian, and the means, μ, and widths, σ, are given.

real data candidates with $ct > 0$. Three curves are overlaid on this histogram: one which includes only the background component from the fit, one which includes the background and the swapped events, and one which includes all three components. The impact of swapping is small, but the correction is included for completeness.

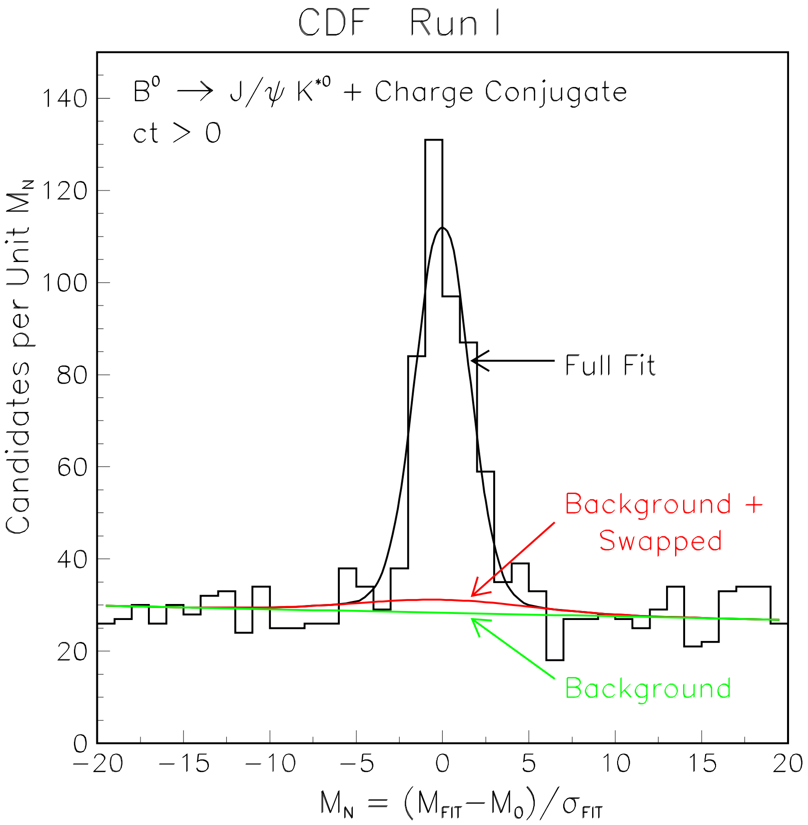

Figure 9-3: The distribution of $J/\psi K^{*0}$ candidates which have $ct > 0$. The line indicates the background distribution indicated by the full likelihood fit. The curve just above the line indicates the distribution of the swapped events, and the top curve shows the distribution of both types of signal, plus background.

9.2.1 Uncertainties on Swapping Parameterization

The difference in the reconstructed $K\pi$ mass for the two permutations of mass assignments for a given pair of tracks is approximately:

$$\Delta M \simeq \left(\frac{M_K^2 - M_\pi^2}{M_{K^{*0}}}\right)\frac{\Delta p}{p} \simeq \frac{\Delta p}{p}\,260\,\mathrm{MeV}/c^2 \tag{9.4}$$

where $M_K = 0.497$ GeV/c^2 is the K^+ mass, $M_\pi = 0.139$ GeV/c^2 is the π^+ mass, $M_{K^{*0}} = 0.896$ GeV/c^2 is the mean of the K^{*0} resonance, Δp is the difference between the magnitudes of the momenta of the two tracks, and $\overline{p}$ is the average of the magnitudes of the two momenta.

When $\Delta p/\overline{p} \approx 0.1$, the mass difference will be approximately the same as the width of the K^{*0} mass peak. Thus, when the K and π momenta are similar, the probability of swapping will be high, and when they are very different, the probability will be low. The momenta of these particles depends on the kinematics of the K^{*0} decay, which in turn depends on the kinematics of the B meson decay and on the B meson momentum. Figure 9-4 shows the P_T distributions for data and Monte Carlo candidates. The agreement is good, indicating that the $p_T(B)$ dependence of the reconstruction efficiency of B mesons is well modeled by the Monte Carlo.

The values of the three swapping parameters are plotted versus $p_T(B)$. From these plots, ranges that encompass the variations in these values are determined, and these ranges are used in Chapter 11 to calculate systematic uncertainties on the dilution measurement in $J/\psi K^{*0}$. The variations in the parameters are all conservative, so the systematic uncertainties calculated from them are probably overestimations of the true systematic uncertainties. But the statistical uncertainty on the dilution measurement in $J/\psi K^{*0}$ is five times larger than the combined systematic uncertainties, so this overestimation has little effect in the total uncertainty on the dilution measurement.

Uncertainty on P_S

Figure 9-5 shows the probability that a candidate is swapped, as a function of $p_T(B)$. The probability remains nearly flat, at 0.1. To be conservative, the uncertainty associated with this parameter is 0.1, as indicated by the dotted lines in the plot.

Uncertainty on μ_S

Figure 9-6 shows the mean of the M_N distribution for swapped and unswapped candidates, as a function of $p_T(B)$. The unswapped candidates all have mean consistent

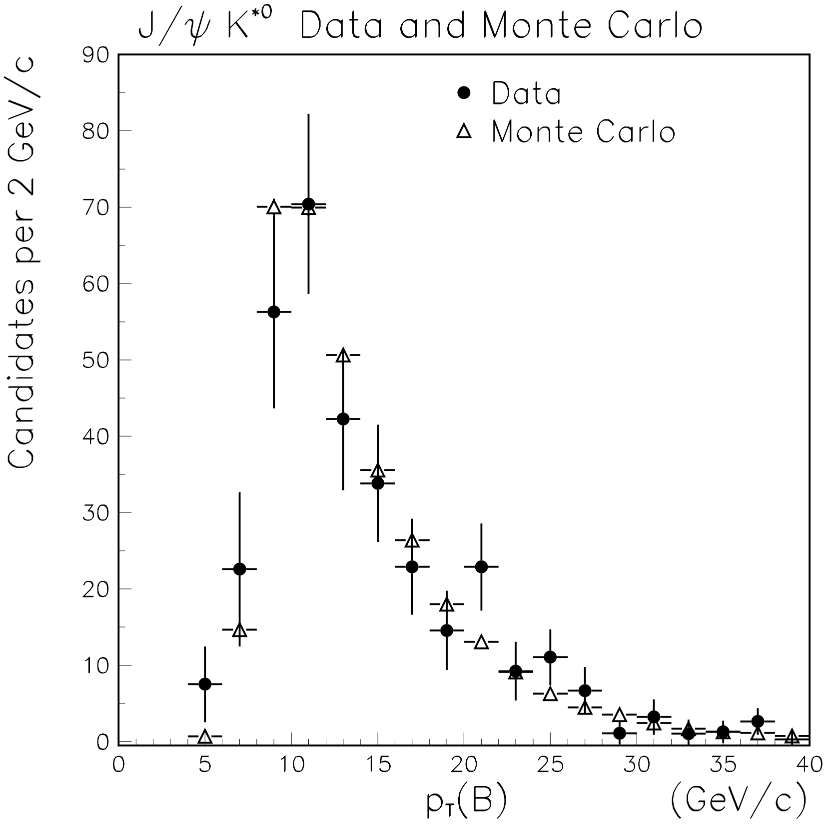

Figure 9-4: The p_T distribution of B^0 candidates in $J/\psi K^{*0}$ decays for data and Monte Carlo. The data have had the distributions for the mass-sideband subtracted, to remove the background; the Monte Carlo has essentially no background. The Monte Carlo histogram is normalized to the same area as the data.

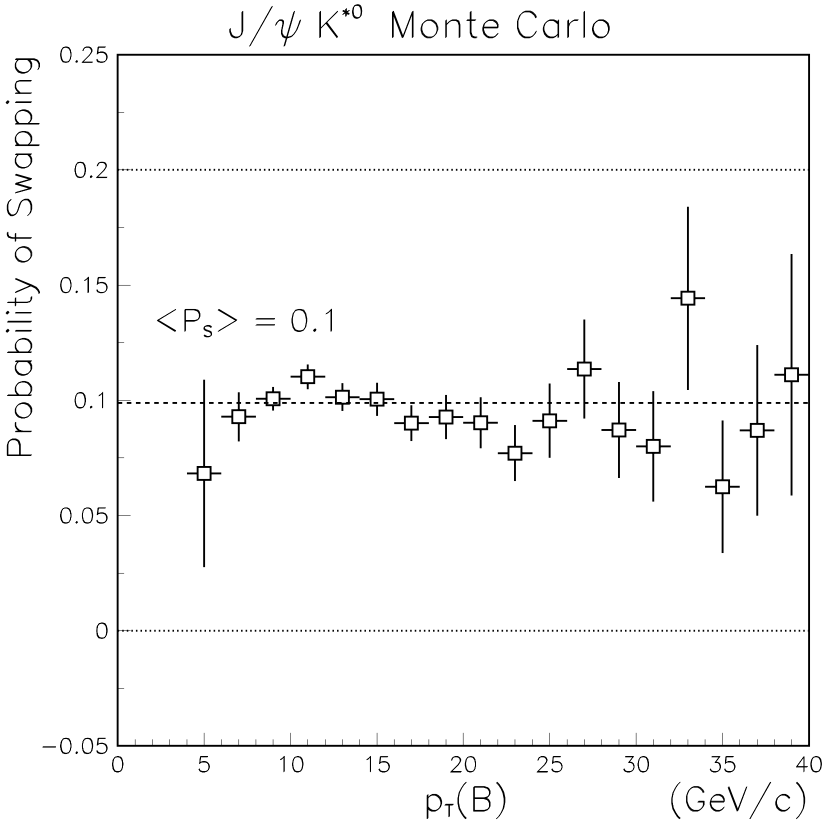

Figure 9-5: The fraction of swapped candidates as a function of the B^0 meson p_T. The dashed line represents the central value used in the fit: 0.1. The dotted lines represent the $\pm 1\sigma$ variations used for systematic uncertainties.

with zero. The swapped candidates generally have mean between -1 and 0, with a small dependence on $p_T(B)$ apparent. The central value used for μ_S is -0.5, represented by the dashed line. The dotted lines show the $\pm 1\sigma$ uncertainties associated with μ_S: ± 0.5

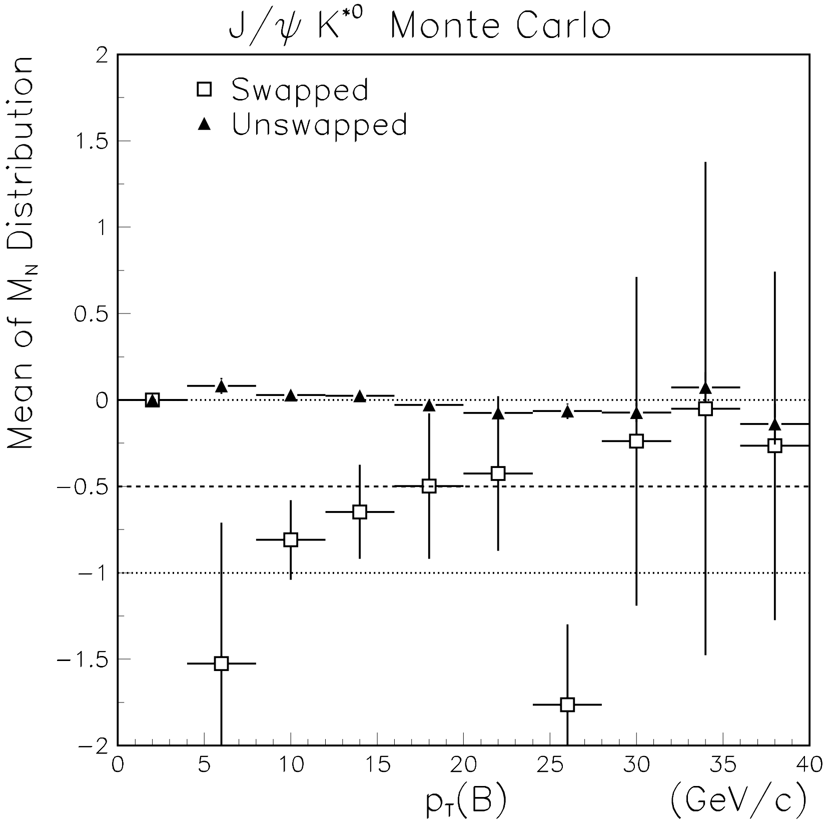

Figure 9-6: The mean of the normalized B mass distribution for swapped and unswapped candidates in bins of the B^0 meson p_T. The central value and $\pm 1\sigma$ variations are indicated by the dashed and dotted lines.

Uncertainty on X_S

Figure 9-7 shows the RMS of the M_N distributions for swapped and unswapped events, as a function of $p_T(B)$. For the unswapped events, the RMS is consistently a little above the dot–dash line at 1.0. For the swapped events, the RMS falls from ~ 7 at low $p_T(B)$ to ~ 3 at high $p_T(B)$. The dashed line shows the central value used for the fit: 5.0. This value is then varied up to 7 and down to 3 (dotted lines) to account for systematic uncertainty due to the uncertainty on X_S.

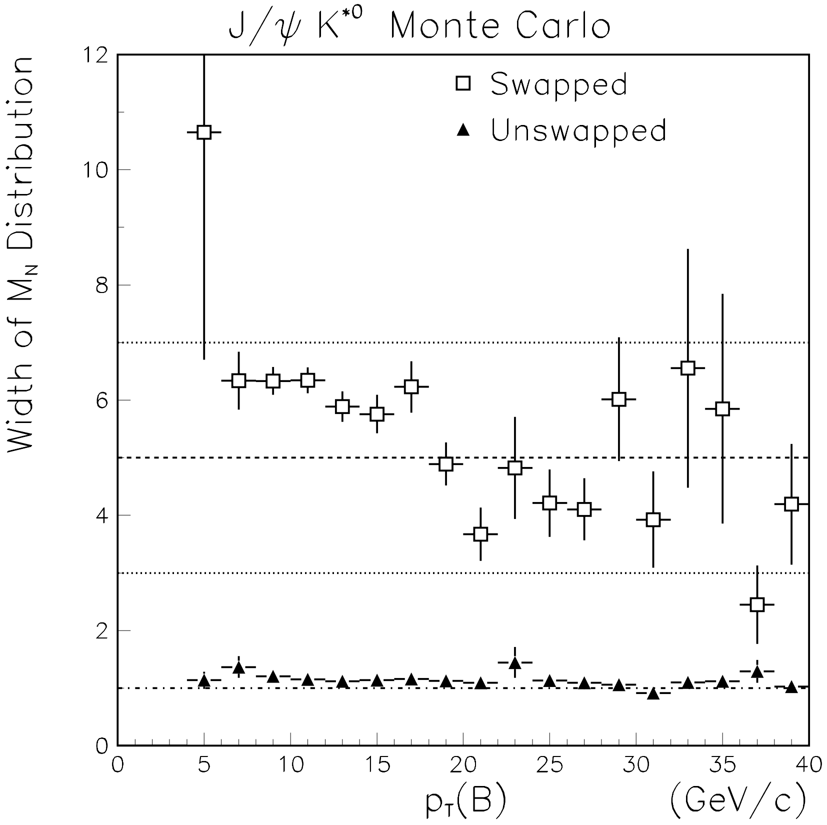

Figure 9-7: The RMS of the normalized B-mass for swapped and unswapped candidates as a function of the B meson p_T. Dashed line is the central value used as input to the likelihood fitter, and the dotted lines show the "1σ uncertainties" used in the likelihood (see text).

9.3 SST Tagging Parameterization

As described in Section 8.2.3, a tagging algorithm requires four parameters to characterize it. The ones used in this analysis are: ϵ, $\mathcal{D}$, α, and γ. The parameter ϵ is the charge-averaged tagging efficiency, and is a fit parameter in each sample. The charge-averaged tagging dilution $\mathcal{D}$ is fit for in $J/\psi K^+$ and $J/\psi K^{*0}$, while in $J/\psi K^0_S$, the combined parameter $\mathcal{D} \sin 2\beta$ is fit for. The dilution $\mathcal{D}$ for $J/\psi K^0_S$ is calculated including the dilutions in other B meson decay modes, as will be described in Chapter 12.

The tagging charge bias α is measured separately, in an inclusive $B \to J/\psi X$ sample. While it would be possible to measure α in $J/\psi K^+$ or $J/\psi K^{*0}$, the inclusive sample is much larger and yields a more accurate measurement. Section 9.3.1 describes this measurement. The tagging efficiency bias γ is not measured, but is mathematically constrained by the value of α, as described in Section 9.3.2.

The backgrounds contain events which either do not contain B mesons, or contain misreconstructed B mesons, so the values of the tagging parameters could be different for the background. Therefore, rather than constraining the tagging parameters for the background to be the same as for the signal, they are allowed to float freely in the fit.

9.3.1 Tagging Charge Bias

The tagging charge bias parameter (from equation 8.24)

$$\alpha = \frac{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)} \tag{9.5}$$

is measured in an inclusive $B \to J/\psi X$ sample, where only the J/ψ is identified. In this sample, J/ψ mesons are reconstructed with the same criteria as described in Chapter 6. Additional selection criteria are then applied to improve the B purity:

- $p_T(J/\psi) > 4$ GeV/c, and
- $L_{xy}(J/\psi) > 200$ μm.

It is known that for a looser L_{xy} cut at 100 μm the J/ψ sample is 90% from b decays [56]; the 200 μm cut consequently gives a very pure b sample. Figure 9-8 shows the mass distribution of the selected dimuon pairs with the signal and sideband regions marked: the signal region is between 3040 and 3160 MeV/c^2, and the sideband regions are 2900 to 3000 MeV/c^2 on the low side and 3200 to 3300 MeV/c^2 on the high side. In all the following plots from this displaced J/ψ sample, the sidebands have been subtracted.

This sample is then tagged with the SST algorithm, approximating the B meson momentum direction by the J/ψ momentum direction. The sample is expected to contain equal numbers of B and $\overline{B}$ hadrons, so the numbers of positive tags (N_+) and negative tags (N_-) should be:

$$N_+ = N_{evt}\,\frac{\mathcal{P}_3(+|+)+\mathcal{P}_3(+|-)}{2}, \qquad \text{and} \tag{9.6}$$

$$N_- = N_{evt}\,\frac{\mathcal{P}_3(-|+)+\mathcal{P}_3(-|-)}{2}, \tag{9.7}$$

where N_{evt} is the number of events in the sample, and $\mathcal{P}_3(s|p)$ is the probability of a B meson produced with flavor p being tagged with charge s, as in Section 8.2.3.

The asymmetry between these two quantities is:

$$\frac{N_+ - N_-}{N_+ + N_-} = \frac{\mathcal{P}_3(+|+)+\mathcal{P}_3(+|-)-\mathcal{P}_3(-|+)-\mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+)+\mathcal{P}_3(+|-)+\mathcal{P}_3(-|+)+\mathcal{P}_3(-|-)} \equiv \alpha. \tag{9.8}$$

where the equivalence is from equation 9.5. The tagging charge bias is measured to be

$$\alpha = 1.9 \pm 1.0\%. \tag{9.9}$$

The dependences of this bias on the p_T and the L_{XY} of the J/ψ meson are also studied. The J/ψ mesons in this sample have similar p_T distribution to those in the fully-reconstructed sample, but not identical. Any dependence on $p_T(J/\psi)$ would introduce a complication which would need correction. Figure 9-9 shows the p_T distribution of the J/ψ mesons used in this study, the efficiency ϵ versus p_T, and the charge bias α versus p_T. The efficiency rises with $p_T(J/\psi)$, but the tagging charge bias does not appear to show any dependence on $p_T(J/\psi)$.

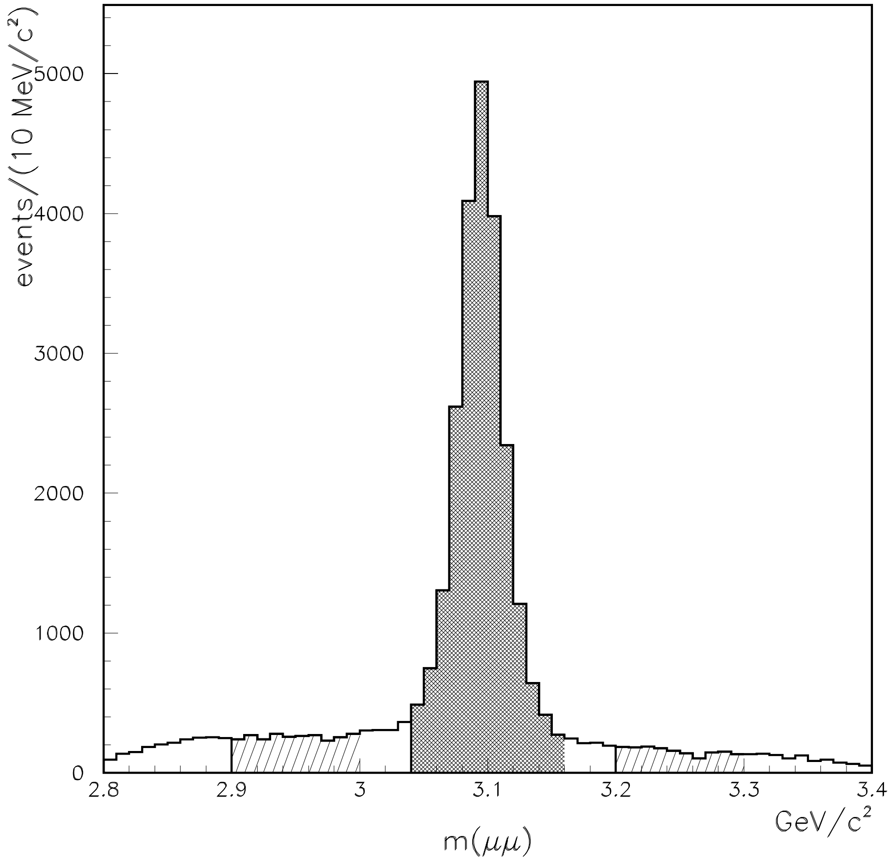

Figure 9-8: J/ψ mass distribution for events in the inclusive $B \rightarrow J/\psi X$ sample. The shaded region indicates the signal region, and the hatched region indicates the sidebands.

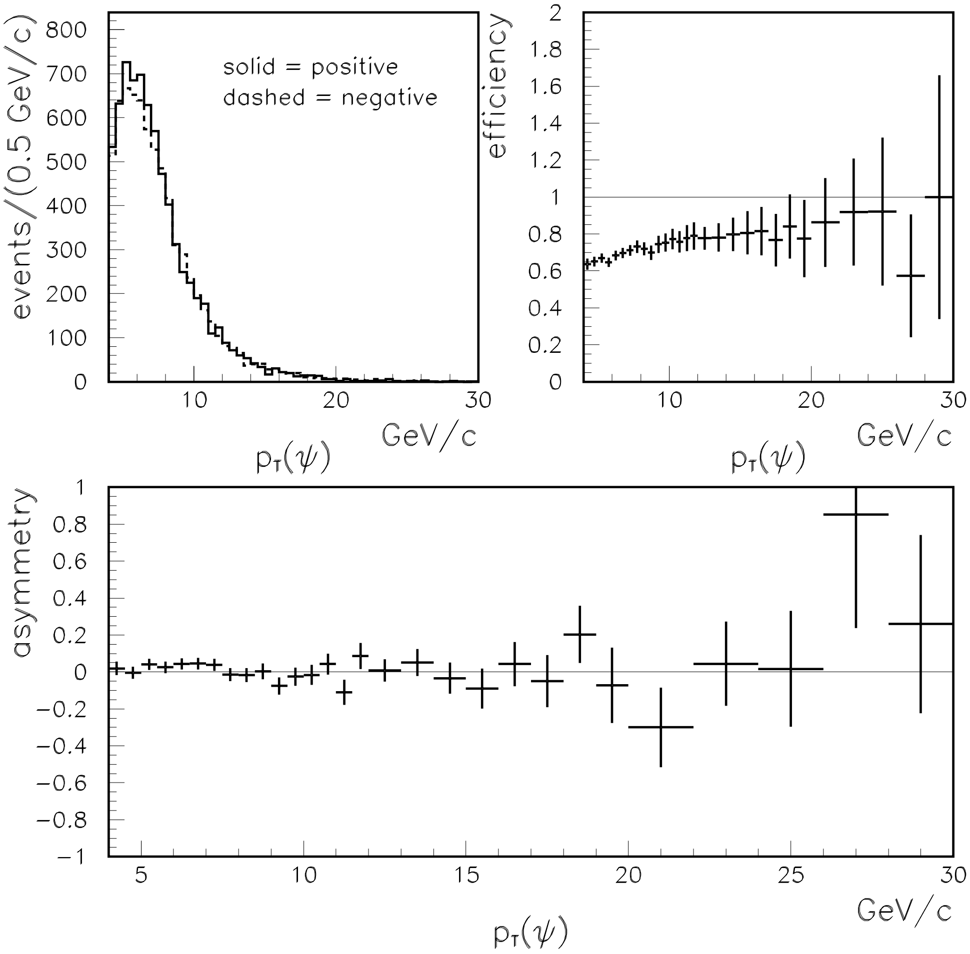

Figure 9-9: The top-left plot shows p_T spectra for positively and negatively tagged J/ψ mesons. The top-right plot shows the charge-averaged tagging efficiency, as a function of $p_T(J/\psi)$. The bottom plot shows the tagging asymmetry as a function of $p_T(J/\psi)$.

Figure 9-10 shows the L_{XY} distribution for J/ψ mesons (the requirement that $L_{XY} > 200\mu$m is removed, for this plot only). The tagging charge bias versus L_{XY} is also plotted. No dependence of α on L_{XY} is evident.

The dependence of the charge bias is also studied versus two other variables: $p_T(tag)$, the transverse-momentum of the track used for the tag, and n_{PI}, the number of primary interactions in the event. If the bias is due to the wire-plane tilt in the CTC, as suspected, a variation with $p_T(tag)$ is expected. Low-momentum tracks should show a large effect, and high-momentum tracks should show virtually no bias. The number of primary interactions affects the *event occupancy*, or the average number of hits per wire in the CTC. When the occupancy gets to be very high, the efficiency begins to drop, and the charge bias might become more pronounced.

Figure 9-11 shows the dependence of the tagging charge bias on each variable. Rather than plotting the dependence versus $p_T(tag)$, it is plotted versus $1/p_T(tag)$, which is proportional to the track curvature (which is the quantity actually measured by the tracker). A clear dependence on both variables is seen.

The selection of the track with the smallest p_T^{rel} (with respect to the B meson) also preferentially selects tracks with low p_T. Therefore the charge bias is also considered for all SST candidates, not just the one with the minimum p_T^{rel}.[2] The B candidates in the inclusive J/ψ sample all have unidentified daughter particles, which could be used as tags. To study the effects of this, the charge bias is considered for SST candidates with a "b vertex veto": tracks with an impact parameter with respect to the J/ψ vertex greater than 2 standard-deviations significant. A fourth category of tracks is also considered: the "side-cone candidates," which satisfy all the SST candidate requirements, except they must have $1.0 < \Delta R < 2.0$, instead of the normal $\Delta R < 0.7$. These tracks point away from the direction of the B momentum.

---

[2] The track chosen by SST depends on the p_T cutoff for tracks considered. However, when all candidates are considered, including lower p_T tracks does not bias the events with higher p_T tracks. Therefore, for candidate plots, the $p_T > 0.4$ GeV/c requirement is dropped. The requirement is still used for the actual tagging.

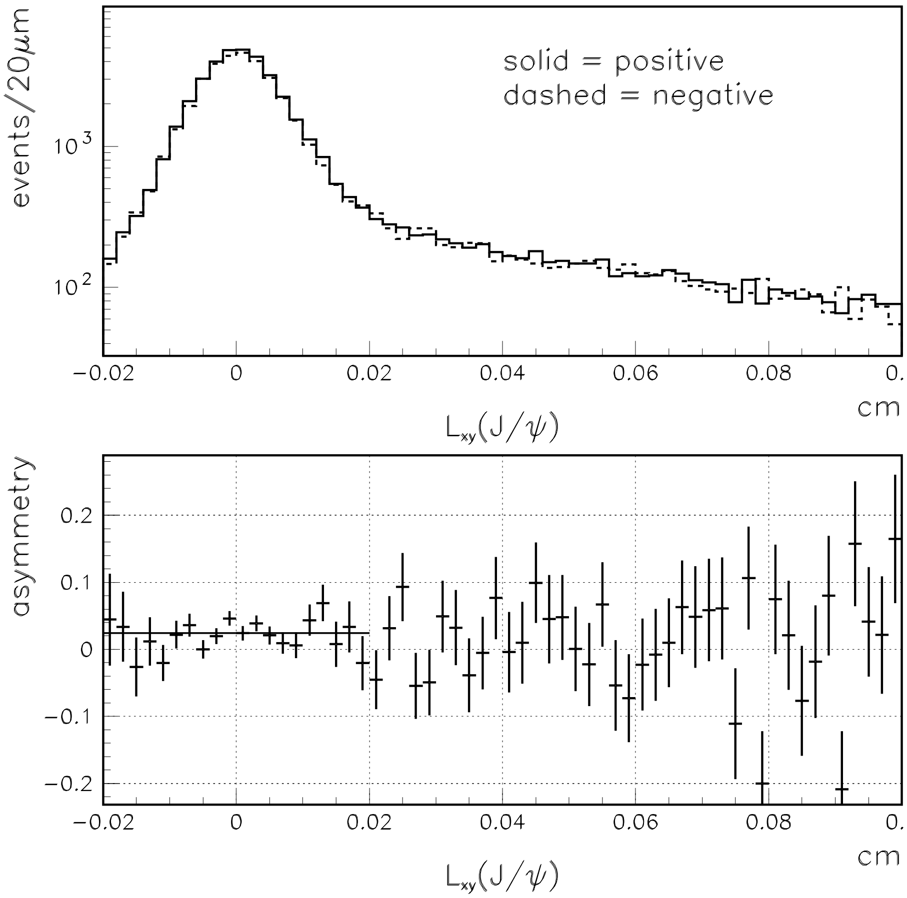

Figure 9-10: The top plot shows the number of positive and negative tags as a function of L_{xy}, the transverse flight distance of the J/ψ. The bottom plot shows tagging asymmetry as a function of L_{xy}. The line between $L_{XY} = -0.02$ and $L_{XY} = +0.02$ indicates the average tagging charge bias (+0.021) for prompt J/ψ mesons.

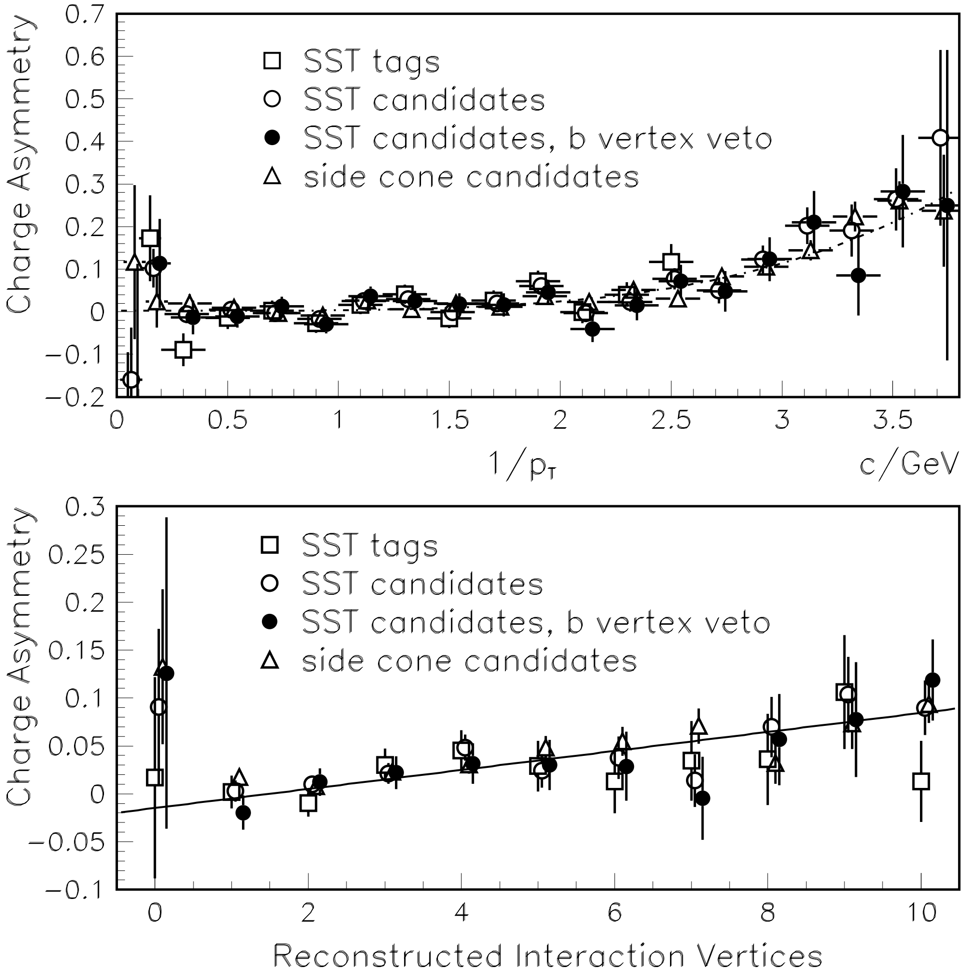

Figure 9-11: Tagging asymmetry dependence on $1/p_T$ (top) and n_{PI} (bottom), superimposed on the charge asymmetry of the SST candidates, that of the "candidates" in the side cone, and that of the SST candidates which are 2σ away from the J/ψ vertex. In the candidate cases we ignore the $p_T > 400$ MeV$/c$ cut since we are plotting a function of p_T. The curves are fit to the last mentioned set of points.

Figure 9-11 shows the charge bias for all four categories of tracks, versus $1/p_T(tag)$ and n_{PI}. There is no discernible difference between the categories, which leads to the conclusion that while the charge bias does depend on $1/p_T(tag)$ and n_{PI}, it does not depend on much else. Therefore α is parameterized in terms of $p_T(tag)$ and n_{PI}, and this parameterization is used for the fully-reconstructed events.

Parameterization of α

The dependence of α on $p_T(tag)$ and n_{PI} is measured in the SST candidates sample, as it is the largest. The results from the other samples are all consistent with the results from this sample. The p_T-dependent asymmetry shape is well fit by the function $\alpha(p_T) = ap_T^{-4}$, and the n_{PI}-dependent shape can be described by a line. Therefore, the following parameterization is used:

$$\alpha(p_T, n_{PI}) = [a_1(n_{PI} - n_0) + b_1](p_T^{-4} - p_{T0}^{-4}) + [a_2(n_{PI} - n_0) + b_2] \quad (9.10)$$

The four quantities a_1, a_2, b_1, and b_2 are fit parameters. The quantities p_{T0} and n_0 are "pivot points," chosen to minimize the correlations between the four parameters. This allows the a and b parameters to be varied independently for the calculation of systematic uncertainties, as is described in Chapter 11.

The pivot values which give the minimum correlation between the four parameters are $p_{T0} = 20\,(c/\mathrm{GeV})^4$ and $n_0 = 3$. The fit values for the parameters are:

$$a_1 = (3.9 \pm 1.8) \times 10^{-4}(\mathrm{GeV}/c)^4 \quad (9.11)$$

$$b_1 = (1.3 \pm 0.4) \times 10^{-3}(\mathrm{GeV}/c)^4 \quad (9.12)$$

$$a_2 = (1.4 \pm 0.4) \times 10^{-2} \quad (9.13)$$

$$b_2 = (2.6 \pm 0.8) \times 10^{-2}. \quad (9.14)$$

9.3.2 Efficiency Bias

The tagging efficiency bias (from equation 8.25)

$$\gamma \equiv \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)} \quad (9.15)$$

is constrained by the value of the tagging charge bias α. This section describes the rational behind this constraint and the region of allowed values.

Depending on the cause of the tagging charge bias, there might also be an associated tagging efficiency bias. If the charge bias is due to lost tracks of one charge, then the B mesons which preferentially tag on tracks of this charge will lose more tags than the other flavor of B meson, and will thus have lower tagging efficiency. If, on the other hand, the charge bias is due to the addition of tracks, all of the same sign, that might happen equally to both types of B meson, introducing no charge asymmetry.

At CDF, both types of effect are expected to be present. The tilt in the wire-planes in the CTC biases the track reconstruction efficiency in favor of positive tracks (*i.e.* the reconstruction of negative tracks is less efficient). As pions (of both charges) travel through the detector, they occasionally interact with atomic nuclei, ejecting protons (an effect called *spallation*). As there are no antiprotons in the matter making up the detector, all these tracks are positive. This should add positive tracks in events with B and $\overline{B}$ hadrons equivalently, so it should not introduce any tagging efficiency bias. Furthermore, the imposition of impact-parameter significance cuts (Section 7.2) reduces spallation to negligible levels.

If a negatively-charged track that would have been used as a tag is not reconstructed, then the tagging algorithm selects the track with the next lowest p_T^{rel}, if there is one that is reconstructed. If not, then the tag would be lost, which would introduce a net positive tagging charge bias. If the new track found is also negatively charged, then the loss has no effect, as it is only the sign of the track that matters. If, on the other hand, a positive track is chosen, then the tagging charge bias would be even larger than that from simply losing the negative track. To evaluate this effect, we proceed as follows.

We introduce a new parameter, η, to denote the fraction of negative tags which are lost (not counting those which retag on another negative track). Also let $f_{1(2)}$ denote the fraction of $B^0(\overline{B}^0)$ mesons which, having lost a negative tag, retag on a positive track. Section 7.1, the introduction to same-side tagging, explains why there are expected to be more positive tracks near B^0 mesons and more negative tracks near $\overline{B}^0$ mesons. From this difference, it is predicted that:

$$0 \leq f_2 \leq f_1 \leq 1. \tag{9.16}$$

Let $\mathcal{D}'$ and ϵ' be the dilution and efficiency of the flavor-tagging algorithm in the absence of a charge asymmetry in the track losses. Including the effects of biased track loss changes the tagging probabilities to:

$$\begin{aligned}
\mathcal{P}_3(+|+) &= \epsilon'\left(\frac{1+\mathcal{D}'}{2}\right) + \epsilon'\left(\frac{1-\mathcal{D}'}{2}\right)\eta f_1 && (9.17)\\
\mathcal{P}_3(-|+) &= \epsilon'\left(\frac{1-\mathcal{D}'}{2}\right)(1-\eta) && (9.18)\\
\mathcal{P}_3(+|-) &= \epsilon'\left(\frac{1-\mathcal{D}'}{2}\right) + \epsilon'\left(\frac{1+\mathcal{D}'}{2}\right)\eta f_2 && (9.19)\\
\mathcal{P}_3(-|-) &= \epsilon'\left(\frac{1+\mathcal{D}'}{2}\right)(1-\eta), && (9.20)
\end{aligned}$$

where equations 9.18 and 9.20 reflect the loss of negative tags, and equations 9.17 and 9.19 reflect the addition of new positive tags. From these equations, the modified tagging parameters can be calculated:

$$\begin{aligned}
\epsilon &= \frac{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}{2}\\
&= \epsilon'\left(\frac{2-\eta+\eta\overline{f}-\mathcal{D}'\eta\Delta f}{2}\right) && (9.21)\\
\mathcal{D} &= \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}\\
&= \frac{\mathcal{D}'(2-\eta-\eta\overline{f})+\eta\Delta f}{2-\eta+\eta\overline{f}-\mathcal{D}'\eta\Delta f} && (9.22)\\
\alpha &= \frac{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) + \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}\\
&= \frac{\eta(1+\overline{f}-\mathcal{D}'\Delta f)}{2-\eta+\eta\overline{f}-\mathcal{D}'\eta\Delta f} && (9.23)\\
\gamma &= \frac{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) + \mathcal{P}_3(-|+) - \mathcal{P}_3(-|-)}{\mathcal{P}_3(+|+) - \mathcal{P}_3(+|-) - \mathcal{P}_3(-|+) + \mathcal{P}_3(-|-)}\\
&= \frac{\eta(\mathcal{D}'-\mathcal{D}'\overline{f}-\Delta f)}{\mathcal{D}'(2-\eta-\eta\overline{f})+\eta\Delta f} && (9.24)
\end{aligned}$$

where the two quantities $\overline{f}$ and Δf have been defined as follows:

$$\overline{f} = \frac{f_1+f_2}{2} \tag{9.25}$$

$$\Delta f = \frac{f_1 - f_2}{2}. \tag{9.26}$$

With these expressions, γ/α can be calculated:

$$\frac{\gamma}{\alpha} = \frac{(\mathcal{D}' - \mathcal{D}'\overline{f} + \Delta f)(2 - \eta + \eta\overline{f} - \mathcal{D}'\eta\Delta f)}{(1 + \overline{f} - \mathcal{D}'\Delta f)(2\mathcal{D}' - \mathcal{D}'\eta - \mathcal{D}'\eta\overline{f} + \eta\Delta f)}. \tag{9.27}$$

The behavior of this rather complicated expression for γ/α is shown in Figure 9-12 for values of $\alpha = 0.02$ and $D' = 0.166$, close to what is expected in data. The function is not defined for $f_2 > f_1$ as it is expected that $f_2 \leq f_1$ (equation 9.16).

When $f_1 = f_2 = 0$ (the leftmost point on the plot), $\gamma/\alpha = 1$. This is the point corresponding to the case where events *never* retag on positive tracks after losing a negative tag. When $f_1 = f_2 = 1$ (the rightmost point), $\gamma/\alpha = 0$. This corresponds to the case where events *always* retag on positive tracks after losing a negative tag. In this case, the track loss does not affect either efficiency, as no tags are lost; they are only switched.

The largest value of γ/α happens at the "rear-most" point, where $f_1 = 1$ and $f_2 = 0$. This point corresponds to the unlikely case that a B^0 will *always* retag on a positive track, but a $\overline{B}^0$ will *never* retag on a positive track. At this point the value of γ/α is:

$$\left(\frac{\gamma}{\alpha}\right)_{MAX} = \frac{(\mathcal{D}' + 1)(4 - (1 + \mathcal{D}')\eta)}{(3 - \mathcal{D}')(4\mathcal{D}' + (1 - 3\mathcal{D}')\eta)}. \tag{9.28}$$

When $f_1 = 1$ and $f_2 = 0$, and η is on the order of a few percent, equations 9.22 and 9.23 can be approximated as:

$$\alpha \simeq 0.7\eta \tag{9.29}$$

$$\mathcal{D} \simeq \mathcal{D}'(1 - \eta/2) \tag{9.30}$$

From these, η and $\mathcal{D}'$ can be derived in terms of measured quantizes: $\eta \simeq 1.4\alpha$, and $\mathcal{D}' \simeq \mathcal{D}$. The neutral dilution measured in [45] is 18 ± 3%. For this range of dilutions, the values of $(\gamma/\alpha)_{MAX}$ are:

$$\mathcal{D} = 0.15 \implies (\gamma/\alpha)_{\mathrm{MAX}} = 2.5 \tag{9.31}$$

$$\mathcal{D} = 0.18 \implies (\gamma/\alpha)_{\mathrm{MAX}} = 2.2 \tag{9.32}$$

$$\mathcal{D} = 0.21 \implies (\gamma/\alpha)_{\mathrm{MAX}} = 2.0 \tag{9.33}$$

The value of γ/α used in the fit is 1.0. For the purpose of evaluating systematic uncertainties (to be described in Chapter 11), it is varied between 0.0 and 2.5.[3]

9.4 Summary of Input Parameters

The results of this chapter are summarized in table 9.1 which shows the input parameters and their uncertainties. The central values in this table are the values of the input parameters used in the likelihood fits, which are described in Chapter 10. The variations from this table are used in Chapter 11 to derive systematic uncertainties on the asymmetry measurements.

---

[3] This is based on the assumption that spallation is truly negligible. However, if spallation were significant, it would increase α without increasing γ, making the ratio γ/α smaller. Thus, the effect of spallation is already included within the range considered for systematic uncertainties.

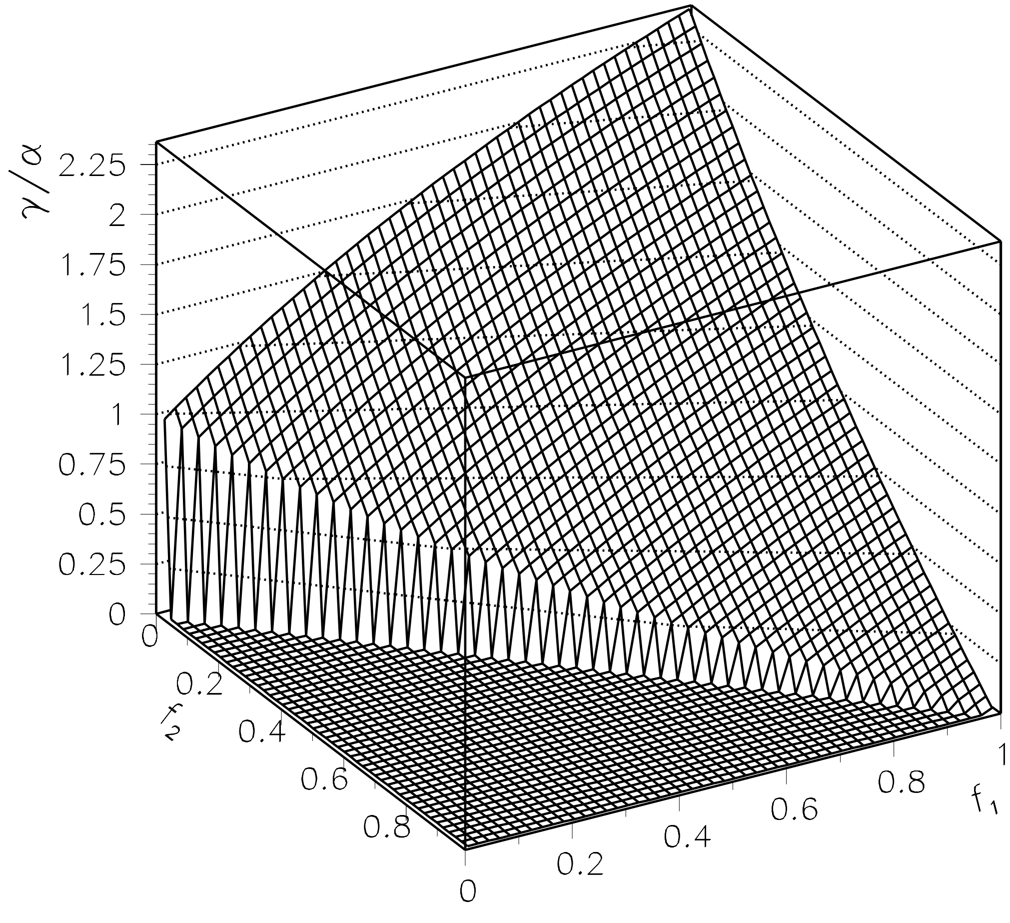

Figure 9-12: Variation of γ/α versus f_1 and f_2 for $\alpha = 2\%$ and nominal dilution $\mathcal{D}' = 16.6\%$. Note that (0,0) is the leftmost corner and (1,1) the rightmost.

Parameter		Central Value	Variation
	B decay Parameters		
τ_{B^0}	(μm)	468	±18
τ_{B^+}	(μm)	486	±18
Δm	(ps^{-1})	0.474	±0.031
	$K\pi$ Swapping Parameters		
P_S		0.1	±0.1
μ_S		-0.5	±0.5
X_S		5.0	±2.0
	Tagging Charge Bias		
a_1	$\times10^{-4}$ (GeV/c)^4	3.9	±1.8
b_1	$\times10^{-3}$ (GeV/c)^4	1.3	±0.4
a_2	$\times10^{-2}$	1.4	±0.4
b_2	$\times10^{-2}$	2.6	±0.8
	Tagging Efficiency Bias		
γ/α		1.0	$+1.5/-1.0$

Table 9.1: Input parameters for the likelihood fit and their uncertainties.

Part IV

Results of the Experiment

Chapter 10

Measurement of Asymmetry

With the input parameters set to the values described in Chapter 9 (in table 9.1), the likelihood fit is performed. Section 10.1 provides the results of the fit and some interpretation of these results, and Section 10.2 provides some plots to allow for visual interpretation of the fit results. In Chapter 11, the input parameters will be varied, within the ranges specified in Chapter 9, to calculate the systematic uncertainties on the fit results.

10.1 Likelihood Fit Results

The results of the fits are summarized in tables 10.1 and 10.2. The main parameters of interest are: the raw CP asymmetry from $J/\psi K_S^0$: $\mathcal{D}_0 \sin 2\beta = 0.31 \pm 0.18$, the neutral dilution from $J/\psi K^{*0}$: $\mathcal{D}_0 = 0.17 \pm 0.11$, and the charged dilution from $J/\psi K^+$: $\mathcal{D}_+ = 0.19 \pm 0.05$. Chapter 12 compares these dilutions to those measured in [45], extrapolates the dilution appropriate for $J/\psi K_S^0$, and calculates to the value of $\sin 2\beta$ most consistent with the data.

10.1.1 Other Fit Parameters

The kinematics of the backgrounds in the three data samples ($J/\psi K^+$, $J/\psi K^{*0}$, and $J/\psi K_S^0$) determine their mass dependencies. The prompt backgrounds all have pos-

Term	$J/\psi K^0_S$	$J/\psi K^{*0}$	$J/\psi K^+$
Background Mass Slopes			
$\zeta_P\,(\times 10^{-3})$	9.4 ± 3.0	4.4 ± 2.4	11.0 ± 0.9
$\zeta_L\,(\times 10^{-3})$	-6.6 ± 5.9	-17.6 ± 5.3	-12.3 ± 2.7
Event Fractions			
f_B	0.117 ± 0.010	0.156 ± 0.009	0.067 ± 0.003
f_L	0.262 ± 0.032	0.222 ± 0.023	0.160 ± 0.008
f_N	0.201 ± 0.037	0.096 ± 0.029	0.137 ± 0.014
$f_{\tau 2}$	0.550 ± 0.108	0.626 ± 0.104	0.781 ± 0.030
Long-lived Background "Lifetimes"			
τ_1 (μm)	423 ± 62	371 ± 65	595 ± 53
τ_2 (μm)	111 ± 21	99 ± 21	135 ± 9
Error Scale-Factors			
X	1.39 ± 0.11	1.54 ± 0.10	1.34 ± 0.05
Y	0.97 ± 0.03	1.06 ± 0.03	0.99 ± 0.01
Reconstruction Asymmetries			
R_B	–	-0.086 ± 0.068	0.077 ± 0.041
R_P	–	0.036 ± 0.029	0.003 ± 0.011
R_L	–	0.095 ± 0.070	0.030 ± 0.034

Table 10.1: First table of results from the likelihood fits in the $J/\psi K$ samples. This table contains all the results not related to the SST tags. The sign convention for the reconstruction asymmetries are such that positive R indicates an excess of reconstructed B^+ or B^0, compared to B^- or $\overline{B}^0$.

Term	$J/\psi K^0_S$	$J/\psi K^{*0}$	$J/\psi K^+$
Tagging Efficiencies			
ϵ_B	0.615 ± 0.041	0.635 ± 0.030	0.624 ± 0.020
ϵ_P	0.626 ± 0.017	0.830 ± 0.011	0.703 ± 0.005
ϵ_L	0.751 ± 0.034	0.778 ± 0.031	0.771 ± 0.014
Background Tagging Charge Biases			
α_P	-0.006 ± 0.045	0.092 ± 0.031	0.033 ± 0.013
α_L	0.167 ± 0.082	-0.044 ± 0.079	0.015 ± 0.037
Background Tagging Efficiency Asymmetries			
δ_P	–	0.012 ± 0.013	-0.002 ± 0.007
δ_L	–	-0.029 ± 0.038	-0.026 ± 0.018
Dilutions and $\mathcal{D} \sin 2\beta$			
$\mathcal{D}_0$, $\mathcal{D}_+$	–	0.165 ± 0.112	0.185 ± 0.052
$\mathcal{D}_0 \sin 2\beta$	0.306 ± 0.179	–	–
$\mathcal{D}_P$	–	-0.003 ± 0.031	-0.069 ± 0.013
$\mathcal{D}_L$	–	-0.050 ± 0.079	-0.089 ± 0.038

Table 10.2: Second table of results from the likelihood fits to the $J/\psi K$ samples. This table contains the results related to the SST tags. The sign convention for the tagging charge bias is that positive α indicates an excess of positive tags. The convention for the tagging efficiency asymmetry is that positive δ indicates higher efficiency for B^+ or B^0, compared to B^- or $\overline{B}^0$. The conventions are also chosen so that $\mathcal{D}_0$, $\mathcal{D}_+$, and $\mathcal{D}_0 \sin 2\beta$ are all expected to be positive. The convention for the background dilutions are that positive dilution indicates an excess of same-sign ($K^\pm$–$\pi^\pm$) versus opposite-sign ($K^\pm$–$\pi^\mp$) tags (where the K is the B daughter, and the π is the SST tag).

itive mass slopes (ζ), and the long-lived backgrounds all have negative mass slopes. This difference in kinematics between the prompt and long-lived backgrounds reinforces the conclusion that they are from different sources.

The fit for the $J/\psi K^0_S$ signal fraction is $f_B = 11.7\%$, indicating that 198 of the 1696 events in the sample are signal events. This 11.7% should not be considered a "signal purity", as it includes events in the sideband regions and in the signal-poor negative ct-region. Counting only the background events under the peak ($|M_N| < 3$), and only in the positive ct region, the signal purity is close to 50%. Most of this background is prompt; for $ct > 200\,\mu\text{m}$, the region most sensitive to the CP asymmetry, the signal purity is 85%.

The fit in the $J/\psi K^{*0}$ sample indicates that 365 of the 2339 events are signal, with slightly better actual purity than $J/\psi K^0_S$. For $J/\psi K^+$, the fit indicates that 846 of the 12,564 events are signal, and that the signal purity is roughly half that of the other two modes. The larger number of events in this sample compensates for the lower signal purity.

The background fractions (f_L, f_N, and $f_{\tau 2}$) are similar between the three decay modes, as are the background lifetimes (τ_1 and τ_2). The differences between these values for the three decay modes are due to the fact that each mode has a different number of tracks associated with the B decay vertex.

The scale-factors on the calculated uncertainties of the invariant masses (X) are similar between the three decay modes, ranging from ≈ 1.3 to ≈ 1.5. Recall that the covariance matrices (see Section 6.3.1) calculated for the CTC information on the track parameters is approximated using the number of hits collected for each track. Inaccuracies in these approximations cause these scale factors to be larger than 1. In contrast, the decay-length uncertainty scale-factors (Y) are all very close to 1. The decay-length measurement is dominated by SVX information, where the uncertainties on the track parameters are better estimated.

The reconstruction asymmetries for the $J/\psi K^0_S$ signal and background are all fixed to 1 (see Section 8.3). The reconstruction asymmetries for the signal events for the other two modes are both $\approx 8\%$, favoring $J/\psi K^+$ over $J/\psi K^-$ and $J/\psi \overline{K}^{*0}$ over

$J/\psi K^{*0}$. These results are only 1 to 2 standard-deviations significant, so it is not possible to conclude that the results cannot be merely statistical fluctuations. But, whether these asymmetries are real detector biases or just fluctuations, including them in the likelihood fit reduces the correlations between the dilutions and tagging charge asymmetries. The backgrounds in these modes all have reconstruction asymmetries consistent with zero.

The efficiencies for tagging signal events are all very similar: $\approx 62\%$. The tagging efficiencies for the long-lived backgrounds are also similar to one another, and are roughly 10% higher than those of the signal. Because the long-lived background events are expected to contain mostly real J/ψ's from incorrectly reconstructed B mesons, these events should often have missing B daughters to tag on, thus increasing the apparent tagging efficiency. The tagging efficiencies for the prompt backgrounds are rather different for the three decay modes. These differences are possibly due to differences in the combinatorial backgrounds between the three decay topologies.

As described in Section 9.3.1, there is a tagging charge bias, which is expected to be due to asymmetric track-reconstruction efficiencies for low p_T tracks. This asymmetry (α) should be the same for all three $J/\psi K$ signals, so their tagging charge biases are all fixed to the values measured in that section.

The background events in these samples are events where either there is no B meson, or where there is a misreconstructed B meson. Therefore, the tagging charge biases for these events might not be the same as those measured in Section 9.3.1, so the parameter α is allowed to float for the backgrounds. Most of the backgrounds have fit values for α consistent with the 1.9% average tagging charge bias measured for the signal events. The two exceptions are the long-lived $J/\psi K_S^0$ background (where $\alpha = 0.167 \pm 0.082$) and the prompt $J/\psi K^{*0}$ background (where $\alpha = 0.092 \pm 0.031$). Each of these is roughly 2 standard-deviations above the expected value. If there is some sort of systematic bias that causes these results to be what they are, then allowing them to float in the fit corrects for that bias. If they are merely statistical fluctuations, then that will be accounted for in the statistical uncertainties on the fit parameters.[1]

As there is no way to determine whether a $J/\psi K_S^0$ candidate was B^0 or $\overline{B}^0$ at the time of decay,[2] the convention chosen in Chapter 8 was to define all the events (both signal and background) to be B^0 at the time of decay. Thus, since there are no "$\overline{B}^0$ background," the background tagging efficiency asymmetries are not defined. The tagging efficiency asymmetries for the other backgrounds are expected to be small, as there is no reason they should not be. The fit results for these terms are all within one standard deviation of zero.

Similarly, the dilutions for the $J/\psi K_S^0$ backgrounds are not defined. The dilutions for the $J/\psi K^{*0}$ backgrounds are both consistent with zero, but both the background dilutions for $J/\psi K^+$ are significantly negative. This last result is expected from charge conservation: since the sum of the charges of the particles leaving an interaction point is zero, it is more likely that a track near a $J/\psi K^+$ ($J/\psi K^-$) background candidate will be negatively (positively) charged. Thus, there is expected to be an excess of OS tags for the $J/\psi K^+$ backgrounds.

10.2 Visual Interpretation

In order to allow for visual interpretation of the results of the likelihood fits, the data are binned in mass and in *ct*. Figures 6-7 to 6-9 in Chapter 6 show the data with $ct > 0$ and $ct < 0$, binned in units of M_N. The curves superimposed over these histograms are interpretations of the results of the likelihood fits.

In order to plot the results versus *ct*, it is convenient to separate the events into two sets: those with $|M_N| < 3$ are called the "signal region," and those with $3 < |M_N| < 20$ are called the "sideband region." The *ct* distributions and time-dependent asymmetries are plotted for both signal and sideband regions.

---

[1] The tagging charge bias for the background *could* have been constrained to the same value for the signal, but that would be an additional assumption, one which was not deemed necessary.

[2] Indeed, for the $J/\psi K_S^0$ signal, the flavor at decay is *not defined.* If it were possible to determine this flavor, then there could be no interference, and thus no CP violation in this decay.

10.2.1 Lifetime Description

The distribution of events in ct for both signal and sideband regions for each mode are shown in figures 10-1 to 10-3. The curves superimposed over the histograms indicate the results of the likelihood fits. The shapes of the distributions in all the plots agree very well with the distribution of events in the data.

As described in Section 9.1, the lifetimes of the signal events in the three modes are all fixed to the world-average B meson lifetimes [32]. If the lifetimes were allowed to float, the results would be $\tau_{B^+} = 482 \pm 19\,\mu\text{m}$ for $J/\psi K^+$, $\tau_{B^0} = 449 \pm 27\,\mu\text{m}$ for $J/\psi K^{*0}$, and $\tau_{B^0} = 396 \pm 35\,\mu\text{m}$ for $J/\psi K^0_S$. The first two are quite consistent with the world averages, $\tau_{B^+} = 486\,\mu\text{m}$ and $\tau_{B^0} = 468\,\mu\text{m}$, and the third is roughly 2 standard deviations low. However, in this case, the raw CP asymmetry changes only slightly, to $\mathcal{D}_0 \sin 2\beta = 0.30$, which indicates that the CP asymmetry is not highly correlated with the B^0 lifetime.

The value of the mixing frequency, Δm, is also fixed in $J/\psi K^{*0}$ and $J/\psi K^0_S$. Allowing it to float gives $\Delta m = 0.316 \pm 0.385\,ps^{-1}$ in $J/\psi K^{*0}$, and $\Delta m = 0.087 \pm 0.567\,ps^{-1}$ for $J/\psi K^0_S$. The statistical uncertainties on these measurements are very large and the results are certainly consistent with the input value, $\Delta m = 0.474$.

10.2.2 Dilutions and $\mathcal{D}_0 \sin 2\beta$

The breakdown of SST tags for the $J/\psi K^0_S$ sample is given in Table 10.3. The sideband-subtracted asymmetry, $(N_- - N_+)/(N_- + N_+)$, where N_- (N_+) is the number of events where the SST tag is on a negative (positive) track, is calculated for each bin in ct.[3]

These asymmetries are plotted in Figure 10-4. The solid curve in this figure shows the asymmetry shape returned by the likelihood fit, a sine-curve with magnitude of $\mathcal{D}_0 \sin 2\beta = 0.31$ (the fit result) and frequency $\Delta m = 0.474$ (an input parameter).

---

[3] The asymmetry is not calculated for the $ct < 0$ bin because the signal region there contains 106 events, of which ≈ 94 are expected to be background. The signal asymmetry in this bin is overwhelmed by background fluctuations.

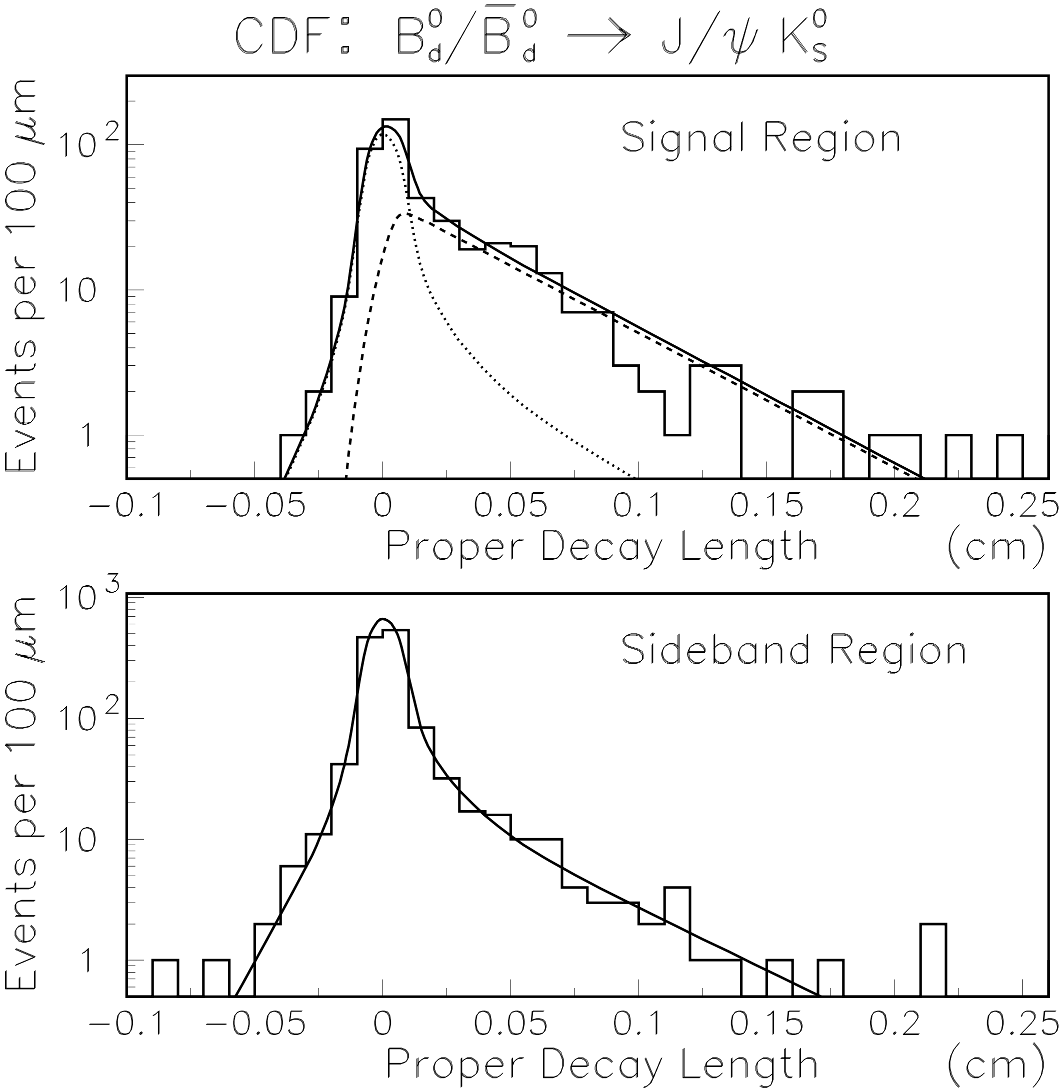

Figure 10-1: Histograms showing the binned distributions vs *ct* for the $J/\psi K_S^0$ sample. The top plot shows the signal region ($|M_N| < 3.0$) with the solid curve showing likelihood fit results. The breakdown into B-signal (dashed) and the background (dotted) contributions is also shown. The lower plot shows the sideband region ($3.0 < |M_N| < 20.0$) and the likelihood result for the the combined backgrounds.

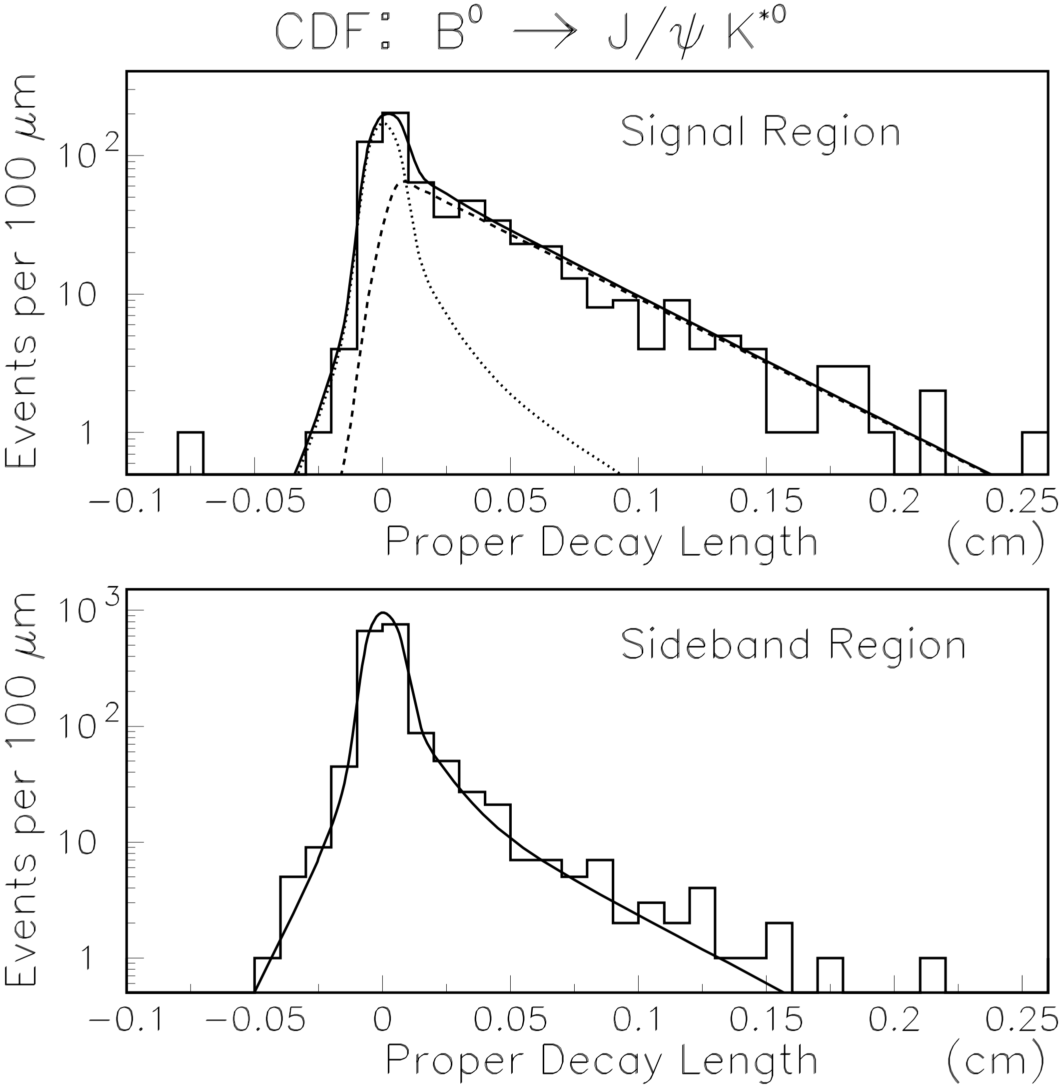

Figure 10-2: Histograms showing the binned distributions vs ct for the $J/\psi K^{*0}$ sample. The top plot shows the signal region ($|M_N| < 3.0$) with the solid curve showing likelihood fit results. The breakdown into B-signal (dashed) and the background (dotted) contributions is also shown. The lower plot shows the sideband region ($3.0 < |M_N| < 20.0$) and the likelihood result for the the combined backgrounds.

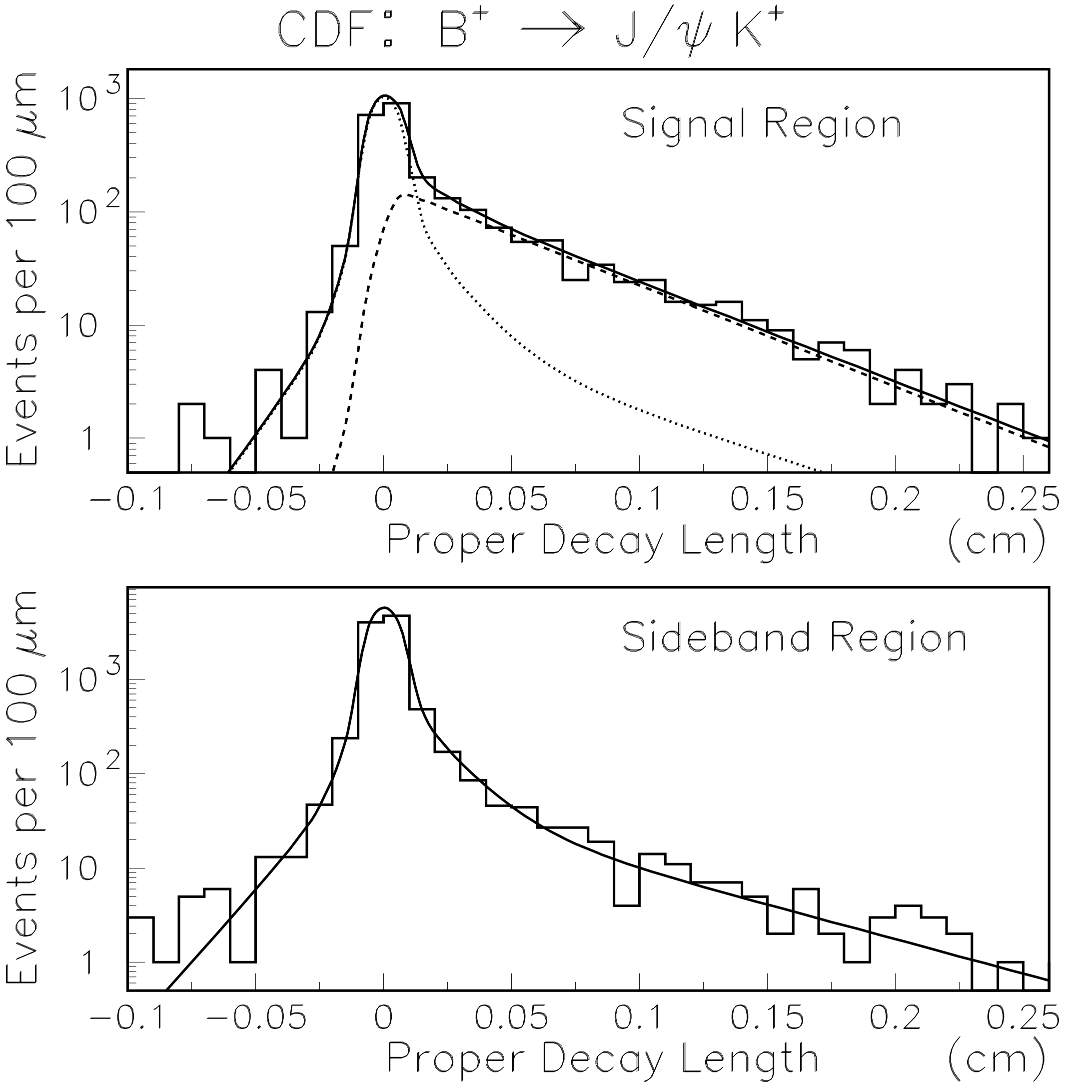

Figure 10-3: Histograms showing the binned distributions vs ct for the $J/\psi K^+$ sample. The top plot shows the signal region ($|M_N| < 3.0$) with the solid curve showing likelihood fit results. The breakdown into B-signal (dashed) and the background (dotted) contributions is also shown. The lower plot shows the sideband region ($3.0 < |M_N| < 20.0$) and the likelihood result for the the combined backgrounds.

Proper Decay Length (μm)	Signal Region			Sideband Region			Asymmetry (%)
	$-$	$+$	0	$-$	$+$	0	
-200 - 0	42	21	43	167	193	174	—
0 - 100	53	48	49	156	175	205	20 ± 25
100 - 200	14	14	15	26	34	24	8 ± 32
200 - 400	12	18	19	17	22	10	-22 ± 24
400 - 800	26	13	22	11	18	11	42 ± 18
800 - 1400	6	4	9	6	6	2	25 ± 40
1400 - 2000	3	1	1	0	0	2	50 ± 43

Table 10.3: Tags for $J/\psi K_S^0$ candidates in proper decay length (ct) bins. The signal region is $|M_N| < 3$, and the sidebands are $3 < |M_N| < 20$. The "+," "−," and "0," headings are for positive, negative, and untagged events. The last column is the sideband-subtracted tagging asymmetry $(N_- - N_+)/(N_- + N_+)$.

As a simple check of the likelihood fit, the points in this figure are fit to $A \sin \Delta m t$ using a simple χ^2 fit (with Δm still fixed). The simple fit (dashed curve) returns $A = 0.36 \pm 0.19$, a value consistent with the one returned by the full likelihood fit (0.31 ± 0.18).

Figure 10-5 shows a similar portrayal of the fit results for $J/\psi K^{*0}$. Here, the asymmetry is $(N_{SS} - N_{OS})/(N_{SS} + N_{OS})$, where N_{SS} (N_{OS}) is the number of events where the SST tag matches (does not match) the flavor of the K in the $J/\psi K^{*0}$ decay. The solid curve shows the $\mathcal{D}_0 \cos \Delta m t$ oscillation indicated by the likelihood fit. A simple χ^2 fit to $A \cos \Delta m t$ gives $A = 0.19 \pm 0.11$, which is in good agreement with the results of the likelihood fit (0.17 ± 0.11).

Figure 10-6 shows a similar plot for $J/\psi K^+$, where no time-dependence is expected for the signal asymmetry. The asymmetry plotted is $(N_{OS} - N_{SS})/(N_{OS} + N_{SS})$, which is expected to be positive for $J/\psi K^+$ signal events. The solid line indicates the result of the likelihood fit, and the dashed line indicates a simple χ^2 fit to a flat line, which has amplitude $A = 0.18 \pm 0.06$, (compared to 0.18 ± 0.05 from the full likelihood fit).

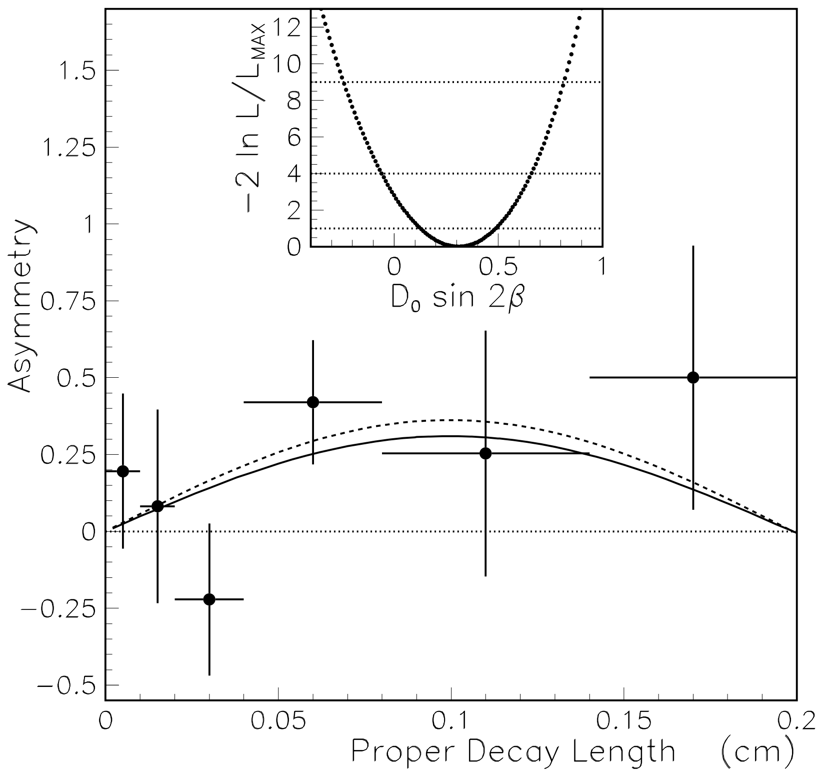

Figure 10-4: The likelihood fit results (solid curve) for the time dependent tagging asymmetry for $J/\psi K_S^0$ superimposed on the sideband subtracted data (points). The dashed curve is the result of a simple least squares fit of the binned (sideband subtracted) data to a constant times $\sin(\Delta m t)$. The inset shows a scan through the full likelihood function by varying $\mathcal{D}\sin(2\beta)$ and reminimizing.

The inset plots in these figures show scans of $-2\ln\mathcal{L}_{tot}$ versus $\mathcal{D}_0\sin 2\beta$, $\mathcal{D}_0$, and $\mathcal{D}_+$, respectively. The shapes are all parabolic, which indicates that the errors on $\mathcal{D}_0\sin 2\beta$, $\mathcal{D}_0$, and $\mathcal{D}_+$ are all Gaussian.

10.2.3 Raw Dilutions and CP Asymmetries

One disadvantage of figures like 10-4 to 10-6 is that the background has been *subtracted*, making it difficult to assess how much effect the background has on the signal measurement. To show the effect of the background, plots of the *raw* asymmetries (*i.e.* with no background subtraction) have also been made. Figures 10-7 and 10-

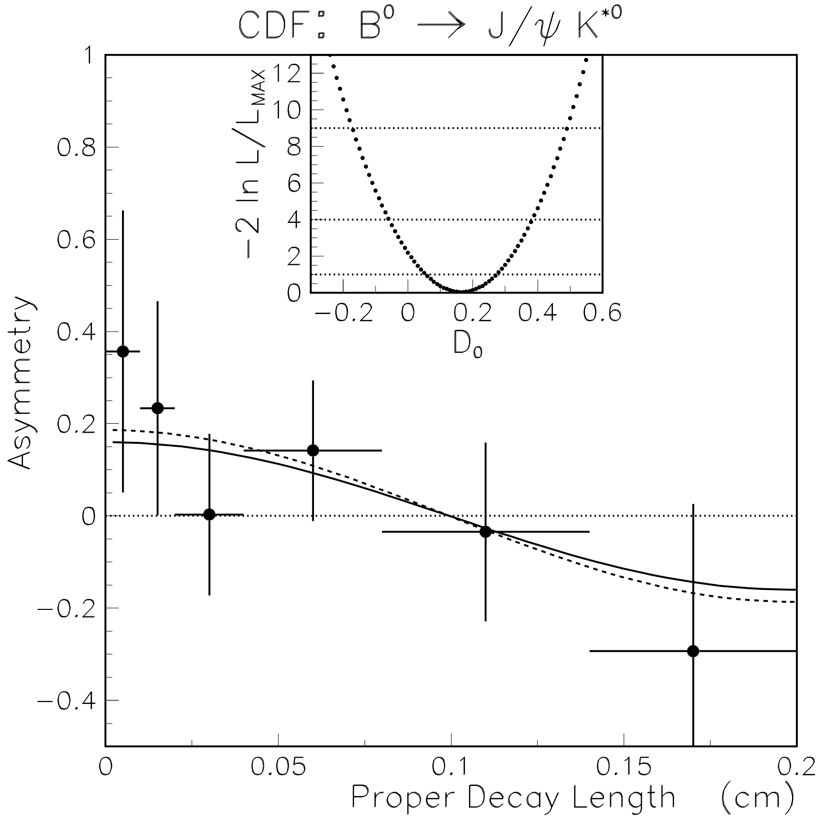

Figure 10-5: The likelihood fit results (solid curve) for the time dependent $(N_{SS} - N_{OS})/(N_{SS} + N_{OS})$ asymmetry for $J/\psi K^{*0}$ superimposed on the sideband subtracted data (points). The dashed curve is the result of a simple least squares fit of the binned (sideband subtracted) data to a constant times $\cos(\Delta m t)$ with Δm fixed. The inset shows a scan through the full likelihood function while varying the dilution, $\mathcal{D}_0$.

8 show the raw $(N_{SS} - N_{OS})/(N_{SS} + N_{OS})$ asymmetries for $J/\psi K^+$ and $J/\psi K^{*0}$. Figure 10-9 shows the raw $(N_+ - N_-)/(N_+ + N_-)$ asymmetry for $J/\psi K_S^0$. In this convention, the signal asymmetry is expected to be *positive* for $J/\psi K^{*0}$ and *negative* for the other two modes.

The top plot in each figure is the asymmetry in the signal region, *without* background subtraction, and the bottom plot is the asymmetry in the sidebands. The plots show the binned data, and the lines are the breakdown of the likelihood contributions.

Since these bins contain both signal and background events, the asymmetry in

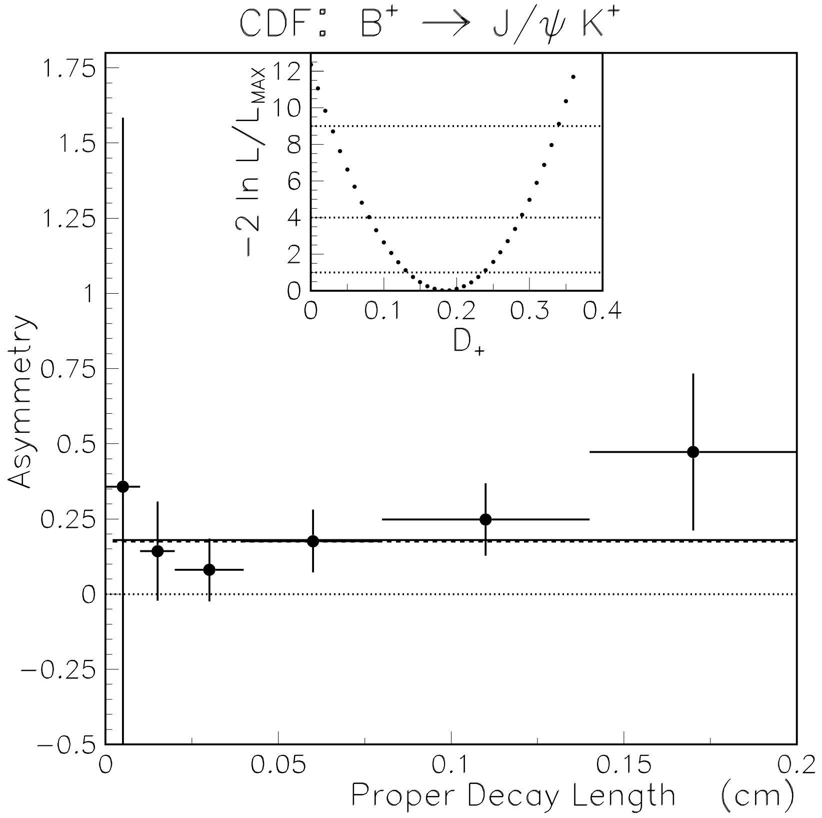

Figure 10-6: The likelihood fit results (solid curve) for the $(N_{OS} - N_{SS})/(N_{OS} + N_{SS})$ asymmetry for $J/\psi K^+$ superimposed on the sideband subtracted data (points). The dashed curve is the result of a simple least squares fit of the binned (sideband subtracted) data to a constant, $\mathcal{D}_+$. The inset shows a scan through the full likelihood function while varying the dilution, $\mathcal{D}_+$.

each bin is the weighted average of the asymmetries of the three types of events:

$$A \;=\; f_P\overline{A}_P + f_L\overline{A}_L + f_B\overline{A}_B \tag{10.1}$$

where f_ϕ is the fraction of events in that bin expected to be of type ϕ, and $\overline{A}_\phi$ is the average asymmetry predicted for events of type ϕ for that bin.

The parameters N_F, N_B, Y, τ_1, τ_2, f_L, f_N, $f_{\tau 2}$, ϵ_B, ϵ_L, and ϵ_P from the likelihood function determine what fraction of the tagged events in each bin are of each type. The dilution asymmetry in the $J/\psi K^{*0}$ signal and the CP-asymmetry in the $J/\psi K^0_S$ signal both vary over time and need to be averaged for each bin, but the asymmetries are not time-dependent for the other modes. These parameters then define the shapes

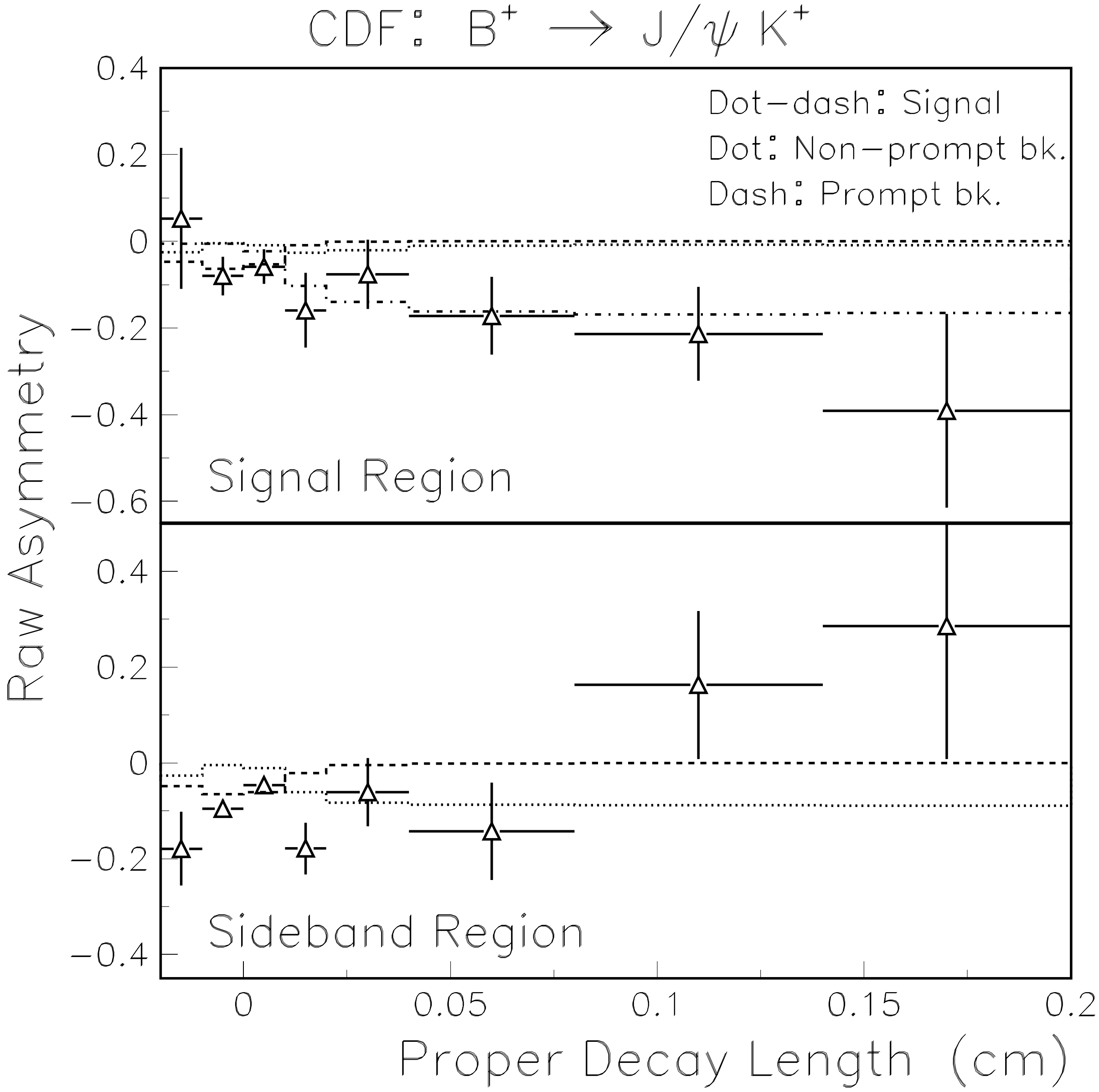

Figure 10-7: The raw dilution asymmetry for $J/\psi K^+$. The points are the raw (*not background subtracted*) asymmetries measured in the data. The lines indicate the bin-by-bin averages of the likelihood fit expectations for the signal (dot-dash), non-prompt background (dot), and prompt background (dash) contributions to the raw asymmetry. Positive asymmetry indicates an excess of SS vs. OS tags. There is expected to be an excess of OS for both signal (the SST effect) and for background (charge conservation).

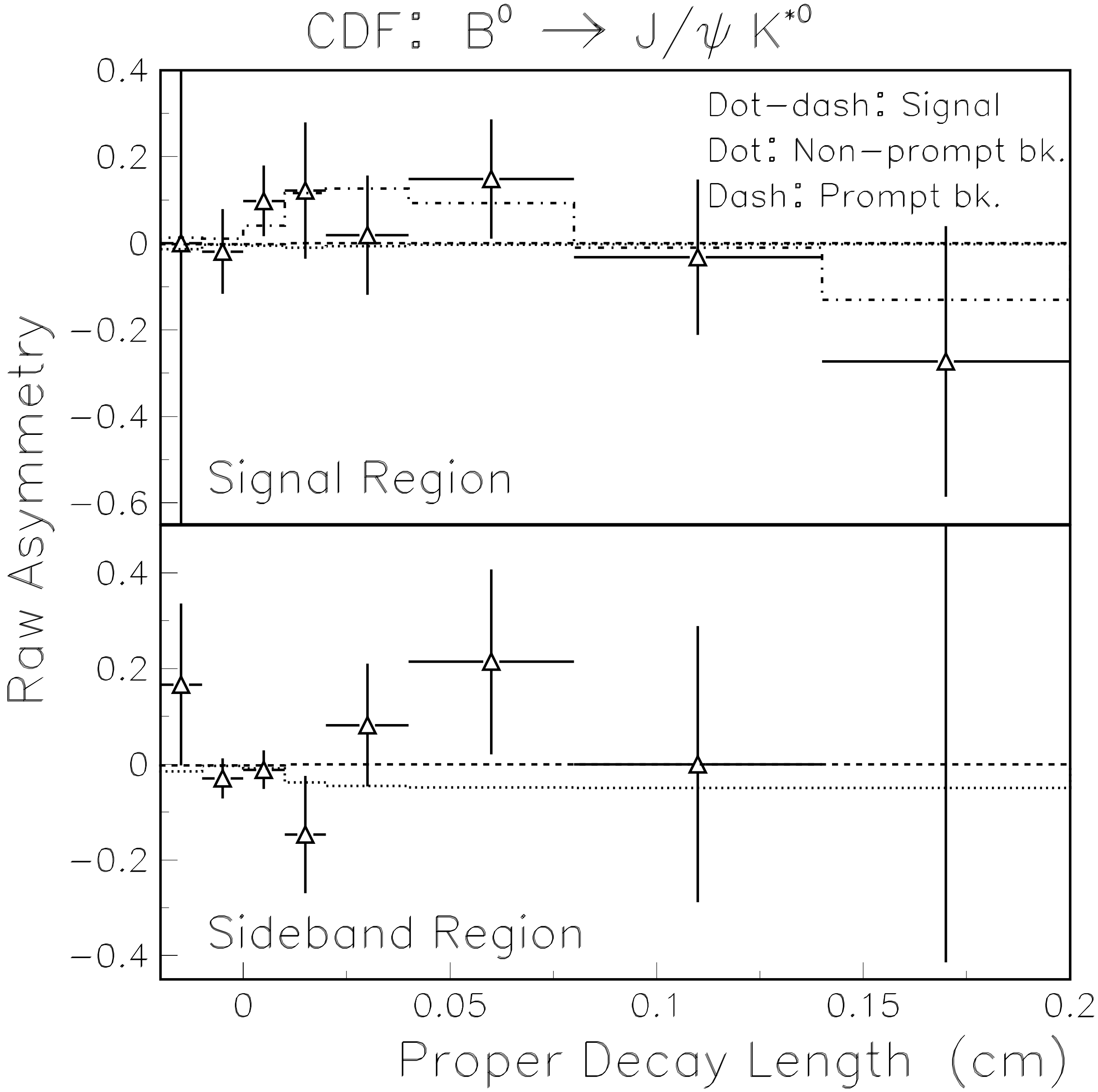

Figure 10-8: The raw dilution asymmetry for $J/\psi K^{*0}$. The points are the raw (*not background subtracted*) asymmetries measured in the data. The lines indicate the bin-by-bin averages of the likelihood fit expectations for the signal (dot-dash), non-prompt background (dot), and prompt background (dash) contributions to the raw asymmetry. Positive asymmetry indicates an excess of SS vs. OS tags. There is expected to be an excess of SS for the signal (times $\cos \Delta mt$), but no particular effect in background.

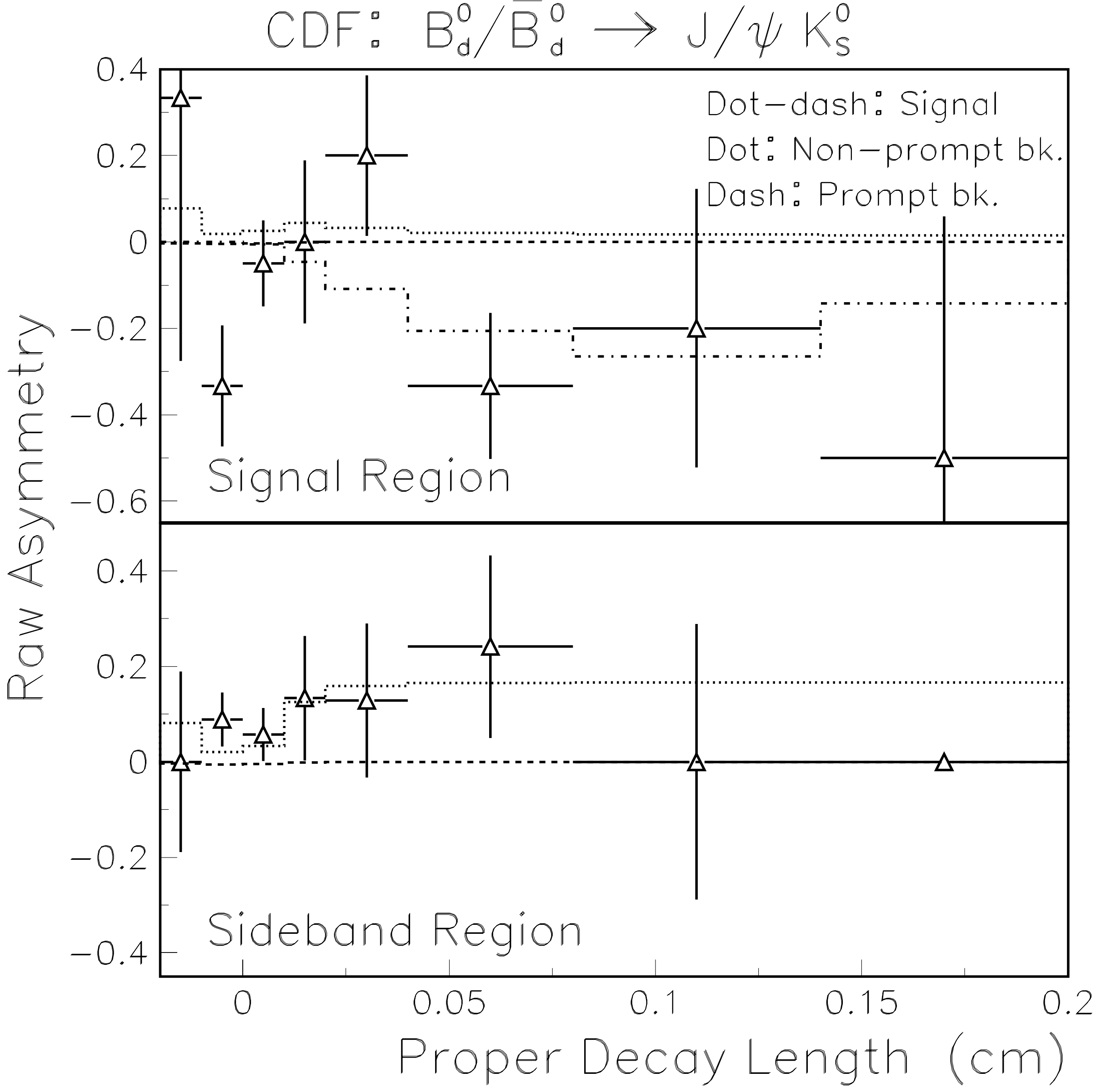

Figure 10-9: The raw tag charge asymmetry for $J/\psi K_S^0$. The points are the raw (*not background subtracted*) asymmetries measured in the data. The lines indicate the bin-by-bin averages of the likelihood fit expectations for the signal (dot-dash), non-prompt background (dot), and prompt background (dash) contributions to the raw asymmetry. Positive asymmetry indicates an excess of "+" vs. "−" tags. The CP-asymmetry is $-D_0 \sin 2\beta \sin \Delta mt$, so the negative asymmetry seen for the signal indicates positive $\sin 2\beta$.

of $f_\phi \overline{A}_\phi$ for each plot. In effect, the likelihood fit adjusts the normalizations of these three "curves" so that their sum best matches the raw asymmetries from the data. These normalizations are the measured dilutions and asymmetries.

For example, the bottom plot in Figure 10-7 shows the dilutions for the $J/\psi K^+$ sidebands. The shape of the non-prompt likelihood "curve" (dots) is fixed by the ct-dependence and fraction of long-lived background in the sample. This fraction is of course high at large ct, drops to nearly zero for $|ct| < 100\mu$m (where the data is dominated by prompt background), and increases again for large negative ct (the negative tail). The prompt curve (dashes), on the other hand, is zero except very near $ct = 0$. The two bins near zero therefore fix the normalization of the prompt background, and the other bins constrain the normalization of the nonprompt background. Both are significantly negative, which is not unexpected for the charged background, where charge conservation predicts more OS tags than SS.

The upper plot in Figure 10-7 shows the signal region for $J/\psi K^+$. The prompt curve has a very similar shape to that in the sidebands. The two bins near zero are still dominated by prompt background, and these points also help constrain the prompt dilution. However, the non-prompt background rises only in the far negative ct bin because the signal dominates the positive ct region; the five high-ct bins on this plot fix the normalization of the signal dilution. It is negative because the SST correlation prefers OS tags for charged B-mesons.

For the $J/\psi K^{*0}$, the backgrounds have much smaller dilutions. The nonprompt background dilution is driven by the five high-ct bins in the bottom of Figure 10-8. The result is about 1σ negative. The prompt dilution, on the other hand, is very nearly zero. The signal curve follows the familiar $\cos \Delta mt$ shape at high ct, but rolls over and drops to zero at low ct. This is due to the presence of prompt background, which dominates the low ct region. The signal curve and the data points agree remarkably well.

Figure 10-9 shows the asymmetry for $J/\psi K^0_S$. The non-prompt background asymmetry is very clearly positive. This corresponds to the high value of 0.167 returned by the likelihood fit. Again, the shape represents the f_L dependence on ct. While this offers no insight into what causes the large asymmetry in the long-lived background, it does show why the fit converges to such a large number: three bins in the high *ct* sidebands have asymmetries near 15%.

The signal curve includes two effects, the time-dependence of f_B, which increases with *ct*, and the variation of the CP-asymmetry as $\sin \Delta mt$. The normalization of the signal curve is clearly set by the four high *ct* bins. Below 200 μm the S/N and $\sin \Delta mt$ are both small, so those events have very little weight in the $\mathcal{D}_0 \sin 2\beta$ measurement. This curve also gives some insight into the magnitude of the effect of the large sideband asymmetry: the nonprompt curve is visible in the signal region, but it is very small at high *ct*. Even a large background asymmetry will have little effect on the measured signal asymmetry.

10.3 Summary of Fit Results

All of the plots in this chapter show curves which agree well with the data, indicating that the model used to generate the likelihood function is appropriate for these data sets. Simple least-squares fits to the background-subtracted asymmetries give results very similar to those of the likelihood fits, indicating that the fits are well-behaved.

Figures 10-7 to 10-9 show that the results of the fit for the signal asymmetries are dominated by the events in high S/N regions. The $J/\psi K^{*0}$ sample is affected most by the background, because the events with high asymmetry ($\cos \Delta mt$ near 1) are in the high-background low-*ct* region. The effect of the background is lessened for $J/\psi K_S^0$ because the signal events with large asymmetry ($\sin \Delta mt$ near 1) are in the high-*ct* region, which has a small contribution from the background.

Furthermore, these figures also indicate that it is the fraction of background present that matters, and not so much what the inherent asymmetries of the backgrounds are. The effect of the large (16%) tagging asymmetry in the long-lived $J/\psi K_S^0$ background on the measured signal asymmetry is small, as indicated by the top plot in figure 10-9. Figure 10-10 shows the normalized mass distribution for $J/\psi K_S^0$ events with $ct > 200\,\mu$m. The S/N in this region is greater than 6, indicating that there is very little background under the $J/\psi K_S^0$ signal in the region where the CP asymmetry is large. Thus, the $\mathcal{D}_0 \sin 2\beta$ measurement is dominated by asymmetries in real $B^0 \rightarrow J/\psi K_S^0$ events, not by fluctuations in the background.

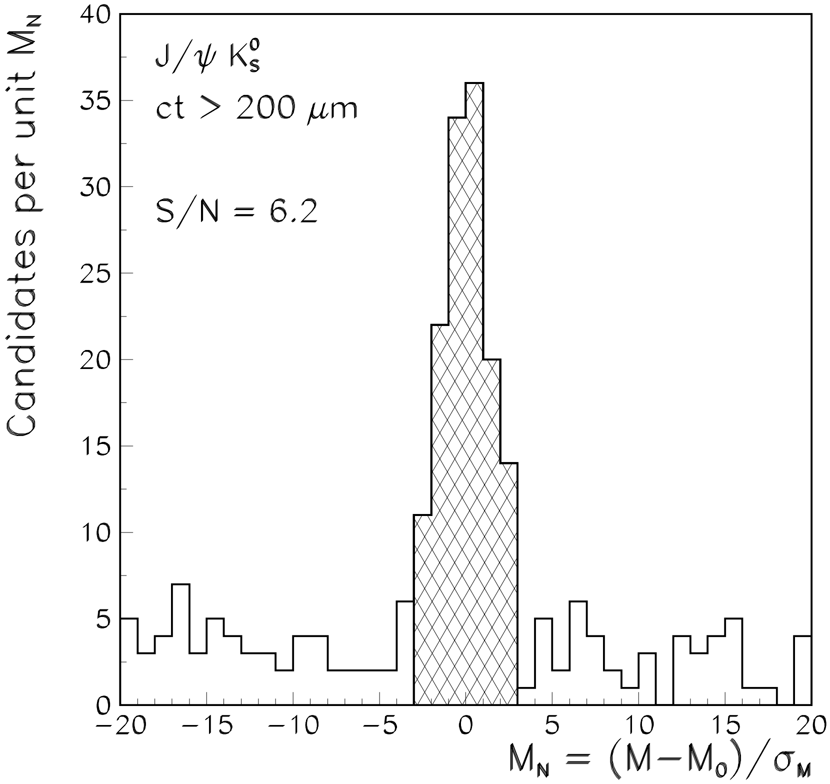

Figure 10-10: Normalized mass distribution for $J/\psi K_S^0$ events with $ct > 200\mu$m. There are 137 events with $|M_N| < 3$, of which 19 are estimated to be background and are ~ 118 estimated to be signal. The ratio of signal to background is $S/N = 6.2$.

Chapter 11

Evaluation of Systematic Uncertainties

This chapter discusses the evaluation of systematic effects which might bias the fit results. In Section 11.1, the determination of systematic uncertainties which arise from the uncertainties on the input parameters is discussed. In Section 11.2, systematic effects due to the presence of a "satellite peak" background is considered. In Section 11.3, the selection criteria used in this analysis are checked to determine if they introduce any systematic biases.

11.1 Uncertainties on the Fixed Parameters

In this section, the systematic uncertainties on the fit parameters arising from the uncertainties on the input parameters are calculated. The input parameters, their values, and their uncertainties are listed in table 9.1, which is reproduced here as table 11.1.

The systematic uncertainties from these parameters are determined by shifting each parameter up or down by the amount specified, remaximizing the likelihood, and observing the shift in the parameter of interest. The parameters of interest are $\mathcal{D}_0 \sin 2\beta$ from $J/\psi K^0_S$, $\mathcal{D}_0$ from $J/\psi K^{*0}$, and $\mathcal{D}_+$ from $J/\psi K^+$.

Parameter	Central Value	Variation
B decay Parameters		
τ_{B^0} (μm)	468	± 18
τ_{B^+} (μm)	486	± 18
Δm (ps^{-1})	0.474	± 0.031
$K\pi$ Swapping Parameters		
P_S	0.1	± 0.1
μ_S	-0.5	± 0.5
X_S	5.0	± 2.0
Tagging Charge Bias		
a_1 $\times 10^{-4}\,(\mathrm{GeV/c})^4$	3.9	± 1.8
b_1 $\times 10^{-3}\,(\mathrm{GeV/c})^4$	1.3	± 0.4
a_2 $\times 10^{-2}$	1.4	± 0.4
b_2 $\times 10^{-2}$	2.6	± 0.8
Tagging Efficiency Bias		
γ/α	1.0	$+1.5/-1.0$

Table 11.1: Input parameters for the likelihood fit and their uncertainties. This table is identical to table 9.1.

11.1.1 B Meson Decay Parameters

The B meson decay parameters fixed in the fits are τ_{B^0}, τ_{B^+}, and Δm. The systematic uncertainties due to these parameters are listed in table 11.2. The most significant shift is the shift in $\mathcal{D}_0 \sin 2\beta$ when Δm is varied. The other shifts are all on the order of 1/10th of 1%, too small to be of concern.

Parameter	Downward	Upward
$J/\psi K^0_S$:	Shift in $\mathcal{D}_0 \sin 2\beta$	
τ_{B^0}	-0.0010	$+0.0009$
Δm	$+0.0291$	-0.0252
$J/\psi K^{*0}$:	Shift in $\mathcal{D}_0$	
τ_{B^0}	-0.0002	$+0.0002$
Δm	-0.0005	$+0.0003$
$J/\psi K^+$:	Shift in $\mathcal{D}_+$	
τ_{B^+}	-0.0012	$+0.0011$

Table 11.2: Systematic uncertainties due to the uncertainties on the input parameters regarding B meson lifetimes and the B^0 mixing frequency. The "Downward" ("Upward") columns represent 1–σ downward (upward) shifts in the input parameters.

11.1.2 Tagging Efficiency Bias

The ratio γ/α of the tagging efficiency bias to the tagging charge bias is constrained, as described in Section 9.3.2. The central value used is 1, and the parameter ranges from 0 to 2.5. Table 11.3 shows the systematic uncertainties associated with these variations. The shifts are all $\sim$ 1% or less.

Parameter	Downward	Upward
$J/\psi K_S^0$:	Shift in $\mathcal{D}_0 \sin 2\beta$	
γ/α	$+0.0054$	-0.0101
$J/\psi K^{*0}$:	Shift in $\mathcal{D}_0$	
γ/α	$+0.0031$	-0.0057
$J/\psi K^+$:	Shift in $\mathcal{D}_+$	
γ/α	$+0.0015$	-0.0034

Table 11.3: Systematic uncertainties due to the constraint on γ/α. The "Downward" ("Upward") columns represent 1–σ downward (upward) shifts in the input parameters.

11.1.3 Tagging Charge Bias

The tagging charge bias, α, is described by the parameters a_1, b_1, a_2, and b_2, as described in Section 9.3.1. The uncertainties on these parameters create systematic uncertainties on the fit parameters. Because γ/α is constrained, the bias on the tagging efficiency and the bias on the tagging charge are correlated. Thus, the systematic effect due to the uncertainties on the parameterization of the tagging charge bias depends on the assumed value of γ/α. For this reason, the systematics are calculated separately for each value of γ/α: 0, 1, and 2.5, and the largest of the three shifts is taken as the systematic uncertainty. Table 11.4 shows the systematic uncertainties from these variations. The quantities used as systematic uncertainties are indicated in bold.

As an additional test, a *residual tagging asymmetry* α' is added to the likelihood function. With this parameter, the tagging charge bias for $J/\psi K^{*0}$ and $J/\psi K^+$ signal events is $\alpha(p_T(\text{tag}), n_{PI}) + \alpha'$, where α' is allowed to float freely.[1] If the parameteri-

---

[1] This term is not included in $J/\psi K_S^0$, as this asymmetry would be indistinguishable from the CP asymmetry. In the other two modes, where the B decay flavor is known, the residual tagging asymmetry can be separated from the dilution.

Parameter	$\gamma/\alpha = 0$		$\gamma/\alpha = 1$		$\gamma/\alpha = 2.5$	
	Downward	Upward	Downward	Upward	Downward	Upward
$J/\psi K^0_S$:	Shift in $\mathcal{D}_0 \sin 2\beta$					
a_1	+0.0013	−0.0013	−0.0002	−0.0001	**−0.0026**	**+0.0024**
b_1	+0.0063	−0.0063	+0.0070	−0.0073	**+0.0085**	**−0.0087**
a_2	−0.0020	+0.0020	−0.0001	−0.0001	**+0.0037**	**−0.0043**
b_2	−0.0112	+0.0112	−0.0120	+0.0117	**−0.0127**	**+0.0122**
$J/\psi K^{*0}$:	Shift in $\mathcal{D}_0$					
a_1	−0.0008	+0.0008	−0.0008	+0.0007	**−0.0011**	**+0.0010**
b_1	−0.0004	+0.0004	+0.0024	−0.0025	**+0.0063**	**−0.0065**
a_2	−0.0002	+0.0001	+0.0013	−0.0014	**+0.0040**	**−0.0044**
b_2	−0.0002	+0.0001	−0.0039	+0.0038	**−0.0092**	**+0.0088**
$J/\psi K^+$:	Shift in $\mathcal{D}_+$					
a_1	+0.0002	−0.0003	−0.0001	+0.0001	**−0.0010**	**+0.0009**
b_1	**+0.0007**	**−0.0007**	+0.0003	−0.0003	−0.0005	+0.0004
a_2	−0.0004	+0.0003	+0.0001	−0.0002	**+0.0015**	**−0.0019**
b_2	**−0.0010**	**+0.0010**	−0.0005	+0.0005	+0.0005	−0.0008

Table 11.4: Systematic uncertainties in the measured asymmetries due to the uncertainties on the tagging charge bias parameterization. Because the tagging charge bias is correlated with the tagging efficiency bias, this systematic depends on the assumed value of γ/α. The uncertainty is therefore calculated separately for each value of $\gamma/\alpha = 0, 1, 2.5$. The "Downward" ("Upward") columns represent 1–σ downward (upward) shifts in the input parameters. The largest positive and negative shifts in each row are taken as the systematic uncertainties (indicated in bold) on the parameters of interest.

zation is appropriate, then α' should be zero for each sample. The results of the fits are:

$$B^0 \quad : \quad \alpha' = -0.0004 \pm 0.078 \tag{11.1}$$

$$B^+ \quad : \quad \alpha' = 0.035 \pm 0.052, \tag{11.2}$$

which are both consistent with zero.

11.1.4 Swapping Parameterization

The presence of $K\pi$ swapping in the $J/\psi K^{*0}$ samples biases the likelihood fit. This is corrected for in the likelihood function through the inclusion of a second signal term, one that contains a fraction P_S of the signal events and has a normalized mass distribution with width X_S and mean μ_S, as described in Section 8.5. The parameters P_S, X_S, and μ_S are determined in Section 9.2. The systematic uncertainties from this parameterization are shown in table 11.5. These are the dominant systematic uncertainties for $J/\psi K^{*0}$.

$J/\psi K^{*0}$:	Shift in $\mathcal{D}_0$	
Parameter	Downward	Upward
P_S	$+0.0086$	-0.0160
X_S	$+0.0111$	-0.0003
μ_S	-0.0008	$+0.0007$

Table 11.5: Systematic uncertainties due to the uncertainty in the $K\pi$ swapping parameterization for $J/\psi K^{*0}$. The "Downward" ("Upward") columns represent 1–σ downward (upward) shifts in the input parameters.

11.1.5 Combined Systematic Uncertainties

These systematic uncertainties are all combined in quadrature, and the combined asymmetries are listed in table 11.6. The dominant systematic uncertainty on $\mathcal{D}_0 \sin 2\beta$

is from the variation in Δm. The dominant systematic uncertainty on the measurement of $\mathcal{D}_0$ in $J/\psi K^{*0}$ is due to the $K\pi$ swapping, and that for the measurement of $\mathcal{D}_+$ in $J/\psi K^+$ is due to the tagging charge bias and the tagging efficiency bias. But, as shown in table 11.6, the systematic uncertainties are all much smaller than the corresponding statistical uncertainties. This indicates that the parameters of interest do not depend significantly on the assumed values for the input parameters.

		Systematic Uncertainty		Statistical
Sample	Parameter	Negative	Positive	Uncertainty
$J/\psi K_S^0$	$\mathcal{D}_0 \sin 2\beta$	-0.032	$+0.033$	± 0.179
$J/\psi K^{*0}$	$\mathcal{D}_0$	-0.021	$+0.018$	± 0.112
$J/\psi K^+$	$\mathcal{D}_+$	-0.004	$+0.003$	± 0.052

Table 11.6: Combined systematic uncertainties for the three modes.

11.2 "Satellite" Peak Background

When a B meson decays to $J/\psi K\pi$, it is possible to reconstruct the $J/\psi K$, ignore the pion, and consider it a $J/\psi K$ event. The mass of the misreconstructed B meson will not be correct, but these candidates will sometimes show up in the low-mass sideband. Since these events would not show up under the signal peak, including them causes the background under the peak to be slightly overestimated and the signal therefore to be slightly underestimated.

Figure 11-1 shows the normalized mass distribution for fully reconstructed $J/\psi K^{*0}$ Monte Carlo candidates (using the same Monte Carlo sample as that used in Section 9.2). Also plotted is the normalized mass distribution calculated when either the π is lost or the K is lost and the π is misidentified as a K. Approximately 10% of the misreconstructed candidates will show up in the lower-mass sidebands.

The partial width of $B \to J/\psi K^*$ is ≈ 1.5 times that of $B \to J/\psi K$, so for each signal event reconstructed, there should be 0.15 satellite events reconstructed.[2] Thus, roughly 5% of the long-lived $J/\psi K_S^0$ background should be satellite events (less for the other modes). This causes the number of signal events to be underestimated by $\approx 1\%$. Since the statistical uncertainties on both of these quantities are much larger than these numbers, this background is considered negligible and ignored.

One further concern is that $B^0 \to J/\psi K^{*0}$ events with $K^{*0} \to K_S^0 \pi^0$ might be CP eigenstates and thus contribute a time-dependent CP asymmetry to the $J/\psi K_S^0$ background. However, it is estimated that this background contributes a net asymmetry of less than 1 tag. This effect is also considered negligible and ignored.

11.3 Robustness Studies

To determine if the measured asymmetries are "real" or are merely artifacts of the particular choices of selection criteria used, *robustness* studies are performed. The p_T cutoff for the B mesons is varied with little noticeable effect. Likewise, the χ^2_{2dtr}, the χ^2_{2dpi}, and the $p_T(K^{(*)})$ cutoffs are varied, and no significant changes in the asymmetries are noted.

Additional robustness tests are performed by varying the selection criteria for SST tagging candidates. Varying the impact-parameter significance and ΔR cuts also has very little effect. Of all the cross-checks performed, only the variation of the $p_T(\mathrm{SST})$ cutoff was found to have a significant effect. Figure 11-2 shows the variations of the three parameters of interest with the p_T threshold for SST candidate selection. The charged dilution rises with $p_T(\mathrm{SST})$, while the neutral dilution shows no clear effect. The CP asymmetry in $J/\psi K_S^0$ rises slightly up to $p_T(\mathrm{SST}) = 500$ MeV/c, but then drops sharply, reaching a minimum at -0.12, for $p_T(\mathrm{SST}) = 700$ MeV/c.

One must be careful when attempting to interpret the information in these plots,

---

[2] Assuming that the reconstruction efficiency for the misreconstructed candidates is the same as that for the properly reconstructed candidates. This is a conservative assumption because the misreconstructed B meson will generally not have momentum pointing directly away from the primary vertex, and the pointing constraint will throw some of these events out. Thus the misreconstructed candidates should have lower reconstruction efficiency than the properly reconstructed ones.

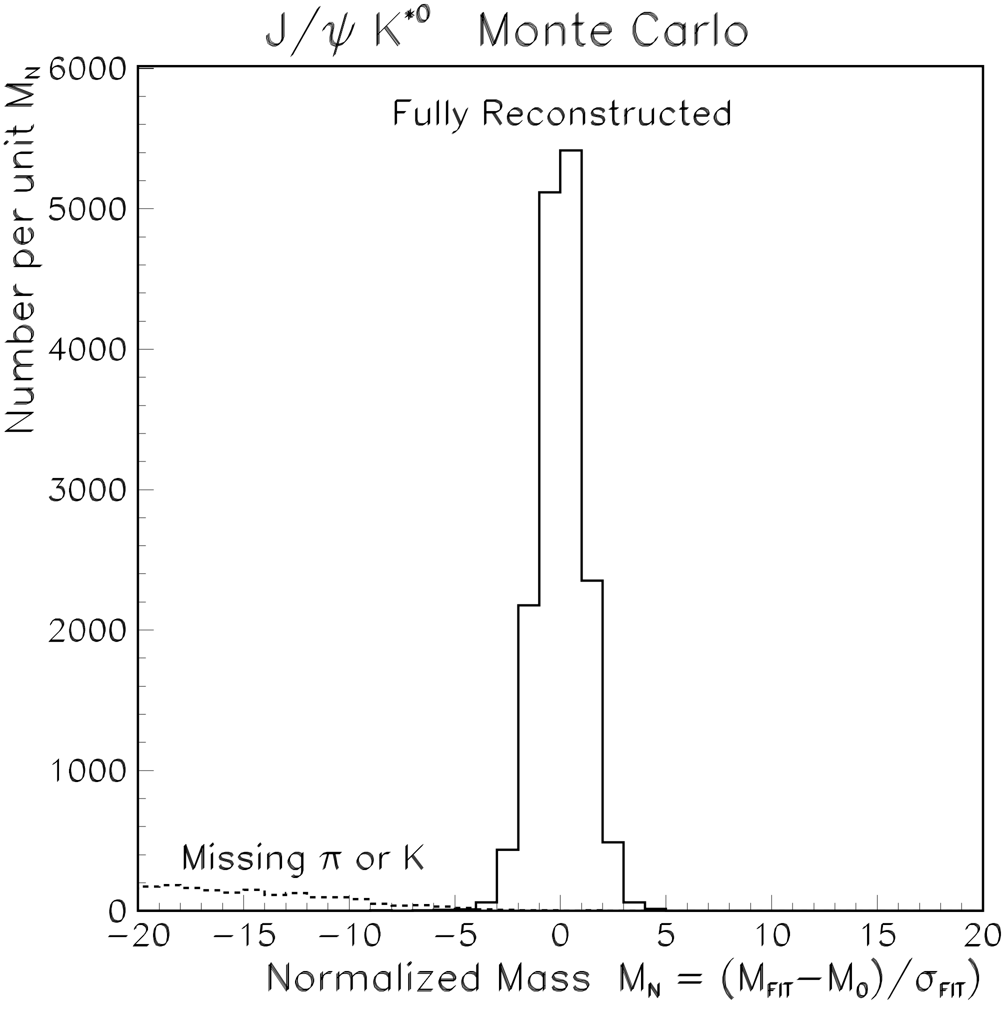

Figure 11-1: Normalized mass distribution for fully and partially reconstructed $J/\psi K^{*0}$ Monte Carlo events.

as the points are all *partially* correlated with one another. When the p_T(SST) cut is raised, any events which tagged on tracks in the range skipped lose their tags, but the rest of the events will remain unaffected. Some of the events which lose their tags will then select new tags at higher p_T. The charges of these new tags are not correlated with those of the tracks that are lost (the charges are actually slightly *anticorrelated*). Thus, points on these plots are partially correlated with one another. This partial correlation makes it impossible to judge the expected point-to-point variations from the information given.[3] This entire issue will be analyzed further in Section 11.3.2.

11.3.1 Comparison with Other Samples

This analysis uses the identical tagging algorithm to that used in [45], and the dilutions are expected to be similar for the $J/\psi K$ samples and the ℓ–D samples used in [45]. Therefore, the variations of the dilutions versus p_T(SST) are also measured for the ℓ–D samples, for comparison with the above results. The ℓ–D samples have much higher statistics than the $J/\psi K$ samples, so the statistical fluctuations in the ℓ–D samples should be smaller. In addition, Monte Carlo simulations of both the $J/\psi K$ and ℓ–D samples have been generated,[4] and the dilution variations are measured in the Monte Carlo samples, as well. Figure 11-3 shows the variations of the $\mathcal{D}_+$ measurements with p_T(SST) for all four samples, and figure 11-4 shows the variations of $\mathcal{D}_0$.

Two conclusions are drawn from these plots. First, the charged dilution rises with p_T(SST) for all four samples, while the neutral dilution stays relatively flat for all four samples. Thus, seeing a p_T(SST) dependence on $\mathcal{D}_+$ in $J/\psi K^+$, but none on $\mathcal{D}_0$ in $J/\psi K^{*0}$, is not an indication of a problem.

The next conclusion is that the shape of the $\mathcal{D}_+$ variation versus p_T(SST) does not agree very well with the shape predicted by the Monte Carlo simulation. Further studies are needed to determine if this difference indicates some sort of systematic

---

[3] This is not to say that these variations cannot be determined, merely that they cannot be determined from the information available in these specific plots.

[4] A description of the Monte Carlo samples used in this study can be found in Appendix A.

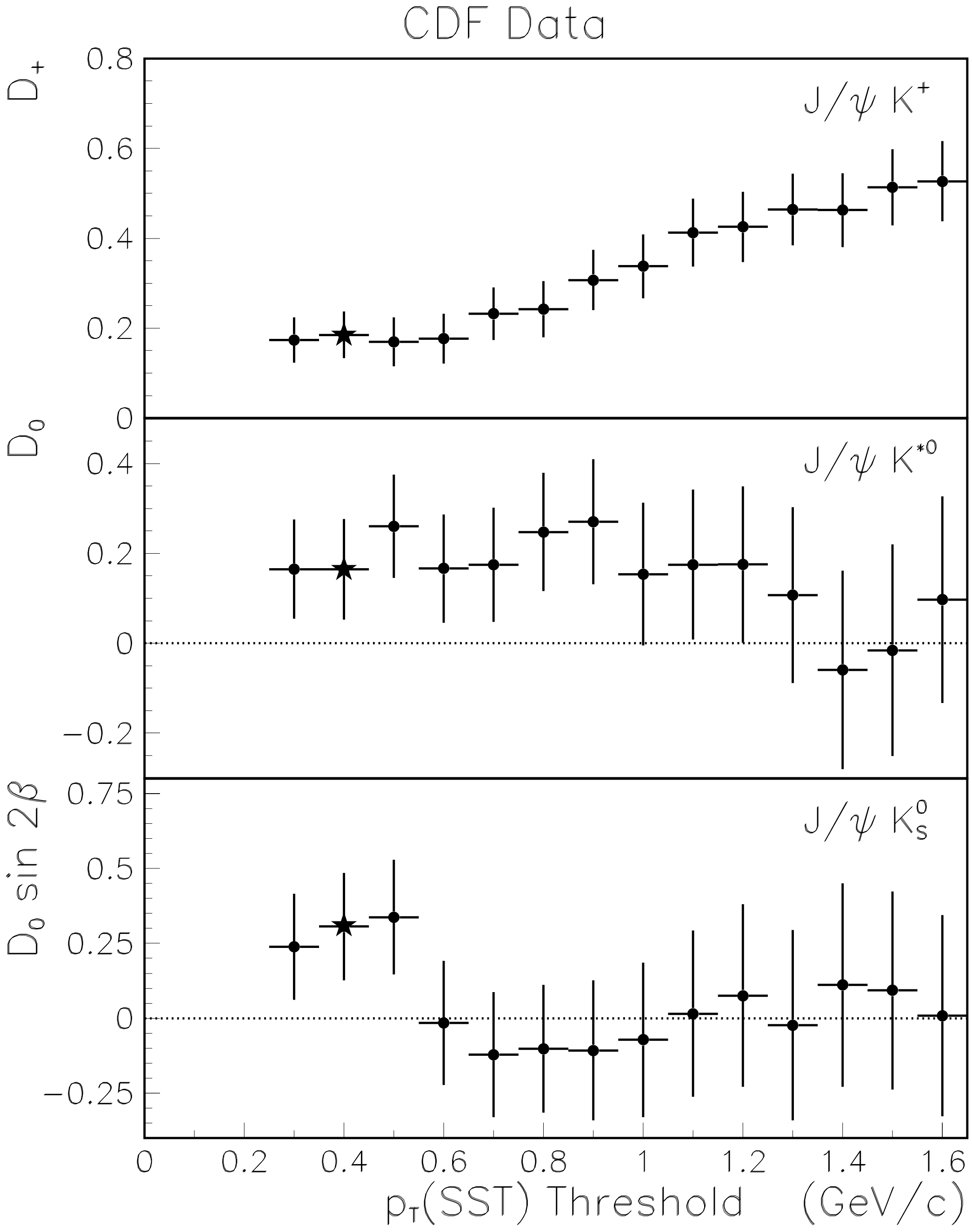

Figure 11-2: The variations of the measured asymmetries with the p_T(SST) cutoff. The nominal cutoff is indicated by the star at 0.4 GeV/c. The points on each plot are all partially correlated with one another, so it is not possible to judge the expected point-to-point variation from the information given.

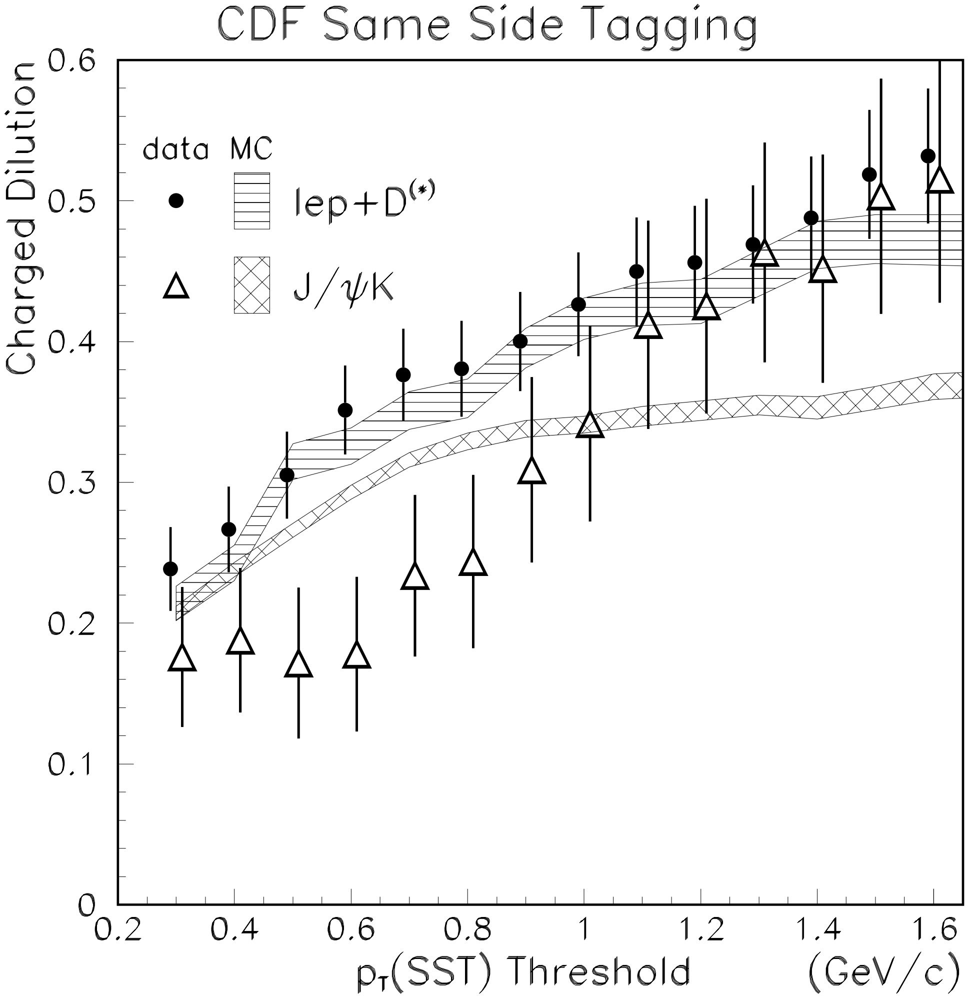

Figure 11-3: The variation of $\mathcal{D}_+$ with the p_T(SST) cutoff for $J/\psi K$ and ℓ–D data and Monte Carlo. The points for each sample are partially correlated with one another, so it is not possible to judge the expected point-to-point variation from the information given.

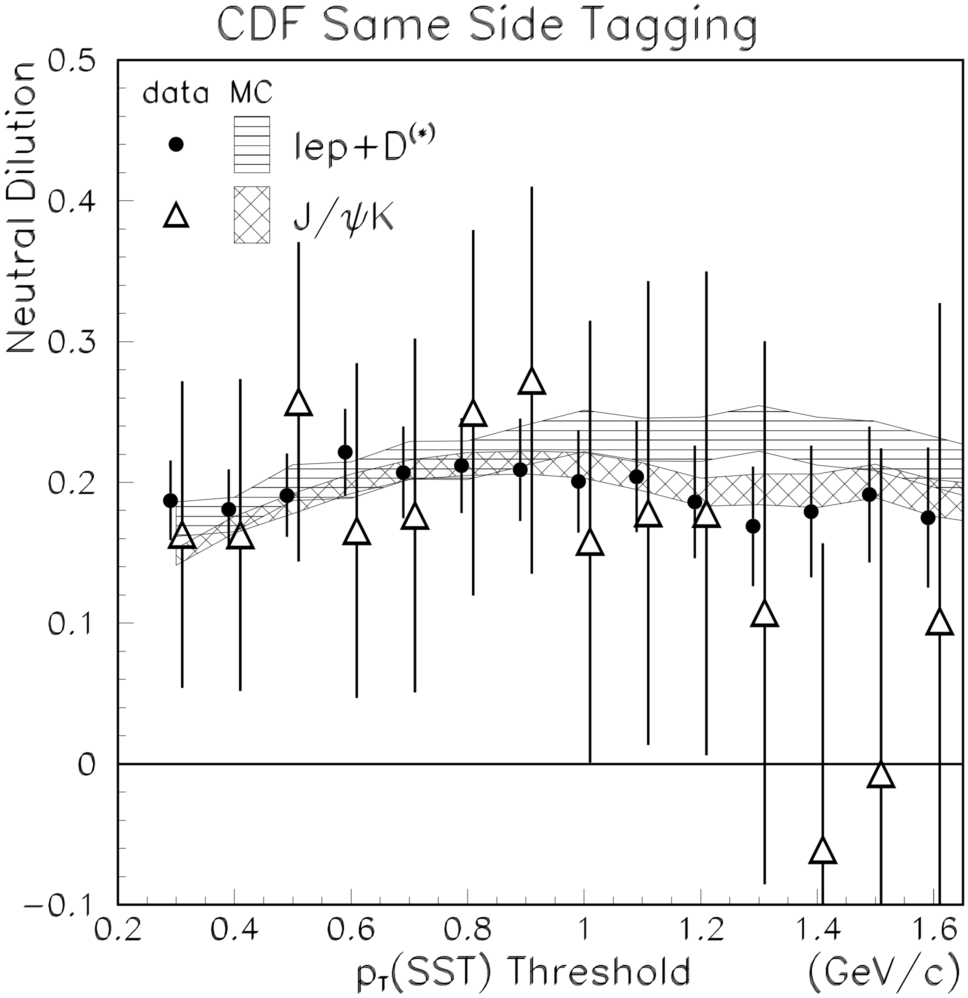

Figure 11-4: The variation of $\mathcal{D}_0$ with the p_T(SST) cutoff for $J/\psi K$ and ℓ–D data and Monte Carlo. The points for each sample are partially correlated with one another, so it is not possible to judge the expected point-to-point variation from the information given.

effect, or if it is consistent with being a statistical fluctuation.

A systematic variation in $\mathcal{D}_0 \sin 2\beta$ with p_T(SST) can only be due to changes in the effectiveness of the tagging algorithm (as $\sin 2\beta$ itself obviously cannot change). The flatness of $\mathcal{D}_0$ versus p_T(SST) for neutral B mesons (see figure 11-4) indicates that the variation cannot be due to variation of the neutral dilution. The p_T(tag) dependence of α seen in Section 9.3.1 indicates that the tagging charge bias will depend on p_T(SST), but the inclusion of this dependence in the likelihood function should counteract that effect. Certainly changes in the tagging charge bias cannot cause variations such as those seen in figure 11-2. Like with the $J/\psi K^+$, more study is needed to determine if the variation seen in the $J/\psi K^0_S$ data is a real effect or if it is just a statistical fluctuation. This is the subject of the following section.

11.3.2 Evaluation of the Anomalous Instabilities

Because the $\mathcal{D}_+$ measurement in $J/\psi K^+$ and the $\mathcal{D}_0 \sin 2\beta$ measurement in $J/\psi K^0_S$ both show unexpected variation with p_T(SST), these effects are studied further. The expected statistical fluctuations are estimated using Monte Carlo, where statistical and systematic effects can be separated.

The $J/\psi K^+$ Monte Carlo sample is divided into 100 subsamples, each of which has equivalent statistics to the $J/\psi K^+$ data. These samples are all then tagged, and the variation of the dilution in each versus p_T(SST) is studied. Figure 11-5 shows the variations of the first eight such samples. The solid curve in each plot indicates the average values (from the entire Monte Carlo sample), the dashed curve indicates the results from the data, and the points are the results of the individual pseudo-experiments. It appears that it is possible for a statistical fluctuation to account for the variation seen in the data.

The $J/\psi K^0_S$ Monte Carlo sample is divided into 200 subsamples, each of which has statistics equivalent to the $J/\psi K^0_S$ data. Figure 11-6 shows the variations of $\mathcal{D}_0 \sin 2\beta$ with p_T(SST) calculated for the first eight subsamples. Some of them show very little dependence, but some show large variations. This indicates that the variation seen in the data is consistent with that from a statistical fluctuation.

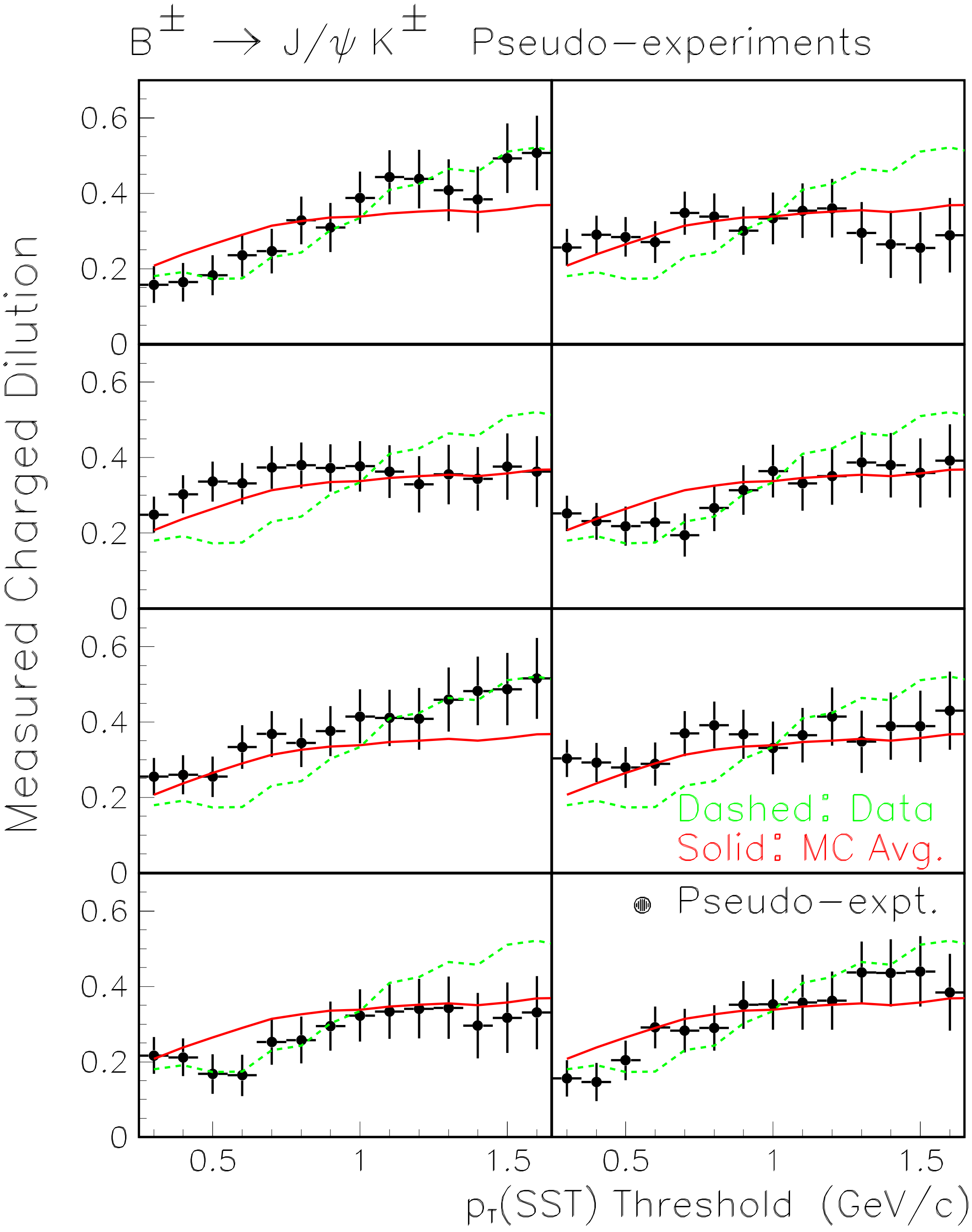

Figure 11-5: The variation of $\mathcal{D}_+$ with the p_T(SST) cutoff for the first eight $J/\psi K^+$ Monte Carlo subsamples. The solid curve indicates the average Monte Carlo values, and the dashed curve indicates the results from the data. The points are the results from individual Monte Carlo subsamples.

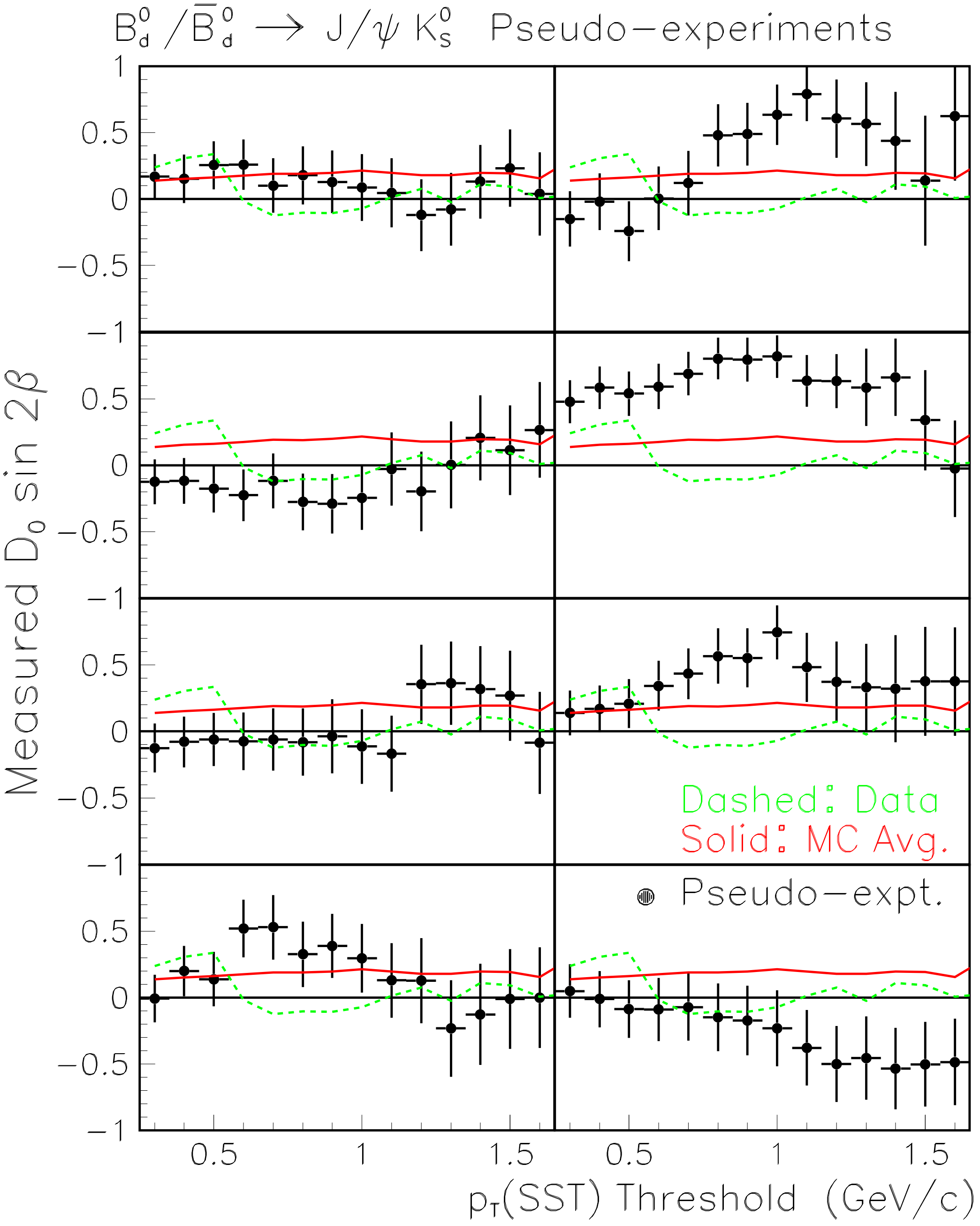

Figure 11-6: The variation of $\mathcal{D}_0 \sin 2\beta$ with the p_T(SST) cutoff for the first eight $J/\psi K_S^0$ Monte Carlo subsamples. The solid curve indicates the average Monte Carlo values, and the dashed curve indicates the results from the data. The points are the results from individual Monte Carlo subsamples.

Quantitative Analysis of the Variations

The variations in both $J/\psi K^+$ and $J/\psi K_S^0$ appear to be *qualitatively* consistent with statistical fluctuations. To perform a *quantitative* test, a χ^2 comparison is performed.

Because the points are correlated with one another, it is not useful to simply perform a χ^2 comparison between the measured asymmetries and the average results from the Monte Carlo. However, the shifts between consecutive points on these plots are less correlated,[5] so they can be used to calculate a meaningful χ^2. The top plots in figures 11-7 and 11-8 show the variations in $J/\psi K^+$ and $J/\psi K_S^0$ (and are the same points as those in figure 11-2). The bottom plots show the shifts between consecutive points and the statistical uncertainties on these shifts.

Using the shifts between consecutive points, the χ^2 is calculated:

$$\chi^2 \quad = \quad \sum_{i=-.3}^{1.5} \left(\frac{\delta_i - \overline{\delta}_i}{\sigma_i} \right)^2 \tag{11.3}$$

where $\delta_i = A_{i+1} - A_i$, A_i is the asymmetry for $p_T(\text{SST}) = i$ GeV/c, $\overline{\delta}_i$ is the Monte Carlo average value for δ_i, and σ_i is the statistical uncertainty on δ_i. The term $\overline{\delta}_i$ contains the expected systematic shift between consecutive points; it is subtracted out to separate the statistical and systematic effects from one another.

This χ^2 is calculated for the data and for each of the Monte Carlo subsamples. As it contains 13 terms, the distribution of Monte Carlo values should be consistent with a χ^2 distribution with 13 degrees of freedom. The histogram in figure 11-9 is the distribution of χ^2 values from the $J/\psi K^+$ Monte Carlo subsamples. This histogram is fit to the χ^2 distribution function:

$$f(z;n) \quad = \quad N \frac{z^{n/2-1} e^{-z/2}}{2^{n/2}, \, (n/2)}, \tag{11.4}$$

where N is the absolute normalization (number of samples), and n is the number of

---

[5] The shift in measured asymmetry that comes when the p_T(SST) cut is raised from 400 to 500 MeV/c is determined by those events which tag on tracks with $400 < p_T < 500$ MeV/c. Likewise, the shift that comes from raising the cut from 500 to 600 MeV/c is determined by the events which tag on tracks with $500 < p_T < 600$ MeV/c. These two sets of events have small overlap, so the two shifts are nearly uncorrelated.

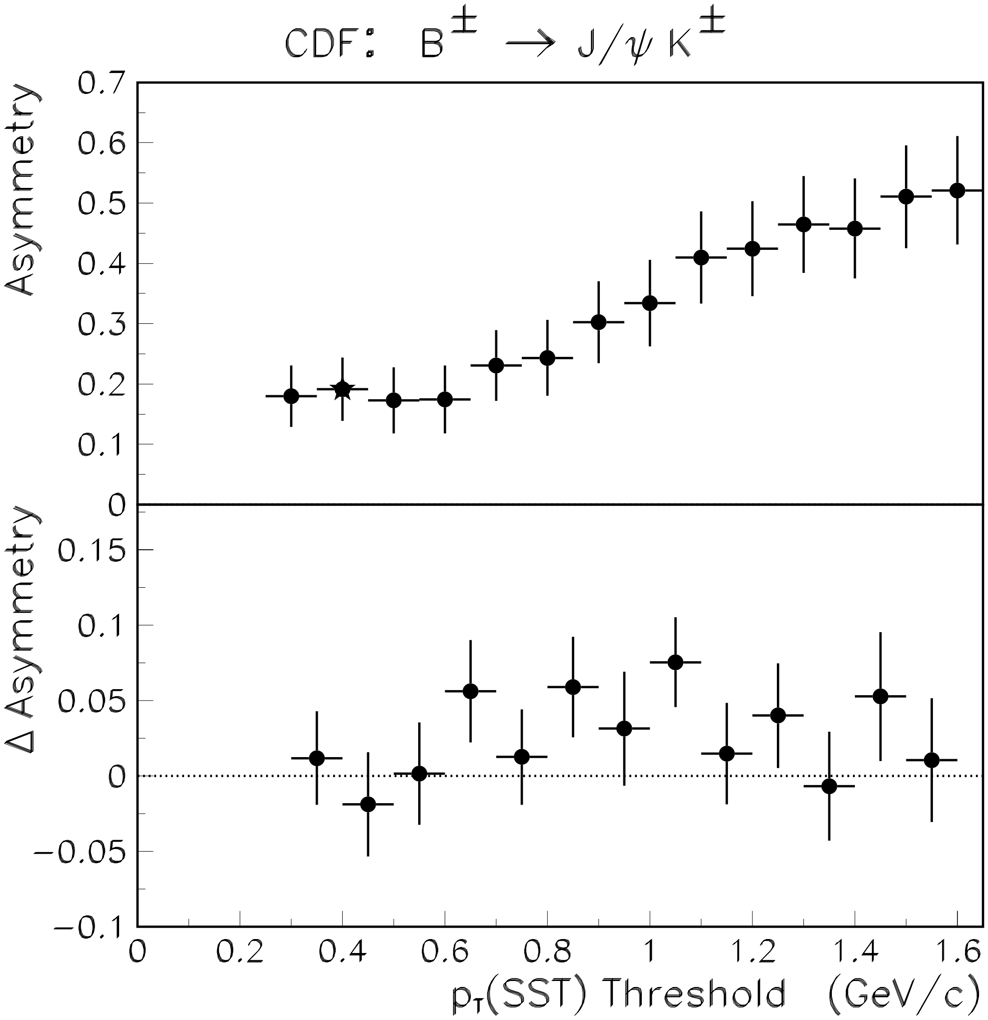

Figure 11-7: The variation of the measured CP asymmetry $\mathcal{D}_0 \sin 2\beta$ with the p_T(SST) cutoff. The nominal value is the star at 0.4 GeV/c. The bottom plot shows the inter-point shifts, with the uncorrelated statistical uncertainties on these shifts.

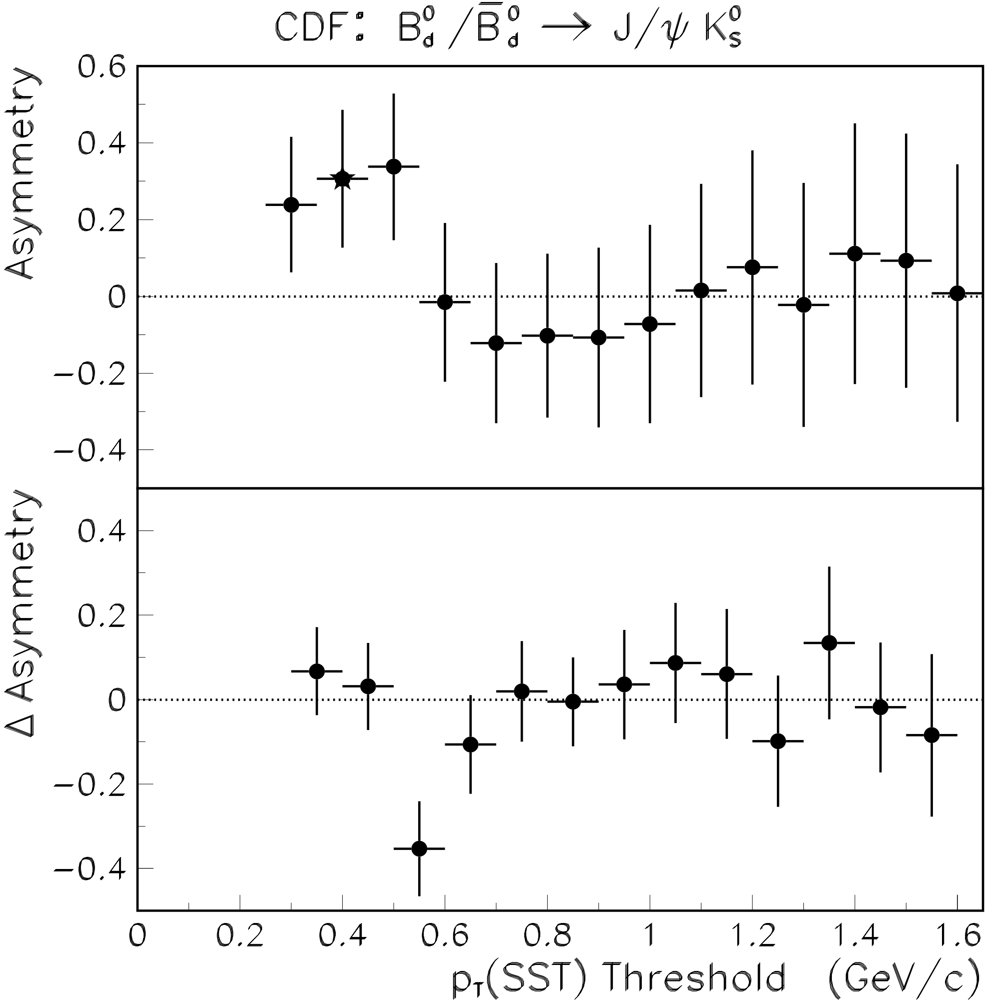

Figure 11-8: The variation of the measured CP asymmetry $\mathcal{D}_0 \sin 2\beta$ with the p_T(SST) cutoff. The nominal value is the star at 0.4 GeV/c. The bottom plot shows the inter-point shifts, with the uncorrelated statistical uncertainties on these shifts.

degrees of freedom.[6] The fit gives $N = 100$ and $n = 12.7$, quite consistent with the expected values. The data has $\chi^2 = 16.7$, and 22% of the Monte Carlo subsamples have higher χ^2 than the data.

Figure 11-10 shows the equivalent distribution for the $J/\psi K_S^0$ Monte Carlo. The fit returns $N = 200$, and $n = 12.4$, indicating the χ^2 formulation behaves as expected. The data has $\chi^2 = 13.0$, and 42% of the Monte Carlo subsamples have higher χ^2 than the data.

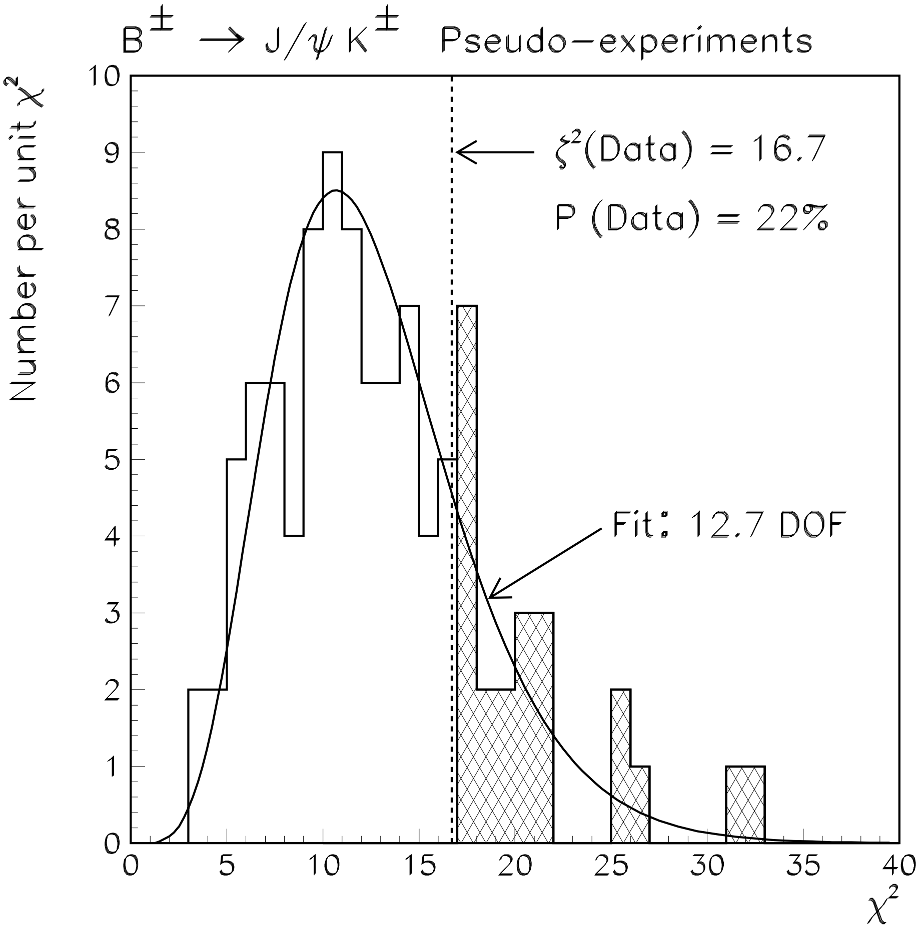

Figure 11-9: The distribution of χ^2 returned by the $J/\psi K^+$ Monte Carlo samples. The fit is to $f(z;n) = z^{n/2-1}e^{-z/2}/2^{n/2}$, $(n/2)$, which corresponds to the standard χ^2 with n degrees of freedom.

---

[6] While, technically, the meaning of a χ^2 distribution with non-integral number of degrees of freedom is not clear, $f(z;n)$ is just a function, which is well-defined for non-integer values of n.

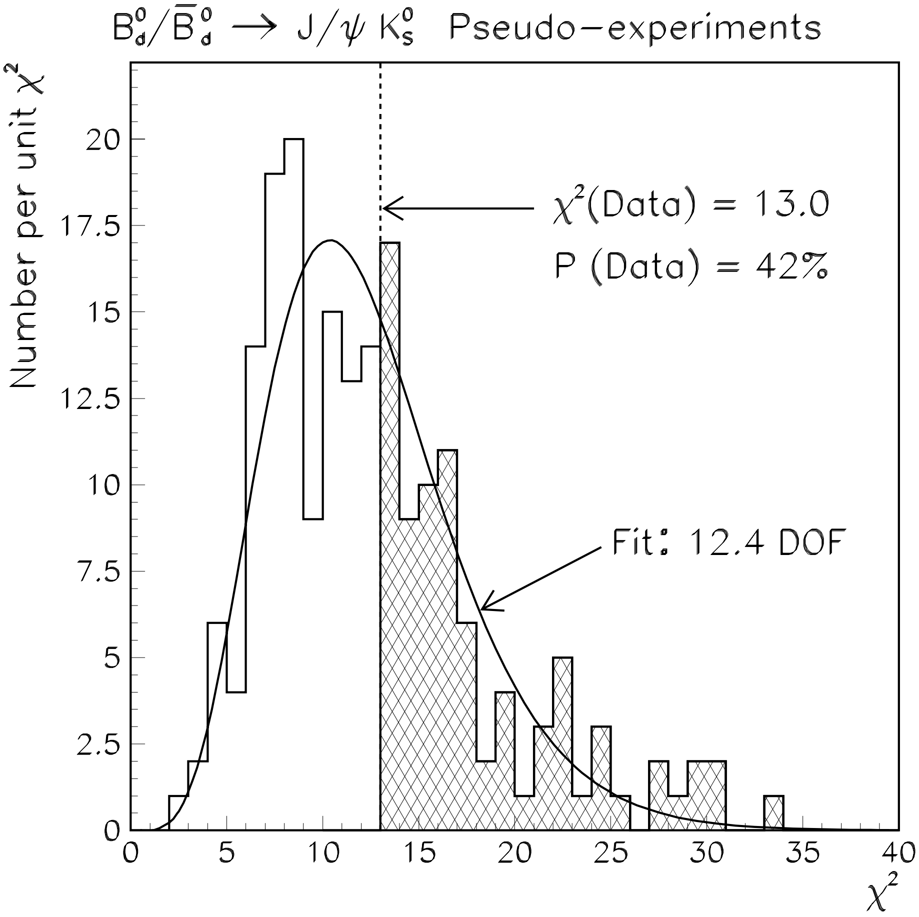

Figure 11-10: The distribution of χ^2 returned by the $J/\psi K_S^0$ Monte Carlo samples. The fit is to $f(z;n) = z^{n/2-1}e^{-z/2}/2^{n/2}$, $(n/2)$, which corresponds to the standard χ^2 with n degrees of freedom.

11.4 Conclusions on Systematic Biases

The systematic uncertainties from the input parameters are calculated and seen to be much smaller than the statistical uncertainties on the parameters of interest. Still, they are included for completeness. The "satellite peak" background is studied and estimated to be negligible.

The selection criteria for both B mesons and SST candidates are varied, and most of them have no noticeable effect. Varying the $p_T(\mathrm{SST})$ cutoff does change the charged dilution, though it does not affect the neutral dilution. This behavior is confirmed with the higher-statistics ℓ–D samples from [45], and with Monte Carlo. The measured value of $\mathcal{D}_0 \sin 2\beta$ shows considerable dependence on the value of $p_T(\mathrm{SST})$ chosen, but these variations are consistent with being statistical fluctuations. Furthermore, none of the systematic effects considered produced changes even 1/10th as large as that seen in the data.

Thus, it is concluded that the original choice of $p_T(\mathrm{SST}) = 0.4$ GeV/c does not introduce any systematic bias into the $\mathcal{D}_0 \sin 2\beta$ measurement. As described in Section 7.2, the choice of $p_T(\mathrm{SST}) = 0.4$ GeV/c was made as a compromise between larger detector bias for low $p_T(\mathrm{SST})$ and loss of efficiency for high $p_T(\mathrm{SST})$. Since no reason is indicated to abandon this choice, it is the one used in Chapter 12 to determine the value of $\sin 2\beta$ most consistent with this data.

Chapter 12

Measurement of $\sin 2\beta$

The raw CP asymmetry was measured using the maximum likelihood fit, described in Chapter 10. The systematic uncertainties on this asymmetry were determined in Chapter 11. All that remains is to extract the value of $\sin 2\beta$ from the raw asymmetry. To do this, the neutral dilution $\mathcal{D}_0$ is needed. Section 12.1 explains how $\mathcal{D}_0$ is determined using the results from the $J/\psi K^+$ and $J/\psi K^{*0}$ dilution measurements, combined with results from [45]. Section 12.2 provides the actual calculation of $\sin 2\beta$, and Section 12.3 explains how the result can be interpreted in terms of confidence limits on the allowed values of $\sin 2\beta$. Section 12.4 summarizes the the results presented in this chapter.

12.1 Dilution Extrapolation

To calculate $\sin 2\beta$ from $\mathcal{D}_0 \sin 2\beta$, one (obviously) needs to know the value of $\mathcal{D}_0$. Three approaches were considered for determining this dilution.

The first approach would be to use the dilution measured in $J/\psi K^{*0}$: $\mathcal{D}_0 = 0.165 \pm 0.112$. The $J/\psi K^{*0}$ and $J/\psi K^0_S$ data samples have essentially identical kinematics, so the dilutions should be very similar. Unfortunately the statistics in $J/\psi K^{*0}$ are not very high, and this approach would give a large systematic uncertainty on $\sin 2\beta$, due to the large uncertainty on the $\mathcal{D}_0$ measurement in $J/\psi K^{*0}$.

The second approach considered would be to use the dilution measured in [45]:

$\mathcal{D}_0 = 0.181 \pm 0.035$. That measurement is made using CDF data and the identical tagging algorithm, so the dilution could be appropriate for the $J/\psi K^0_S$ sample.[1] The advantage of using the ℓ–D sample is that the statistics of the ℓ–D sample are much higher, and the dilution is measured much more accurately. The difficulty with using this measurement is that the ℓ–D sample has more energetic B mesons. Figure 12-1 shows the $p_T(B)$ distributions for all five samples, and table 12.1 shows the means and RMS's for the $p_T(B)$ distributions. The average $p_T(B)$ for the ℓ–D samples is roughly 10 GeV/c higher than for the $J/\psi K$ samples. If the dilution depends on the energy of the B meson, then the dilution measured in ℓ–D would not be appropriate for $J/\psi K^0_S$.

Data Sample	$\langle p_T(B)\rangle$ (GeV/c)	$\sigma_{p_T(B)}$ (GeV/c)
Neutral ℓ–D	21.8	6.0
Charged ℓ–D	22.0	6.1
$J/\psi K^{*0}$	14.6	6.6
$J/\psi K^{\pm}$	12.8	6.3
$J/\psi K^0_S$	11.6	5.6

Table 12.1: Average values ($\langle p_T(B)\rangle$) and RMS's ($\sigma_{p_T(B)}$) for the $p_T(B)$ distributions for $J/\psi K$ and ℓ–D data.

The third approach considered would be to simply use Monte Carlo simulation to determine what the dilution should be for $J/\psi K^0_S$. Monte Carlo would clearly not suffer from problems with statistics, as very large simulated samples can be generated. Likewise, a Monte Carlo sample would not suffer from problems of dilution dependence on $p_T(B)$, as it could be set to generate B mesons with the proper kinematics. The problem with using Monte Carlo is that the details of $p\overline{p}$ collisions are not very

---

[1] The ℓ–D sample suffers from two complications not present in the $J/\psi K^0_S$ sample. There is ***crosstalk***, where B^0 mesons are identified as B^+ and vice-versa, and it is also possible for the ℓ–D events to tag on B daughters. But the dilution measurement in ℓ–D is corrected for these effects, and can thus be used for tagging fully-reconstructed B mesons.

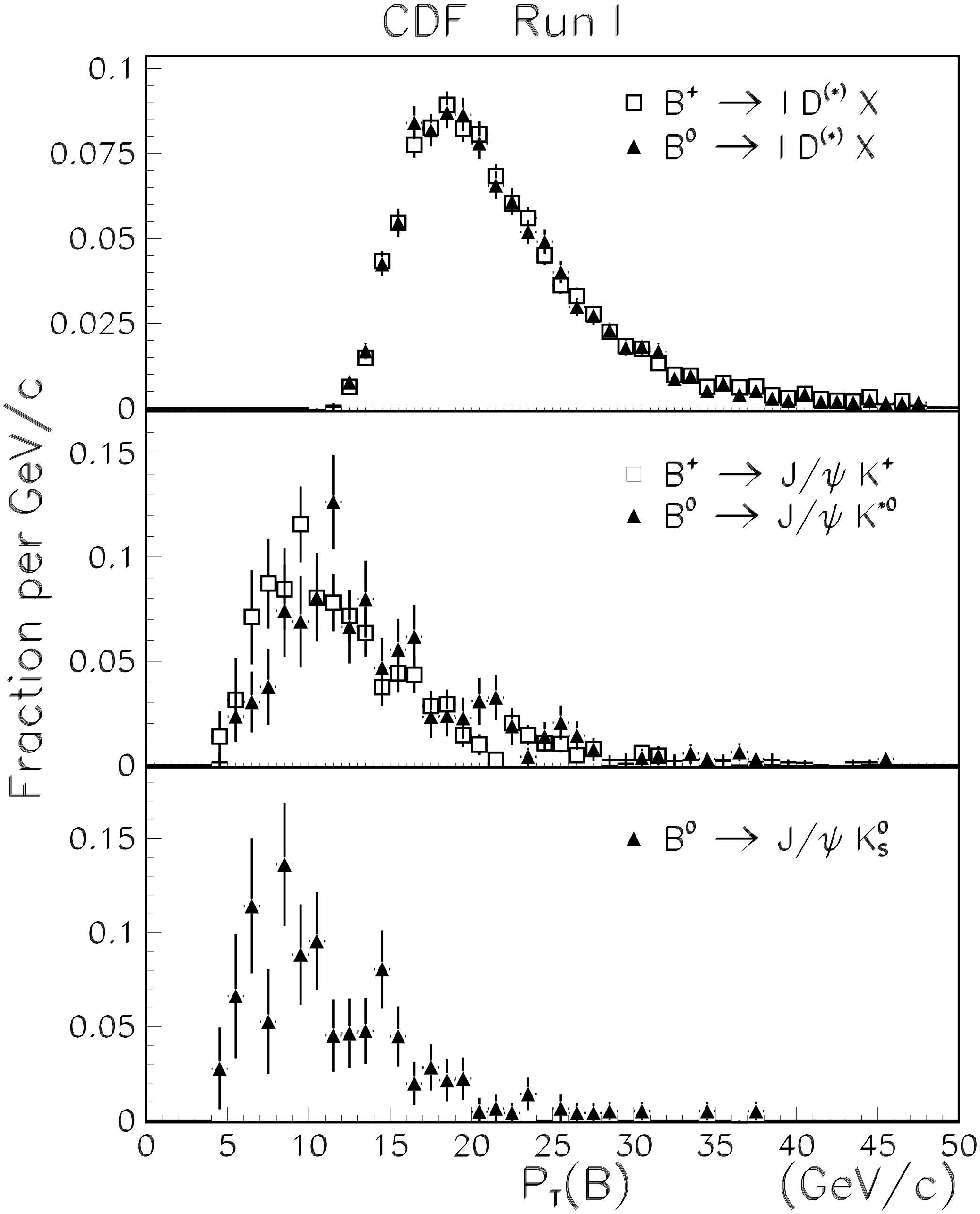

Figure 12-1: Distributions of $p_T(B)$ for the $J/\psi K$ and ℓ–D data.

well understood, as the basic process that one needs simulate, *i.e.* hadronization, is a low Q^2 process for which only phenomenological models exit (see Section 3.1). It is possible to specify minor changes in the input parameters for the Monte Carlo generation which significantly change the predicted dilutions.

The compromise chosen is to combine the three approaches. The Monte Carlo simulation is used to determine what dependence the dilution has on $p_T(B)$, and that dependence is used to extrapolate from the dilution measured in ℓ–D to the dilution that would be appropriate for $J/\psi K^0_S$. In addition, the Monte Carlo is used to predict the ratio of the charged dilution to the neutral dilution: $\mathcal{D}_+/\mathcal{D}_0$. This ratio is then used to extrapolate from the dilutions measured in the charged ℓ–D sample and the $J/\psi K^+$ sample to what would be appropriate for $J/\psi K^0_S$. These three extrapolations are then averaged together with the dilution measured in the $J/\psi K^{*0}$ sample to determine the dilution for $J/\psi K^0_S$.

12.1.1 Monte Carlo Simulation

The Monte Carlo used for this study requires much higher statistics than previous Monte Carlo studies, so a simplified Monte Carlo generation is used (described in Appendix A.3). The tagging efficiency and dilution of the Monte Carlo and their dependencies on $p_T(B)$ will both be needed to perform the extrapolation.

The tagging efficiency is calculated for the Monte Carlo in bins of $p_T(B)$, as shown in figure 12-2. The efficiency rises with $p_T(B)$, indicating that the fragmentation tracks (Section 3.1.3) are more numerous and/or more energetic around higher energy B mesons.

The dependence of the dilution on $p_T(B)$ is shown in figure 12-3. The vertical bands indicate the typical $p_T(B)$ ranges for the $J/\psi K$ and ℓ–D samples. The dilution at very low $p_T(B)$ is considerably lower than at high $p_T(B)$, but for the $p_T(B)$ range relevant to the data, the variation is only $\sim 20\%$ (relative). The charged dilution is consistently higher than the neutral dilution (as was predicted in Section 7.3).

For comparison, the ratio of the charged dilution to the neutral dilution, $\mathcal{D}_+/\mathcal{D}_0$, is plotted versus $p_T(B)$ in figure 12-4. The ratio exhibits very little dependence on

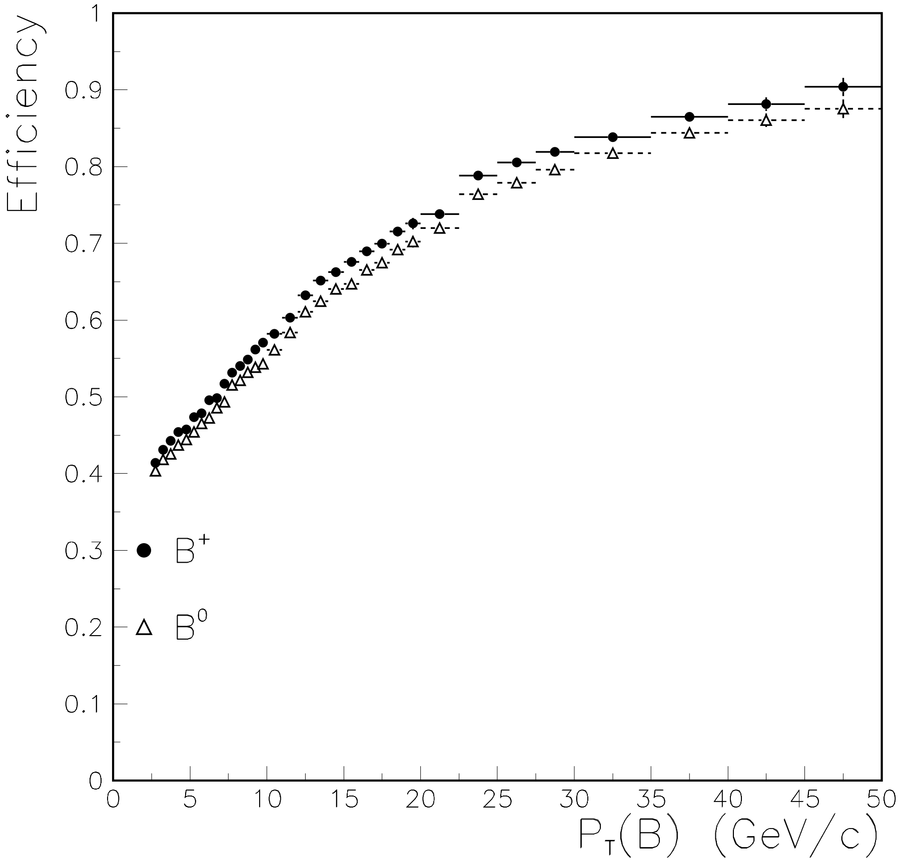

Figure 12-2: Tagging efficiency vs $p_T(B)$ for MC.

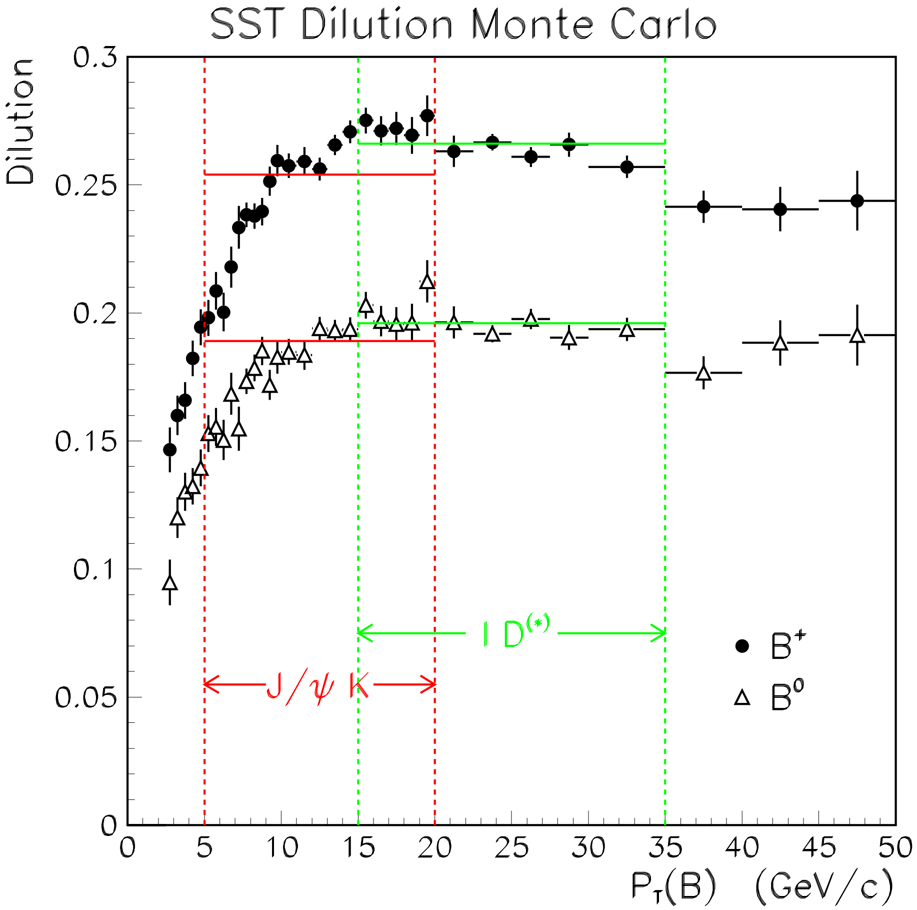

Figure 12-3: Dilution vs $p_T(B)$ for MC. The vertical bands indicate typical $p_T(B)$ regions for the $J/\psi K$ and ℓ–D samples. The horizontal lines indicated the Monte Carlo predicted dilutions for the $J/\psi K$ and ℓ–D samples (below).

$p_T(B)$; a flat-line fit with $\mathcal{D}_+/\mathcal{D}_0 = 1.35$ is quite consistent with the Monte Carlo results.

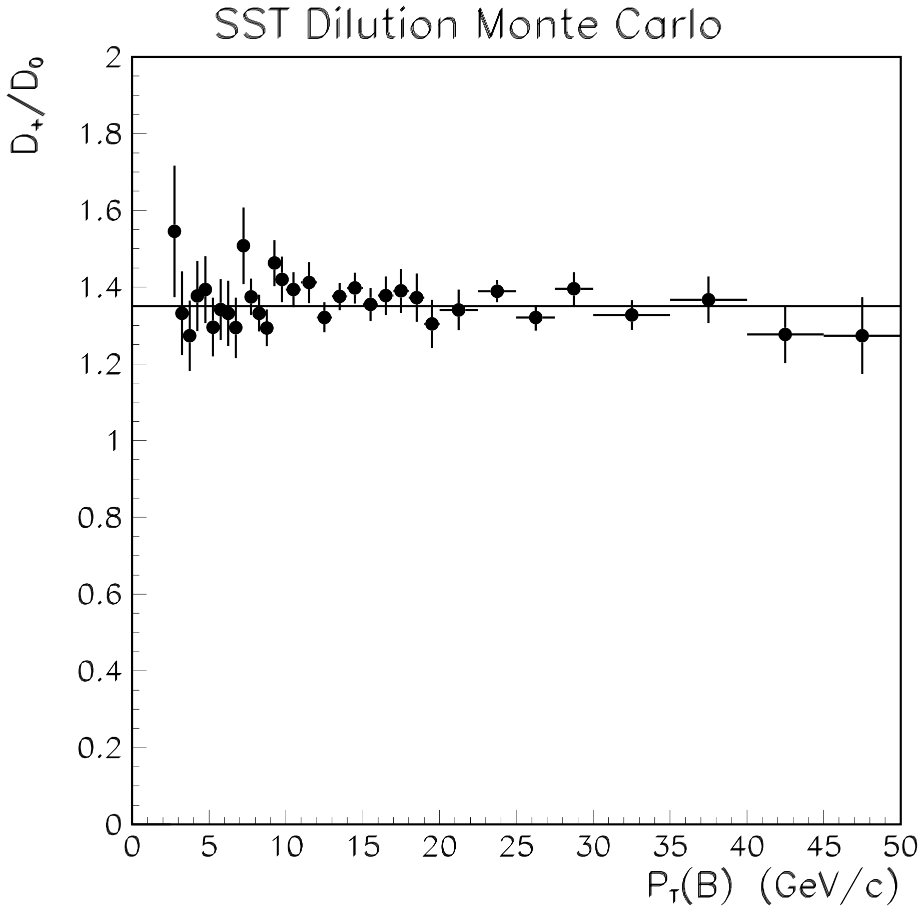

Figure 12-4: Ratio $\mathcal{D}_+/\mathcal{D}_0$ vs $p_T(B)$ for MC. The horizontal line is a flat-line fit at $\mathcal{D}_+/\mathcal{D}_0 = 1.35$.

Comparison with Data

For comparison, the dilution dependence on $p_T(B)$ is also calculated for the data (shown in figure 12-5). The lines indicate flat-line fits to the data, and the curves indicate the dilution predictions from the Monte Carlo. The data clearly do not have sufficient statistics to differentiate between the fairly weak $p_T(B)$-dependence predicted by the Monte Carlo and no $p_T(B)$-dependence at all.

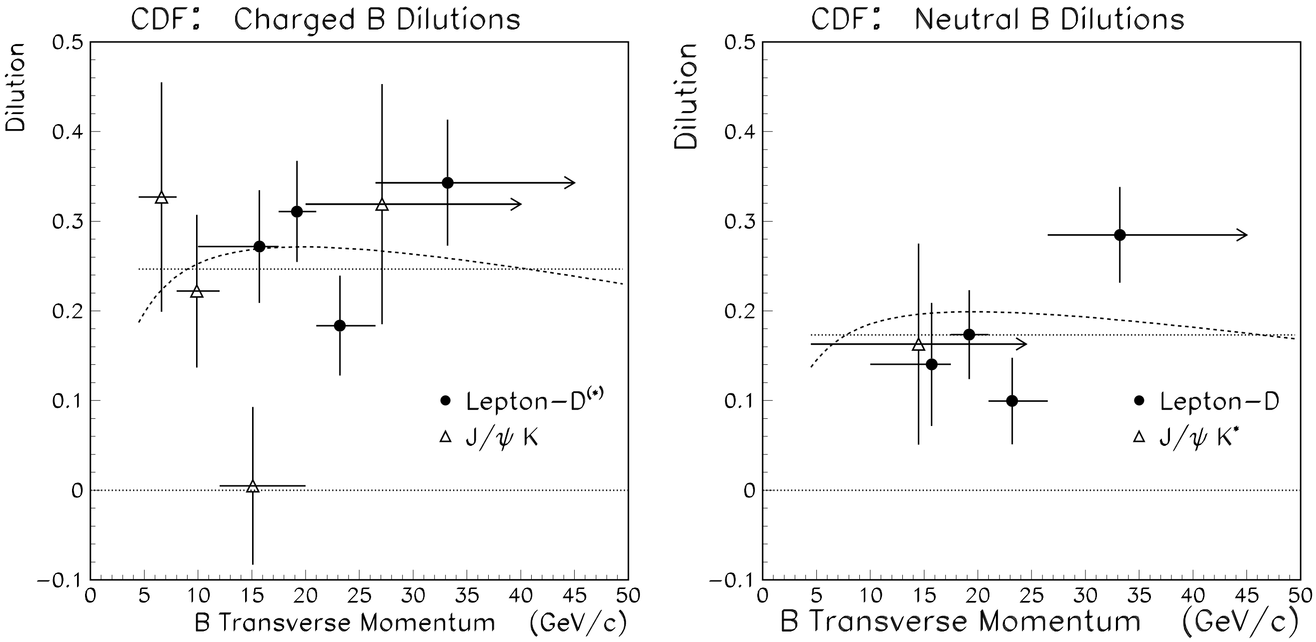

Figure 12-5: Dilutions for the data samples, versus $p_T(B)$. The curves are predictions from Monte Carlo simulation. Flat-line fits to the data yield $\mathcal{D}_+ = 24.6 \pm 2.6\%$ and $\mathcal{D}_0 = 17.3 \pm 2.6\%$.

12.1.2 Dilution Extrapolation for $J/\psi K_S^0$

To calculate the dilution that the Monte Carlo predicts for a specific data sample, the Monte Carlo dilutions are weighted by the $p_T(B)$ distribution from the data:

$$\mathcal{D}_{MC} = \frac{\sum_i \mathcal{D}_i N_i \epsilon_i}{\sum_i N_i \epsilon_i} \tag{12.1}$$

where i is index of the $p_T(B)$ bin, $\mathcal{D}_i$ and ϵ_i are the dilution and efficiency from the Monte Carlo, and N_i is the number of events from the data.

Using the $p_T(B)$ distributions from figure 12-1, the $p_T(B)$-weighted dilutions, $\mathcal{D}_{MC}$, are calculated for each sample. The horizontal lines in figure 12-3 indicate the values of $\mathcal{D}_{MC}$ for $J/\psi K^{*0}$, $J/\psi K^+$, and for the two ℓ–D samples. The values of $\mathcal{D}_{MC}$ for all five samples are shown in table 12.2, along with the dilutions measured in the data ($\mathcal{D}_{data}$) for each sample (except for $J/\psi K_S^0$, where there is of course no dilution measurement in the data), and the ratio of the two: $\mathcal{D}_{data}/\mathcal{D}_{MC}$.

The average of these ratios is $\mathcal{D}_{data}/\mathcal{D}_{MC} = 0.906 \pm 0.101$.[2] This ratio is within 1

---

[2] The two ℓ–D dilutions are measured in a simultaneous fit, and are correlated [45]. This corre-

Data Sample	Measured ($\mathcal{D}_{data}$)	Predicted ($\mathcal{D}_{MC}$)	Ratio ($\mathcal{D}_{data}/\mathcal{D}_{MC}$)
Neutral ℓ–D	0.181 ± 0.035	0.196	0.923 ± 0.179
Charged ℓ–D	0.267 ± 0.037	0.266	1.004 ± 0.139
$J/\psi K^{*0}$	0.165 ± 0.112	0.189	0.873 ± 0.593
$J/\psi K^{\pm}$	0.185 ± 0.052	0.254	0.728 ± 0.205
$J/\psi K^0_S$	-	0.183	-

Table 12.2: Measured and predicted dilutions for the various data samples. There is, of course, no measured $J/\psi K^0_S$ dilution.

standard-deviation of 1.0, and the four individual values are each within 1 standard-deviation of the average. This ratio is taken as a Monte Carlo "scale-factor." Multiplying this scale-factor by the $\mathcal{D}_{MC}$ for $J/\psi K^0_S$ gives:

$$\mathcal{D}_0 = 0.166 \pm 0.018. \tag{12.2}$$

This is the value used for $\mathcal{D}_0$ for $J/\psi K^0_S$. It incorporates the four data measurements, and uses the shape of the $p_T(B)$ dependence and the ratio $\mathcal{D}_+/\mathcal{D}_0$ from the Monte Carlo. It is very close to both the $\mathcal{D}_0$ measured in $J/\psi K^{*0}$ and that measured in ℓ–D, indicating that the effect of the Monte Carlo extrapolation is not large. To determine the systematic uncertainties on this extrapolation, variations are made in the Monte Carlo generation, and the changes in the predicted dilution are studied.

12.1.3 Systematic Checks of the Dilution Extrapolation

To test the validity of this dilution calculation, the Monte Carlo samples were regenerated with modified input parameters. The $p_T(B)$ dependence of the dilution, the charged/neutral ratio, $\mathcal{D}_+/\mathcal{D}_0$, and the final dilution prediction $\mathcal{D}_0(J/\psi K^0_S)$ are all studied for each variation of the input parameters. If the $p_T(B)$ dependence and the charged/neutral ratio show little dependence on the Monte Carlo input parameters, then relying on this information from the Monte Carlo should be safe.

---

lation (+0.372) is taken into account when performing the average.

The four parameters varied are those that seem most likely to affect the tagging, and they are listed in table 12.3. The central values and the variations are taken from [54] (see Appendix A). Since some tagging tracks are expected to come from B^{**} decays, the number of B^{**} mesons produced in the hadronization of the b quarks is varied. The Peterson fragmentation parameter ϵ_B affects the amount of energy in the fragmentation tracks, relative to the energy of the B meson, so it is varied. The fragmentation "p_T width" ($\sigma(p_T)$) affects the p_T distribution of fragmentation tracks. Harder fragmentation might be expected to lead to higher dilution. Finally, the scale-factor for the amount of underlying-event is varied. Tracks from the underlying event generally have no correlation with the flavor of the B meson, so increasing this cross-section would be expected to lower the dilution. These variations are very conservative [54].

Parameter	Nominal	Low	High
B^{**} fraction	30%	24%	38%
Peterson ϵ_B	0.0063	0.004	0.008
$\sigma(p_T)$ (MeV/c)	600	360	800
Und. Event Scale-Factor	1.66	1.0	2.5

Table 12.3: Values of four input parameters for the Monte Carlo generation. The parameters and ranges chosen are from [54].

Figure 12-6 shows the fractional change in the predicted dilutions for the four data samples, for each of the Monte Carlo generations. The uncertainties on the plot are from Monte Carlo statistics; roughly 10^6 Monte Carlo events were generated for each variation. The dilutions do change significantly, particularly with $\sigma(p_T)$ and the underlying event scale-factor. These variations confirm that relying on the Monte Carlo alone would have been inappropriate. However, the ratios between the dilutions are more stable and using the Monte Carlo for extrapolation of these ratios should be fine.

Figure 12-7 shows the variation of the ratio of the charged to neutral dilution for each Monte Carlo generation. The largest shift from the average value, 1.35, is -0.05.

Figure 12-8 shows the variation of the ratio of the dilution for low-p_T (in the $J/\psi K$ range) B mesons to the dilution of high-p_T (in the ℓ–D range) ones. The average value for this ratio is 0.96, and the largest shift is $\sim +0.08$, which happens for both charged and neutral B mesons when $\sigma(p_T) = 360$ MeV/c. None of the points in this plot are more than $\sim 6\%$ away from 1.0, which indicates that the $p_T(B)$ dependence of the dilution is not strong and also that it does not depend significantly on the Monte Carlo input parameters.

Figure 12-9 shows the extrapolated $J/\psi K^0_S$ dilution for each generation. The plot shows both the correlated uncertainties due to the data statistics and the much smaller uncorrelated uncertainties from the Monte Carlo statistics. The largest variation in the predicted $J/\psi K^0_S$ dilution is for $\sigma(p_T) = 360$ MeV/c, where $\Delta\mathcal{D}_0(J/\psi K^0_S) = +0.01$.

The ratio of charged to neutral dilutions $\mathcal{D}_+/\mathcal{D}_0$ and the $p_T(B)$ dependence of the dilutions are both seen to be very stable, changing only slightly when the Monte Carlo input parameters are varied. The variation which causes the largest shift in $\mathcal{D}_0(J/\psi K^0_S)$, lowering $\sigma(p_T)$ to 360 MeV/c, is a very conservative shift [54]. Thus, it is concluded that this approach yields a dilution which should be appropriate for $J/\psi K^0_S$. The largest shift seen in the Monte Carlo variations is taken as a 1–σ systematic uncertainty:

$$\mathcal{D}_0(J/\psi K^0_S) \quad = \quad 0.166 \pm 0.018\,(\text{stat.}) \pm 0.010\,(\text{syst.}), \qquad (12.3)$$

The statistical uncertainty is the combination of the uncertainties from the data, and the systematic uncertainty is from varying the Monte Carlo parameters.

Additional Systematic Study: f_K and f_p

The above systematic studies vary Monte Carlo input parameters to determine how much the dilution calculation depends on them. These parameters change the track multiplicities, energies, and angular dependencies, but they do not change the species of the particles being generated. In each of the above simulations, roughly the same

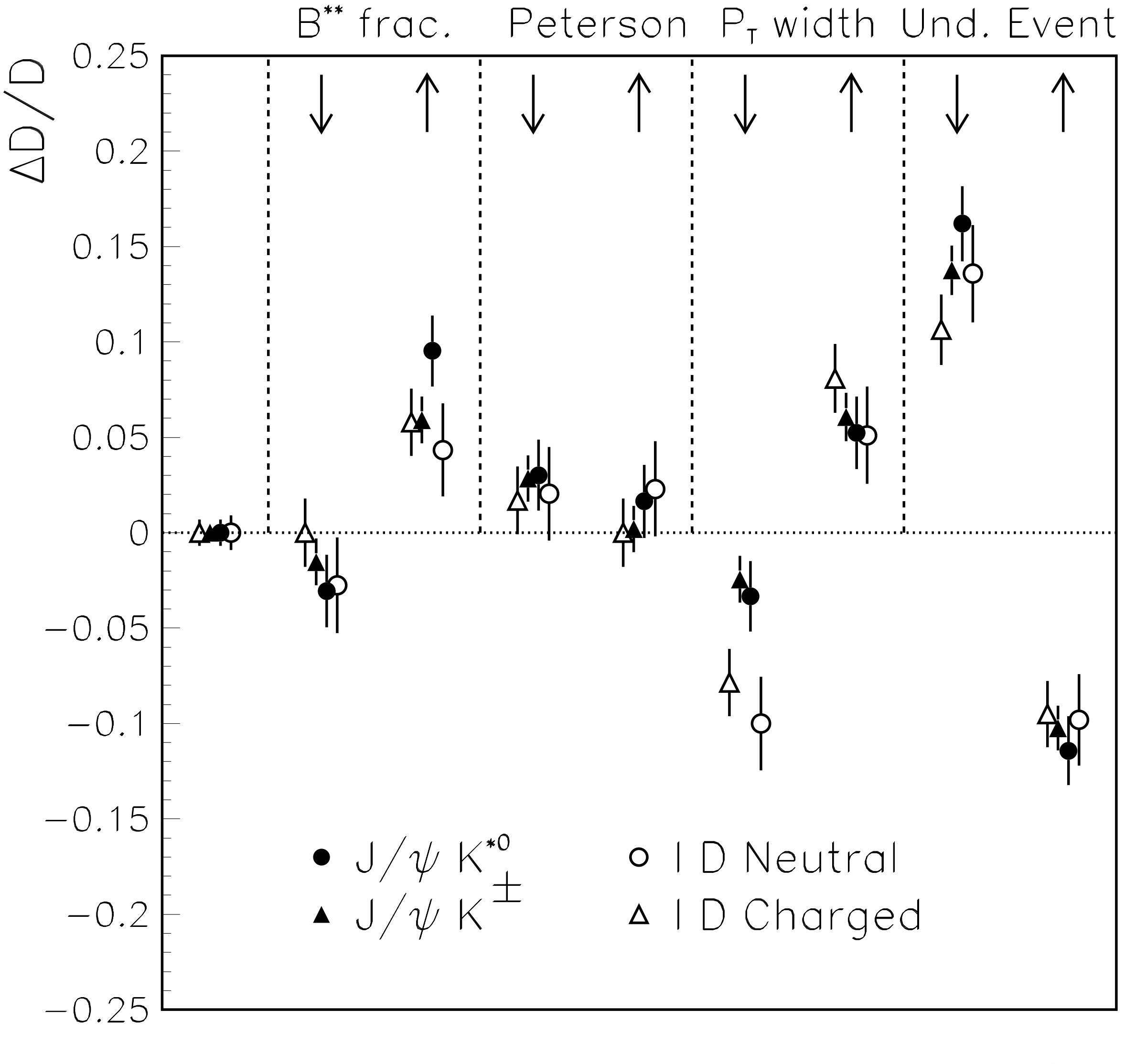

Figure 12-6: Fractional change in dilution predicted for each data sample. The horizontal scale indicates which Monte Carlo sample is being compared to the nominal Monte Carlo sample. Uncertainties are due to MC statistics.

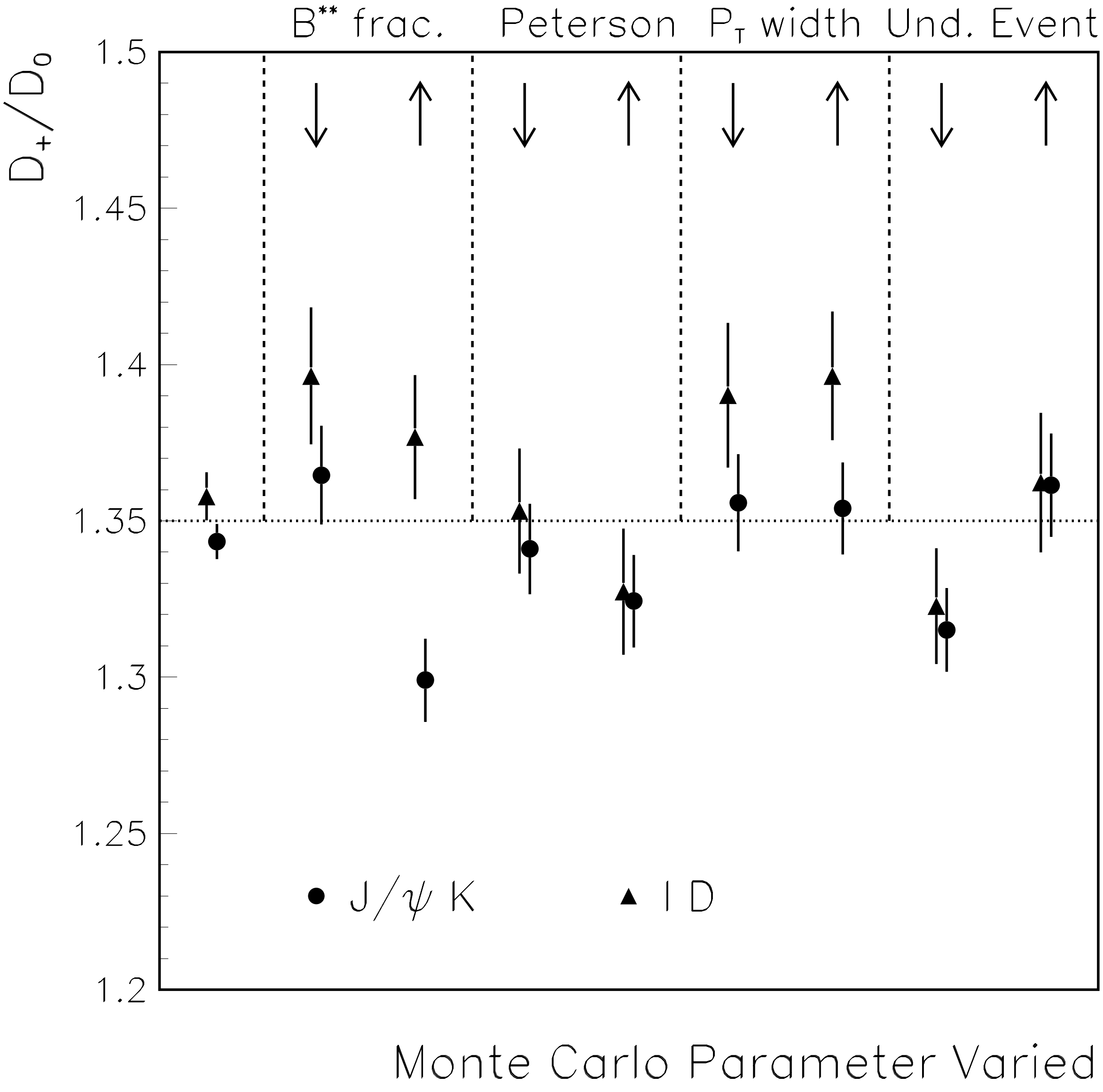

Figure 12-7: Ratio of charged to neutral dilution. The horizontal scale indicates which Monte Carlo sample is being compared to the nominal Monte Carlo sample. Uncertainties are due to MC statistics. Zero is suppressed on the vertical scale.

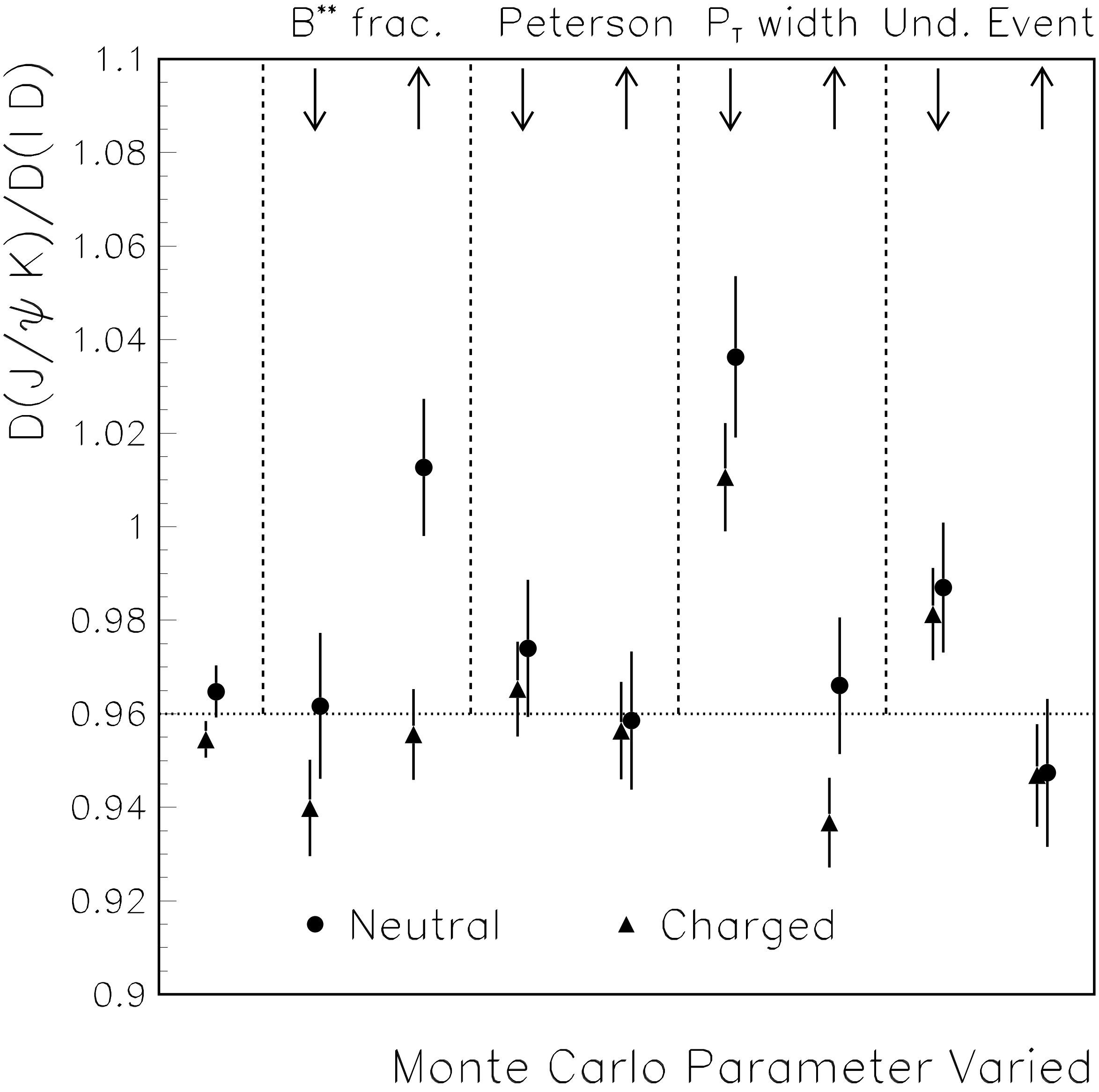

Figure 12-8: Ratio of dilution for low p_T B mesons to that of high p_T B mesons. The horizontal scale indicates which Monte Carlo sample is being compared to the nominal Monte Carlo sample. Uncertainties are due to MC statistics. Zero is suppressed on the vertical scale.

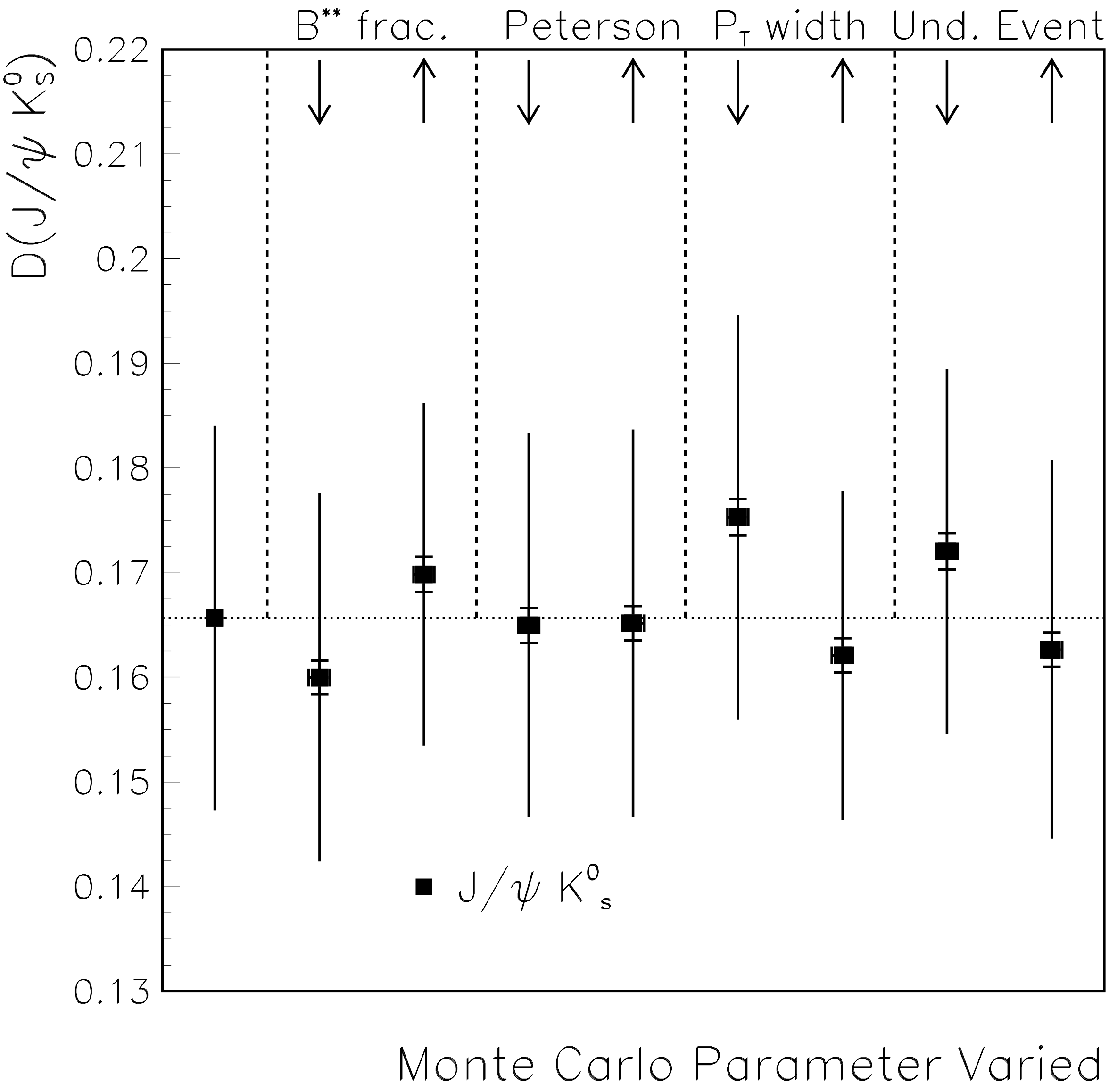

Figure 12-9: Predicted dilution for $J/\psi K_S^0$ sample. The horizontal scale indicates which Monte Carlo sample is being compared to the nominal Monte Carlo sample. The full error bars represent uncertainties from data statistics, which are 100% correlated between bins. The horizontal bars represent the uncorrelated uncertainties due to MC statistics. Zero is suppressed on the vertical scale.

numbers of kaons and protons are included among the SST tags. In addition to the parameters above, the fraction of kaon tags (f_K) and the fraction of proton tags (f_p) are also varied.

Table 12.4 shows the dilutions for SST tags in the Monte Carlo, separated by the species of particle used for the tag.[3] Roughly 50% of both B^+ and B^0 mesons tag on pions, and the dilutions for these tags are both 22%. This agreement is not unexpected. The difference between tagging B^+ and B^0 mesons is expected to come from tagging on protons and kaons (see Section 7.3).

Tag Flavor	Fraction (f)	Dilution ($\mathcal{D}$)	$f \times \mathcal{D}$
B^0 Mesons			
π^+	0.504	0.219	0.1104
K^+	0.046	-0.120	-0.0056
p	0.021	-0.078	-0.0016
other	0.0017	-0.025	-0.00004
Combined	0.573	0.180	0.1032
B^+ Mesons			
π^+	0.502	0.221	0.1110
K^+	0.063	0.404	0.0256
p	0.027	0.334	0.0091
other	0.0017	0.023	0.00004
Combined	0.594	0.245	0.1457

Table 12.4: Monte Carlo calculated dilutions for charged and neutral mesons, broken up by the flavor of the tagging particle. The combined fraction is the tagging efficiency ($\sum f_i$). The combined value for $f\mathcal{D}$ is the sum $\sum f_i \mathcal{D}_i$. The combined dilution is the weighted average $\sum f_i \mathcal{D}_i / \sum f_i$.

The charged B mesons tag on kaons 6.3% of the time, while only 4.6% of neutral

---

[3] In this study, no $p_T(B)$ weighting is performed. Varying the kaon and proton fractions changes $\mathcal{D}_+/\mathcal{D}_0$ (as will be shown), and this ratio does not appear to depend on $p_T(B)$.

B mesons tag on kaons. Furthermore, the dilution for a B^+ to tag on a kaon is 40%, while the dilution for a B^0 to tag on a kaon is -12%, much smaller than for B^+ mesons, and *with the opposite correlation to the pion tags for* B^0. Similarly, B^+ mesons tag on protons 2.7% of the time, with dilution 33%, whereas B^0 mesons tag on protons 2.1% of the time, with dilution -8%. Tagging on kaons and protons *increases* the dilution for B^+ and *decreases* the dilution for B^0.[4]

The average dilution in each sample is:

$$\overline{\mathcal{D}} = \frac{\sum_i f_i \mathcal{D}_i}{\sum_i f_i}, \tag{12.4}$$

where f_i is the fraction of events that tag on species i, and $\mathcal{D}_i$ is the dilution of those tags. To determine the systematic effects of f_K and f_p, these parameters are varied, and $\overline{\mathcal{D}}$ is recalculated for each variation. The kaon fraction, f_K, is varied by ±30% (relative), and the proton fraction, f_p, is varied by ±50% (relative). These variations are taken from a CDF measurement [57], and are quite conservative. Table 12.5 shows the effects of these variations. Increasing either fraction increases the charged dilution and decreases the neutral dilution, thus increasing $\mathcal{D}_+/\mathcal{D}_0$.

Fraction	Variation	$\Delta\mathcal{D}_+$	$\Delta\mathcal{D}_0$	$\Delta(\mathcal{D}_+/\mathcal{D}_0)$
f_K	±30%	±0.005	∓0.007	±0.084
f_p	±50%	±0.002	∓0.005	±0.045
Combined	–	±0.005	∓0.008	±0.095

Table 12.5: Effects on charged and neutral dilution calculations due to variation in the fractions of kaon and proton tags. The bottom row is the combination of the other two in quadrature (we assume the two effects are unrelated).

Varying f_K by ±30% changes $\mathcal{D}_+/\mathcal{D}_0$ by ±8.4%, and varying f_p by ±50% changes $\mathcal{D}_+/\mathcal{D}_0$ by ±4.5%. The effect of varying f_K is larger because there are more kaon

---

[4] Both samples also occasionally tag on tracks in the "other" category, which includes electrons and muons, generally from Dalitz decays of π^0's and leptonic decays of η's (decay products from long-lived particles like B, D, or K mesons are excluded by the simulation). These tags are very rare, so they are not considered further.

tags. Combining these two effects in quadrature leads to a variation of ±9.5% on $\mathcal{D}_+/\mathcal{D}_0$. Including this variation in the calculation of $\mathcal{D}_0(J/\psi K^0_S)$ is equivalent to adding a systematic uncertainty of ±0.008. Combining this systematic uncertainty with value from equation 12.3 yields:

$$\mathcal{D}_0(J/\psi K^0_S) = 0.166 \pm 0.018\,(\mathrm{stat}) \pm 0.013\,(\mathrm{syst}), \tag{12.5}$$

where the statistical uncertainty is the combination of the statistical uncertainties in the data samples, and the systematic uncertainty is due to the uncertainties associated with the dilution extrapolation using the Monte Carlo.

The effects of such large variations in the fractions of kaon and proton tags on the final extrapolated dilution for $J/\psi K^0_S$ are thus seen to be fairly small. This is for two reasons. First, the ratio $\mathcal{D}_+/\mathcal{D}_0$ measured in the ℓ–D data is 1.48 ± 0.29. The uncertainty on the data measurement of this ratio is considerably larger than the uncertainty from the Monte Carlo calculation, so including the Monte Carlo uncertainty has little effect. In addition, this variation only affects how the charged dilutions are averaged into the $J/\psi K^0_S$ dilution prediction; the extrapolations from the two neutral modes are unaffected.

12.2 The Value of $\sin 2\beta$

The raw CP asymmetry measured in $J/\psi K^0_S$ is:

$$\mathcal{D}_0 \sin 2\beta = 0.306 \pm 0.179\,(\mathrm{stat}) \pm 0.033\,(\mathrm{syst}). \tag{12.6}$$

Dividing $\mathcal{D}_0 \sin 2\beta$ (equation 12.6) by $\mathcal{D}_0$ (equation 12.5) gives:

$$\begin{aligned} \sin 2\beta &= 1.84 \pm 1.08\,(\mathrm{stat}) \pm 0.32\,(\mathrm{syst}). \\ &= 1.84 \pm 1.12 \qquad (12.7) \end{aligned}$$

12.3 Setting a Limit on $\sin 2\beta$

The central value for $\sin 2\beta$ is not in the "physical" region (*i.e.* $[-1,1]$).[5] This is because the raw CP asymmetry is larger than the tagging dilution factor. There has been much debate about how one should calculate limits on the allowed region of a parameter when the measurement indicates a value outside the physical region [32]. Recently, a method recommended by Feldman and Cousins [58] has gained favor. As it would be unwieldy to include a lengthy description of this method here, the reader is referred to reference [59], which describes this method and its merits in detail.

12.3.1 The Limit Calculation

In order to calculate a confidence limit on a parameter μ, one needs to know the expected distribution of the measured value x in terms of μ: $f(x;\mu)$. This function is simply the probability of measuring the value x, given that the true value is μ.

The Distribution Functions for $\sin 2\beta$

The raw CP asymmetry (to be denoted as "y") measured in this analysis has Gaussian distribution, with combined statistical and systematic uncertainty of $\sigma_y = 0.182$. Thus, if the dilution were known to be $\mathcal{D}'$, then the distribution of y would be:

$$f(y;\mu,\mathcal{D}') = \frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{(y-\mathcal{D}'\mu)^2}{2\sigma_y^2}\right] \tag{12.8}$$

where μ is the actual value of $\sin 2\beta$.

The "true" dilution is not known, but is estimated from the measured dilution $\mathcal{D}$, which has uncertainty $\sigma_{\mathcal{D}}$. If one treats the true dilution as a probabilistic variable,[6]

---

[5] Recall from equations 2.33 and 2.34 that the branching fraction for B^0 ($\overline{B}^0$) to $J/\psi K_S^0$ is proportional to $(1 \mp \sin 2\beta \sin \Delta m t)$. If $|\sin 2\beta| > 1$, then one of these branching fractions would be negative at $t = \pi/2\Delta m$, something which clearly cannot be possible. Thus it must be the case that the actual value of $|\sin 2\beta|$ is less than or equal to 1.

[6] While this is not technically the correct thing to do, it is a good approximation, as the uncertainty on $\mathcal{D}$ is much smaller than the value of $\mathcal{D}$.

then the probability distribution of the true dilution would be:

$$P(\mathcal{D}') = \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{D}}} \exp\left[-\frac{(\mathcal{D}' - \mathcal{D})^2}{2\sigma_{\mathcal{D}}^2}\right] \tag{12.9}$$

Including this with equation 12.8 and integrating over the unknown $\mathcal{D}'$ yields the following distribution for y:

$$f(y;\mu) = \frac{1}{\sqrt{2\pi}\omega(\mu)} \exp\left[-\frac{(y - \mathcal{D}\mu)^2}{2\omega(\mu)^2}\right] \tag{12.10}$$

where $\omega(\mu) = \sqrt{\sigma_y^2 + \mu^2\sigma_{\mathcal{D}}^2}$ contains both the uncertainty on y and the uncertainty on $\mathcal{D}$.

The distribution is still Gaussian, but its width depends on the actual value of $\sin 2\beta$. This should be expected; when $\sin 2\beta = 0$, there is no asymmetry, and y must be entirely a statistical fluctuation; the actual value of $\mathcal{D}$ is irrelevant. But when $\sin 2\beta = 1$, the uncertainty on $\mathcal{D}$ leads to uncertainty on the predicted asymmetry, which widens the distribution of measured values. However, the effect is not very large, since $\omega(1)/\omega(0) = 1.007$.

The measured value of $\sin 2\beta$ is $x \equiv y/\mathcal{D}$. The uncertainty on $\mathcal{D}$ is included in the distribution function for y and is therefore not included again. Thus, the distribution function for x in terms of μ is:

$$f(x;\mu) = \frac{1}{\sqrt{2\pi}(\omega(\mu)/\mathcal{D})} \exp\left[-\frac{(x - \mu)^2}{2(\omega(\mu)/\mathcal{D})^2}\right]. \tag{12.11}$$

The "Scaled Likelihood" Function

From the probability distribution function, one can define the "scaled likelihood" function:

$$R(x;\mu) = \frac{f(x;\mu)}{f(x;\mu_{best})}, \tag{12.12}$$

where μ_{best} is the value of μ which is most likely to result in a measurement of x. This is the ratio of the probability of an experiment measuring x when the true value is μ to the maximum probability of an experiment measuring x for any value of μ. If x is inside the physical region, μ_{best} will generally be equal to x, but if x is outside the physical region, μ_{best} will be the boundary between the physical and unphysical regions.

For this analysis, the boundaries of the physical region are ± 1. Thus, when $x > 1$, μ_{best} will be 1, when $x < -1$, μ_{best} will be -1, and when $-1 < x < 1$, μ_{best} will equal x. Thus, the "scaled likelihood" functions for this analysis are:

$$R(x,\mu) \quad = \quad \begin{cases} \frac{\omega(-1)}{\omega(\mu)} \ \exp\left[-\frac{(x-\mu)^2}{2(\omega(\mu)/\mathcal{D})^2} + \frac{(x+1)^2}{2(\omega(-1)/\mathcal{D})^2}\right] & (x < -1) \\ \\ \frac{\omega(x)}{\omega(\mu)} \ \exp\left[-\frac{(x-\mu)^2}{2(\omega(\mu)/\mathcal{D})^2}\right] & (-1 < x < 1) \\ \\ \frac{\omega(1)}{\omega(\mu)} \ \exp\left[-\frac{(x-\mu)^2}{2(\omega(\mu)/\mathcal{D})^2} + \frac{(x-1)^2}{2(\omega(1)/\mathcal{D})^2}\right] & (x > 1) \end{cases} \qquad (12.13)$$

where, again, $\omega(\mu) = \sqrt{\sigma_y^2 + \mu^2\sigma_{\mathcal{D}}^2}$. Thus, when x is in the physical region, $R(x;\mu)$ follows a Gaussian distribution, but normalized by $\omega(x)/\omega(\mu)$. This normalization term reflects the fact that the distribution function has a width that depends on the value of μ. Outside the physical region, $R(x;\mu)$ is the ratio of the probability of μ fluctuating to x to the probability of $+1$ or -1 fluctuating to x. Again, the effects of different widths are included.

Determining the Confidence Interval

This function is used to rank the values of x, to determine which ones are most consistent with a given μ. For each value of μ, the limits γ_1 and γ_2 are calculated such that

$$\int_{\gamma_1}^{\gamma_2} f(x;\mu)dx \quad = \quad 1 - \epsilon, \qquad (12.14)$$

where $1 - \epsilon$ is the desired exclusion (*eg.* 95%). An additional constraint is needed to define γ_1 and γ_2. The choice recommended by reference [58] is:

$$R(\gamma_1;\mu) \quad = \quad R(\gamma_2;\mu). \qquad (12.15)$$

The limits, γ_1 and γ_2, partition the values of x into two sets: those inside the range $\gamma_1 < x < \gamma_2$, and those outside that range. The probability of an experiment measuring a value inside the region is $1-\epsilon$, and the probability of measuring a value outside the region is ϵ. Furthermore, the values of x inside the region all have values of $R(x;\mu)$ larger than those of x that are outside the region.[7]

The values of γ_1 and γ_2 are calculated for each value of μ. Then the experiment is performed, and x is measured. From this, the "confidence interval" is defined to be those values of μ where $\gamma_1(\mu) < x < \gamma_2(\mu)$. Generally, this is a continuous region $\mu \in (c_1, c_2)$. By construction, the probability that the region (c_1, c_2) contains the true value of μ is $1-\epsilon$.

The horizontal lines in figure 12-10 show the confidence bands for each value of μ,[8] for $1-\epsilon = 95\%$. The value of $\sin 2\beta$ measured by this experiment (1.84) is indicated by the vertical line. For this value of $\sin 2\beta$, the confidence interval is defined by $c_1 = -0.20$, and $c_2 = 1$. Values of $\sin 2\beta$ less than -0.20 are *excluded* at 95% confidence level.

12.3.2 Experimental Sensitivity

The measured value of $\sin 2\beta$ is in the unphysical region, which is part of the reason the lower limit is so high. If the true value of $\sin 2\beta$ were 1, and this experiment were repeated many times, the median lower limit set would be -0.89. This number is termed the "sensitivity" of the experiment. The difference between this value and the limit actually set (-0.20) indicates how "fortunate" this experiment was to be able to set such a high limit.

Another measure of the experimental sensitivity is the probability of setting any limit at all. With the statistical accuracy of this measurement, any measured value of $\sin 2\beta$ above $+0.85$ would exclude -1 at $\geq 95\%$ C.L. If the true value of $\sin 2\beta$ were 1, then 56% of repeated experiments would measure a value ≥ 0.85 and thus exclude -1 at $\geq 95\%$ C.L. In addition, the fraction of these experiments that would

---

[7] Actually, the recommendation of reference [58] is to partition the set of values of x based on this last requirement, that those within the region have values of $R(x;\mu)$ greater than those outside the region. In this analysis, the two requirements are equivalent, but the first is simpler to explain.

[8] Recall that μ is the *actual* value of $\sin 2\beta$, which *cannot* be outside the physical region.

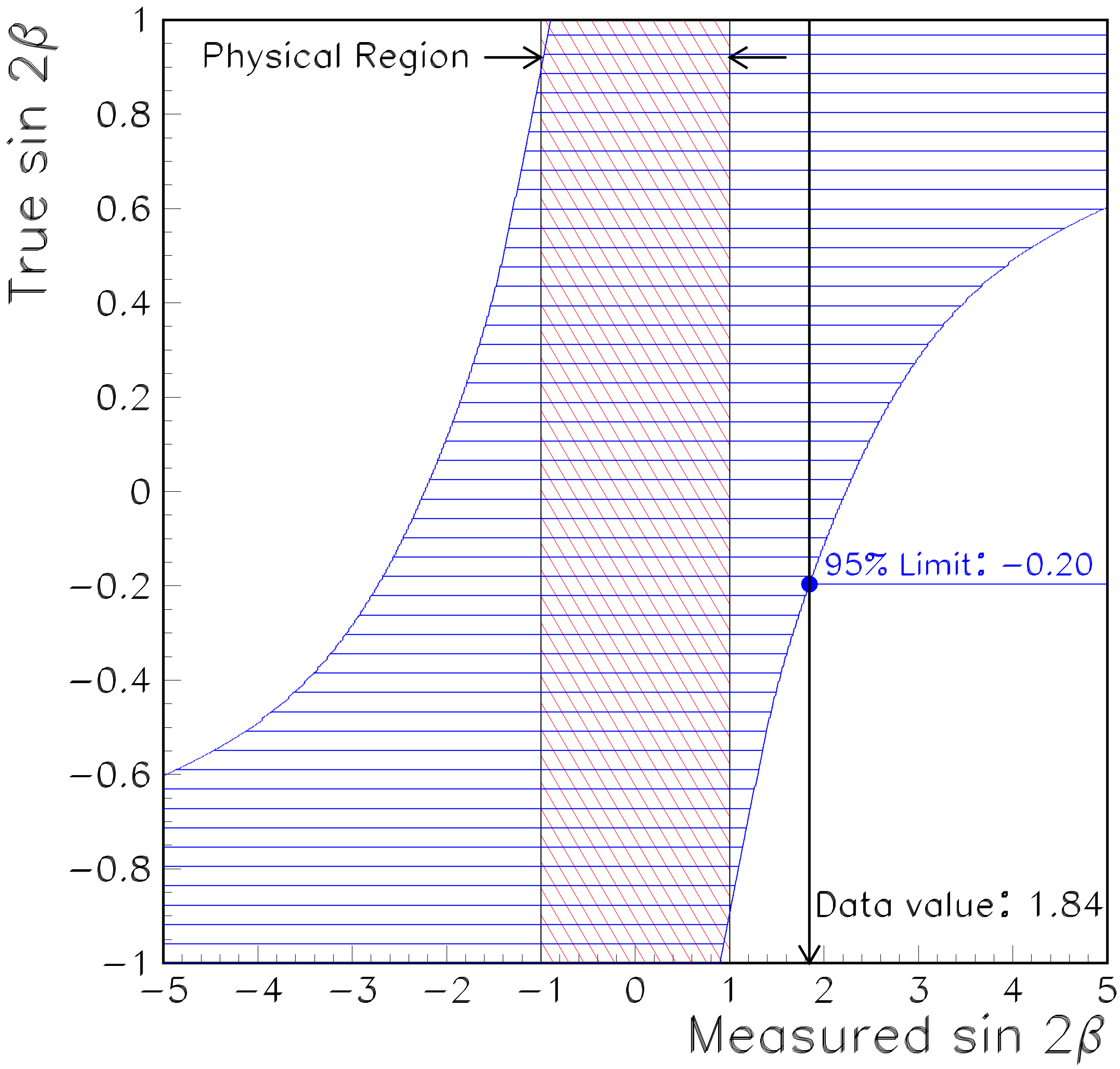

Figure 12-10: Limit plot for $\sin 2\beta$. Vertical scale is true value of $\sin 2\beta$. Horizontal scale is the measured value of $\sin 2\beta$. The band shows the "confidence belts" for 95% exclusion.

set a higher limit than the one set here is 23%. The results of this experiment are fortuitous, but not extraordinarily so.

12.3.3 Exclusion of $\sin 2\beta \leq 0$

The confidence level with which $\sin 2\beta \leq 0$ is excluded is *independent* of the dilution. If $\mathcal{D} \sin 2\beta \leq 0$, then either $\mathcal{D} \leq 0$ (which is clearly not the case) or $\sin 2\beta \leq 0$. Thus, the exclusion of $\sin 2\beta \leq 0$ can be calculated directly from the measured value of $\mathcal{D} \sin 2\beta$ and its statistical and systematic uncertainties. This calculation indicates that the region $\sin 2\beta \leq 0$ is excluded at 90% C.L.

12.4 Summary

The dilution appropriate for $J/\psi K^0_S$ is extrapolated from the values measured in the $J/\psi K^+$ and $J/\psi K^{*0}$ data, as well as from the ℓ–D data used in [45]. Monte Carlo simulation is used to extrapolate the $p_T(B)$ dependence of the dilution, as well as the ratio of dilution factors for tagging charged versus neutral B mesons ($\mathcal{D}_+/\mathcal{D}_0$). The extrapolated dilution is:

$$\mathcal{D}_0(J/\psi K^0_S) = 0.166 \pm 0.018\,(\text{stat.}) \pm 0.013\,(\text{syst.}), \tag{12.16}$$

where the statistical uncertainty is the combination of the statistical uncertainties from the data, and the systematic uncertainty is due the the uncertainties associated with the dilution extrapolation using the Monte Carlo.

In the $J/\psi K^0_S$ sample, the raw CP asymmetry has been measured to be:

$$\mathcal{D}_0 \sin 2\beta = 0.306 \pm 0.179\,(\text{stat.}) \pm 0.033\,(\text{syst.}). \tag{12.17}$$

From this raw asymmetry, the value of $\sin 2\beta$ most consistent with the data is calculated to be:

$$\begin{aligned} \sin 2\beta &= 1.84 \pm 1.08\,(\text{stat.}) \pm 0.32\,(\text{syst.}). \\ &= 1.84 \pm 1.12 \qquad (12.18) \end{aligned}$$

Using the prescription set forth in [58], confidence intervals are calculated. Values of $\sin 2\beta$ less than -0.20 are excluded at 95% C.L. The experimental sensitivity for 95% C.L. lower limit for this analysis is -0.89. While setting a limit as high as this one was fortuitous, if $\sin 2\beta$ were 1, roughly 56% of experiments would exclude -1 at $\geq$ 95% C.L. and 23% would set a limit higher than -0.20. Values of $\sin 2\beta \leq 0$ are excluded at 90% C.L., and this exclusion is *independent* of the value of the dilution.

Chapter 13

Conclusions

B^0 and B^+ mesons have been reconstructed via the decays $B^+ \rightarrow J/\psi K^+$, $B^0 \rightarrow J/\psi K^{*0}$, and $B^0 \rightarrow J/\psi K^0_S$, using 100 pb^{-1} of $p\overline{p}$ collusions recorded using the CDF detector at the Fermilab Tevatron. After background-subtraction, it is estimated that 846 $J/\psi K^+$ decays, 365 $J/\psi K^{*0}$ decays, and 198 $J/\psi K^0_S$ decays were reconstructed.

These data were flavor-tagged using the same-side tagging algorithm developed in [45]. An unbinned maximum-likelihood fit was used to determine the charged dilution ($\mathcal{D}_+$) in $J/\psi K^+$, the neutral dilution ($\mathcal{D}_0$) in $J/\psi K^{*0}$, and the raw CP asymmetry ($\mathcal{D}_0 \sin 2\beta$) in $J/\psi K^0_S$. The measured values are:

$$\begin{aligned}
\mathcal{D}_+ &= 0.185 \pm 0.052\,(\text{stat.}) \pm 0.004\,(\text{syst.}), && (13.1)\\
\mathcal{D}_0 &= 0.165 \pm 0.112\,(\text{stat.}) \pm 0.020\,(\text{syst.}), \text{ and} && (13.2)\\
\mathcal{D}_0 \sin 2\beta &= 0.306 \pm 0.179\,(\text{stat.}) \pm 0.033\,(\text{syst.}). && (13.3)
\end{aligned}$$

The charged and neutral dilutions agree well with the values measured in the ℓ–D samples of [45] ($\mathcal{D}_+ = 0.267 \pm 0.037$ and $\mathcal{D}_0 = 0.181 \pm 0.035$). Although the statistical significance of the $\mathcal{D}_0$ measurement in $J/\psi K^{*0}$ is weak, it does appear to follow the expected $\cos \Delta m t$ time-dependence (shown in figure 10-5). The dependence of the raw CP asymmetry is also consistent with the expected $\sin \Delta m t$ shape (shown in figure 10-4).

The above dilution measurements are combined (including the ones from ℓ–D), using Monte Carlo simulation, to arrive at a neutral dilution appropriate for the

$J/\psi K_S^0$ sample:

$$\mathcal{D}_0(J/\psi K_S^0) = 0.166 \pm 0.018\,(\text{stat.}) \pm 0.013\,(\text{syst.}). \tag{13.4}$$

where the statistical uncertainty is the combination of the statistical uncertainties from the data, and the systematic uncertainty is due to the uncertainties associated with the dilution extrapolation using the Monte Carlo.

The raw CP asymmetry from $J/\psi K_S^0$ is divided by this dilution to arrive at the value of $\sin 2\beta$:

$$\begin{aligned} \sin 2\beta &= 1.84 \pm 1.08\,(\text{stat.}) \pm 0.32\,(\text{syst.}). \\ &= 1.84 \pm 1.12 \end{aligned} \tag{13.5}$$

Using the prescription set forth in [58], confidence intervals are calculated. Values of $\sin 2\beta$ less than -0.20 are excluded at 95% C.L. The experimental sensitivity for 95% C.L. lower limit for this analysis is -0.89. While setting a limit as high as this one was fortuitous, if $\sin 2\beta$ were 1, roughly 56% of experiments would exclude -1 at $\geq$ 95% C.L. and 23% would set a limit higher than -0.20. Values of $\sin 2\beta \leq 0$ are excluded at 90% C.L., and this exclusion is *independent* of the value of the dilution.

13.1 Projections for Future CDF Measurement of $\sin 2\beta$

In the near future, CDF will commence a new data taking run (Run II).[1] The Tevatron is undergoing the final stages of upgrades to increase the instantaneous luminosity, the number of proton and antiproton bunches, and the center-of-mass energy of the $p\overline{p}$ collisions [60]. Over the course of the two-year running period, 2 fb^{-1} of integrated luminosity are expected to be recorded (roughly 20 times the 100 pb^{-1} recorded in Run I, the data used in this analysis).

In addition, the CDF detector is being upgraded for the new running conditions [61]. Among other improvements, the CTC and SVX are being replaced by

---

[1] Run II is scheduled to start in early 2000.

similar, but more powerful, chambers. The p_T thresholds on the muon triggers will be lowered, which will increase the detector acceptance. Roughly $10,000$ $B^0 \rightarrow J/\psi K^0_S$ and $18,000$ $B^0 \rightarrow J/\psi K^{*0}$ events are expected to be collected with the improved detector.

A simple extrapolation based on the increased statistics indicates that the measurement of $\mathcal{D}_0$ in $J/\psi K^{*0}$ alone will have a statistical uncertainty of 0.016, and the raw CP asymmetry in $J/\psi K^0_S$ will have a statistical uncertainty of 0.025. If the measured neutral dilution is the same as for Run I, the uncertainty on $\sin 2\beta$ measured in Run II would be $\pm 0.16\,(\text{stat.}) \pm 0.10\,(\text{syst.})$, where the statistical error is from the uncertainty on the raw CP asymmetry measurement in $J/\psi K^0_S$ and the systematic uncertainty is due to the uncertainty on the measurement of $\mathcal{D}_0$ in $J/\psi K^{*0}$.[2] Improvements in the SST algorithm and the addition of other tagging lgorithms are expected to reduce this uncertainty by an additional factor of 2.

The improved $\sin 2\beta$ measurement obtained in Run II should, if the Standard Model is correct, be sufficient to observe CP violation in B mesons. The B^0–$\overline{B}^0$ system would then provide the only observation of CP violation outside the K^0–$\overline{K}^0$ system. This measurement, along with many others to be studied in Run II (such as CP violation in other B decay modes, and B^0_s mixing [61]), will help significantly overconstrain the CKM matrix, and may ultimately help resolve the mystery surrounding the origin of CP violation. The present analysis has been a new step on that journey.

Portions of this analysis have been submitted for Publication. The article "Measurement of the B^0_d-$\overline{B}^0_d$ flavor oscillation frequency and study of same side flavor tagging of B mesons in $p\overline{p}$ collisions" has been submitted to *Phys. Rev. D*; the preprint for this article is **FERMILAB-Pub-98/188-E**. The article "Measurement of the CP-Violation Parameter sin(2β) in $B^0_d/\overline{B}^0_d \rightarrow J/\psi K^0_S$ Decays" has been submitted to *Phys. Rev. Lett.*; the preprint for this article is **FERMILAB-Pub-98/189-E**.

---

[2] The other systematic uncertainties considered in this analysis are due to uncertainties on parameters (*eg.* τ_{B^0} and Δm) which will be measured more accurately with the higher statistics available in Run II. It is expected that the systematic uncertainty due to the uncertainty on the dilution measurement will be the dominant systematic uncertainty on $\sin 2\beta$ in Run II.

Part V

Appendices

Appendix A

Monte Carlo Simulation

In this analysis, three "full" Monte Carlo simulations have been employed.[1] The first is used to study the kinematics of $J/\psi K^{*0}$ decays, and is described in Section A.1. The second is used to study the p_T(SST) threshold dependence for tagging, and is described in Section A.2. The third is used to study the $p_T(B)$ dependence of the tagging dilution, and is described in Section A.3.

A.1 Monte Carlo for Kinematic Studies

When studying the kinematics of B mesons decay products, one only needs to model the kinematics and the decay of the B meson; the underlying event and fragmentation particles are not needed. The BGENERATOR program [62] is used to generate a large sample of B^0 mesons. For each event, BGENERATOR generates a single b quark, according to the p_T spectrum derived from the next-to-leading-order QCD calculations of inclusive b production by Nason, Dawson, and Ellis [63]. The b quark is hadronized into a B^0 meson using the Peterson fragmentation model [64] (using the Peterson parameter $\epsilon_B = 0.006$). No particles other than the B^0 meson are generated.

The QQ program [65] is then used to decay the B^0 mesons to $J/\psi K^{*0}$, the J/ψ to $\mu^+\mu^-$, and the K^{*0} to $K^+\pi^-$. QQ is a Monte Carlo simulation created by the CLEO collaboration [66] that simulates the decays of B mesons and of the daughters of B

---

[1] In addition, a Toy Monte Carlo is used. The Toy Monte Carlo is the subject of Appendix B.

mesons. This simulation is tuned using branching fractions measured in data. QQ can also be set to "force" certain decays, as above.

The QFL' program [67] is used to simulate the CDF detector response. This program extrapolates the paths of the charged particles through the CTC and generates a CTC track (with covariance matrix) for each particle. It uses a parameterized efficiency and resolution, rather than simulating wire hits and the full track reconstruction algorithm. This allows the simulation to be much faster than it would be for detailed CTC simulation. QFL' also generates simulated SVX hits, muon chamber hits, calorimeter information, and VTX information.

This simulated information is then studied exactly as if it were real data: muon candidates are formed by matching hits in the muon chambers with CTC tracks. SVX information is combined with CTC information, where appropriate. The VTX information is used to locate primary vertices. J/ψ and B meson candidates are reconstructed from the track information.

This sample is used to study the decays of B^0 mesons to $J/\psi K^{*0}$. It is used in Section 9.2 to study the effects of $K\pi$ swapping, and in Section 11.2 to study the partially reconstructed "satellite peak" background.

A.2 Monte Carlo for Tagging Studies

The above Monte Carlo is not useful for studying same-side tagging, as the B fragmentation tracks and the underlying event are not generated. To study SST, another Monte Carlo is used. This Monte Carlo uses the PYTHIA program [68] to generate $p\overline{p}$ events.

A.2.1 Full Event Simulation with PYTHIA

The PYTHIA program was developed by the LUND group [68]. It uses the string fragmentation model [39] and can be tuned to simulate different experimental environments. Unfortunately, the default settings for $p\overline{p}$ collisions at 1.8 TeV do not simulate the CDF environment perfectly: both the fragmentation and the underlying-event tracks are "softer" (have lower average p_T) and are less numerous than in the data. Reference [54] discusses comparisons with the data and tuning done on the PYTHIA parameters to make the simulation agree better with the data.

The three parameters which have the most effect on these distributions are "PARP(31)," the scale-factor for the number of tracks produced by the underlying event, "PARJ(55)," which is the Peterson fragmentation parameter (ϵ_B), and "PARJ(21)," called $\sigma_{p_T}^{frag}$, which controls the energy distribution of tracks produced in fragmentation. With these parameters set to new values, the track energy and multiplicity distributions match the data very well.

Once the event is generated by PYTHIA, the B hadrons are decayed using the QQ program, as for the first Monte Carlo. QQ forces $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K_S^0$, and $J/\psi \to \mu^+\mu^-$. To avoid correlations that might occur when multiple B mesons are tagged in a single event, only B^+ and B^0 events have their decays forced; the B^- and $\overline{B}^0$ mesons can decay to any of the available final states. As the branching ratio to the above modes ($J/\psi K^+$, $J/\psi K_S^0$) is only $\sim 10^{-4}$, the fraction of events with multiple tagged B mesons is negligible. After passing through QQ, the events are passed through QFL', and are then reconstructed as if they were real data. These reconstructed B candidates are then tagged using the SST algorithm. This simulation is used in Section 11.3 to study the dependence of the tagging dilution on the p_T(SST) threshold.

A.3 Monte Carlo for Dilution Dependencies

In Section 12.1, a Monte Carlo simulation is used to determine the dependence of the tagging dilution on the p_T of the B meson being tagged. For this simulation, very high statistics are needed, as the simulated events are to be split into many bins in $p_T(B)$. Generating a sample of this magnitude using the above method would have taken many CPU–years, and was therefore considered impractical.

To speed up the simulation, it is simplified. The PYTHIA program is used to generate the entire event as above, but the rest of the simulation is skipped; neither the QQ program nor the QFL' simulation is used. The momentum and flavor of the B mesons are taken directly from the simulation and the decay products of the B are ignored.[2] The B mesons are restricted to have $|\eta| < 1$, as a rough simulation of the detector acceptance. To avoid possible correlations that can arise from tagging multiple B mesons in a single event, only B^+ and B^0 mesons (and *not* their charge-conjugates) are used. Events with more than one B^+ or B^0 (or one of each) are discarded for the same reason.

The solid histogram in figure A-1 shows the $p_T(B)$ distribution of the simulated B mesons. This distribution falls very rapidly with rising $p_T(B)$, decreasing by four orders of magnitude over the range shown. This shape is similar to that for the data, except that, due to the B meson reconstruction and selection requirements, the data has lower efficiency at low $p_T(B)$. As the dilution will be measured in bins of $p_T(B)$, the distribution is not required to match that of the data.

In fact, a flatter distribution would be better, as higher statistics at high $p_T(B)$ would allow for better estimates of the $p_T(B)$ dependence. To flatten the distribution, additional samples are generated, with cuts on the p_T of the b quark[3] of 5, 10, and 20 GeV/c. The dashed histogram in figure A-1 shows the $p_T(B)$ distribution for the final sample. To avoid possible biases in events near the $p_T(b)$ cutoff, the B mesons are required to have p_T at least 2.5 GeV/c higher than the cutoff for the b quarks. Approximately one million each of B^+ and B^0 are generated for each $p_T(b)$ cut, and roughly half of these pass the selection requirements. This sample drops only two orders of magnitude in rate over the $p_T(B)$ range shown.

In addition to B mesons, the simulation produces other particles. Unstable parti-

---

[2] Since fragmentation is a *strong* process and B meson decay is a *weak* process, the two should be independent (and are, in any case, handled independently by simulations). Thus, the tagging dilution and efficiency should be independent of the B decay mode.

[3] The p_T cut is placed on the b quark, rather than the B meson, as the b quark p_T is determined early in the simulation, while the B meson p_T is not determined until the simulation is nearly complete (*i.e.* after the parton shower and hadronization processes have been simulated). Thus, a lower-limit on the p_T of the b quark speeds up the simulation considerably, rejecting the low p_T events immediately. A p_T cut on the B meson would be much less effective.

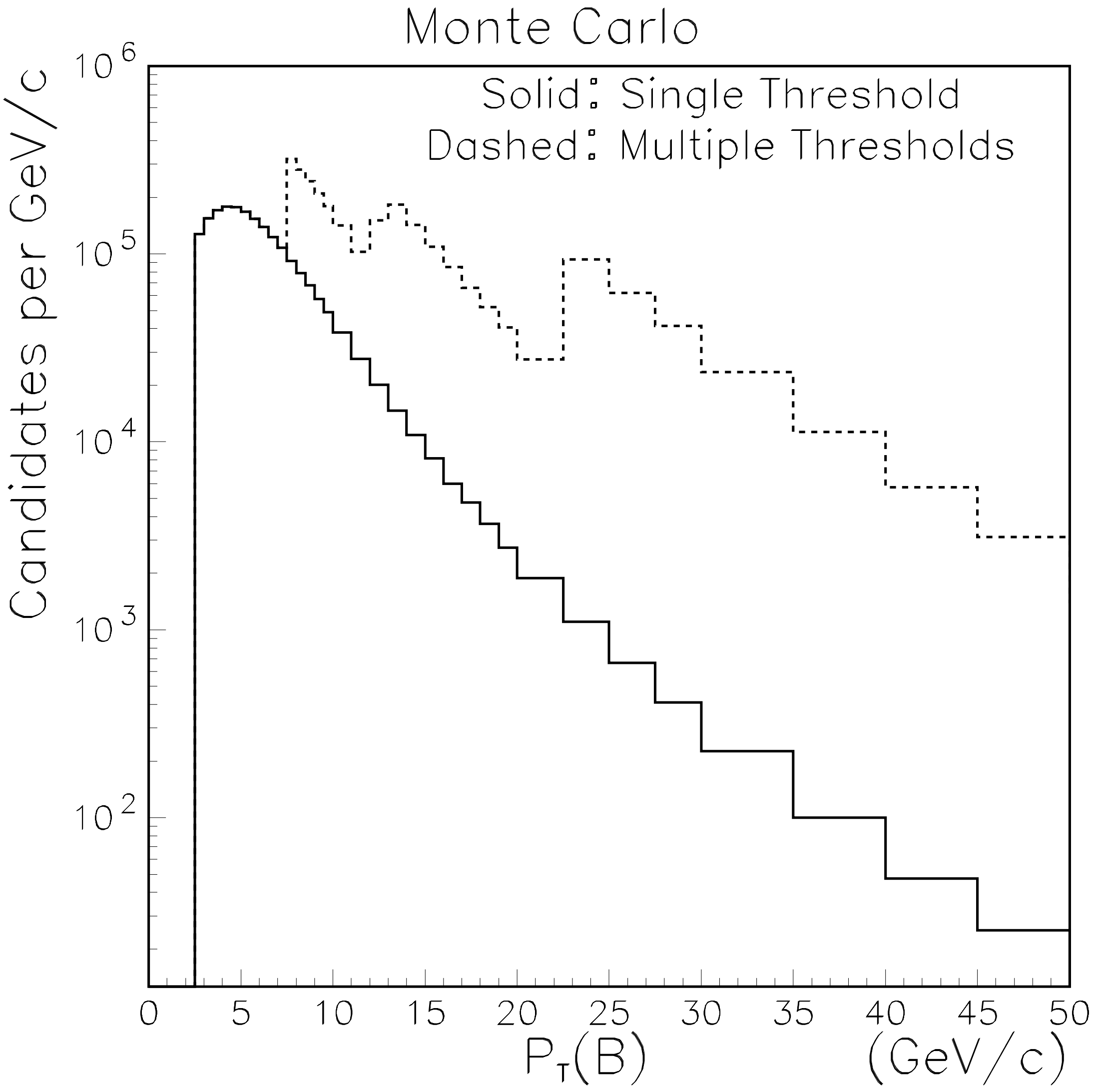

Figure A-1: p_T distribution for generated B mesons. The histogram is normalized by bin-width, so the vertical scale is "Candidates per GeV/c," for all bins, despite varying bin sizes. The solid histogram is for the sample generated with $p_T(b)$ threshold of 0. The dashed histogram also includes the other three samples, which are generated with thresholds of 5, 10, and 20 GeV/c.

cles are decayed to other particles, which may themselves be decayed, until "stable" particles are reached. The "stable" charged particles (e, μ, π, K and p) that have momenta extrapolating through the CTC are assumed to leave tracks in the CTC. The simulation reconstructs these tracks with perfect resolution and with an efficiency based on the track p_T, using the curve displayed in figure A-2. This efficiency curve is from an internal CDF note [69].

SST candidates are sought among these simulated tracks. The exit-radius cut applied to the data is applied to the simulation; this is nearly equivalent to requiring $|\eta| < 1$. Tracks from particles which are the descendants of weakly-decaying particles are discarded from consideration, as a rough simulation of the impact-parameter significance requirement for real-data tracks. This requirement also removes the possibility of tagging on B daughters, which is appropriate, as the B mesons in the data are fully-reconstructed and tagging on B daughters is not possible for the data. As for the data, the tracks are also required to have $p_T > 0.4$ GeV/c and be within $\Delta R < 0.7$ of the B meson. If any tracks pass these criteria, the one with the minimum p_T^{rel} is chosen as the SST tag.

This rough simulation of the CDF detector should be adequate, as it is only *relative* dilution measurements that are needed, because only *ratios* of dilutions from the simulation will be used. The effects of eschewing a more complicated detector simulation should be similar for B mesons of different energies, so the relative effects of the simplifications should be small. To study this, the simulation is simplified one more step: the p_T dependent track-reconstruction efficiency is skipped, and tracks are reconstructed with 100% efficiency. This raises the tagging efficiencies by a few percent, but has no significant effect on any of the dilution calculations. This indicates that a more complicated detector simulation would not be likely to have significantly different results, either.

The tagging efficiency and dilution are then measured in bins of $p_T(B)$ and $\eta(B)$. Figures A-3 and A-4 show the efficiency and dilution dependencies on $\eta(B)$. The efficiency is flat for $|\eta(B)| < 0.3$, and then drops off as $|\eta(B)|$ increases. This is not unexpected, as the tracks are required to have $|\eta| < 1$ (the exit-radius cut), and

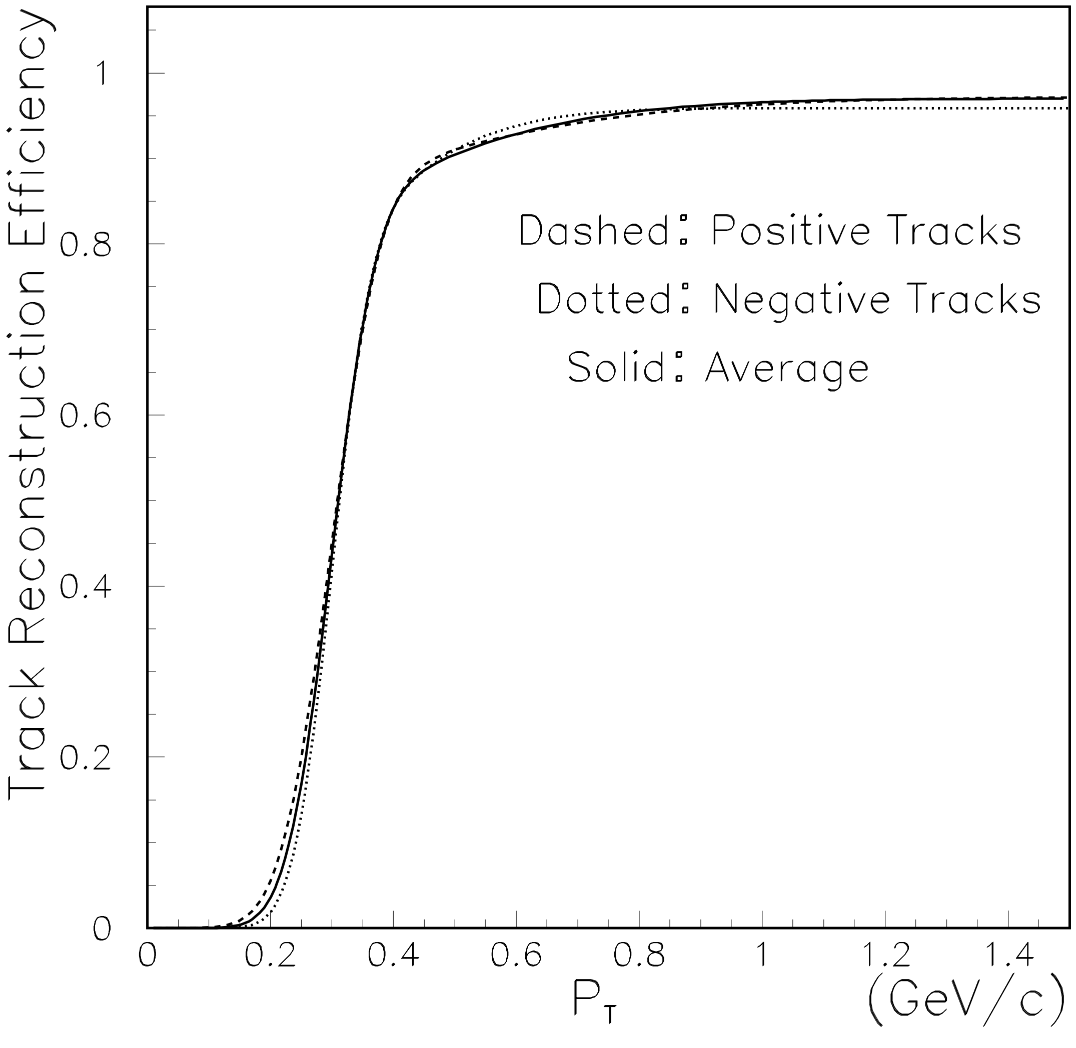

Figure A-2: The parameterization used in the track-reconstruction efficiency simulation. Reference [69] measures separate efficiencies for positive and negative tracks, but they are so similar that the average is used in this analysis. The efficiencies above 1.5 GeV/c are essentially constant.

be within $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.7$ of the B meson. When $|\eta(B)| < 0.3$, the exit-radius cut has no effect, and the efficiency is maximum and constant. When $|\eta(B)| > 0.3$, the exit-radius cut discards some tracks that would have been tagging candidates, so the efficiency drops. The tagging efficiency for B^+ mesons is slightly higher than that for B^0 mesons, but the shapes are nearly identical.

The tagging dilution for B^+ is higher than that for B^0, but neither dilution shows any significant dependence on $\eta(B)$ (see figure A-4). The purpose of this Monte Carlo is to determine what dependence the dilution has on $p_T(B)$, so the different $\eta(B)$ bins will be added together. If there were a dependence of the dilution on $\eta(B)$, this sum would need to be weighted by the $\eta(B)$ distribution from the data. As there is no dependence on $\eta(B)$, the weighting is not needed.

The tagging efficiency and dilution dependencies on $p_T(B)$ are discussed in Section 12.1, where they are used to extrapolate the dilution appropriate for the $J/\psi K^0_S$ data. Also discussed in that section are variations made to the Monte Carlo generation, which are used to determine systematic uncertainties on the dilution extrapolation. These variations result in significant changes to the p_T distributions and multiplicities of reconstructed tracks (see Chapter 12). These effects should be considerably larger than the effects of the approximations described above, so any systematic effect due to these approximations should be covered by the systematic uncertainties on the dilution extrapolation.

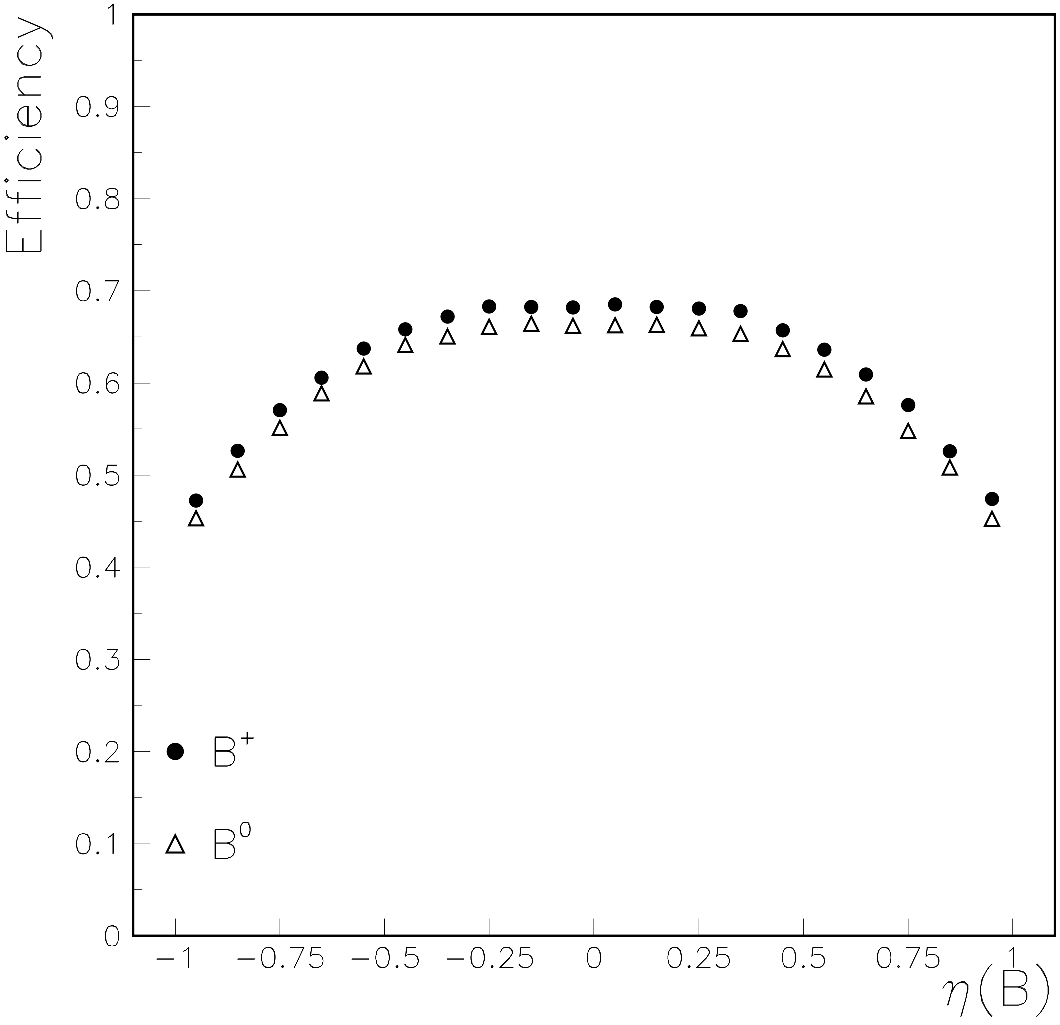

Figure A-3: Tagging efficiency vs $\eta(B)$ for MC. Tracks are selected within a cone of $\Delta R < 0.7$ around B mesons, so B mesons with $|\eta| > 0.3$ have decreasing efficiency, due to CTC fiduciality requirements on tagging tacks.

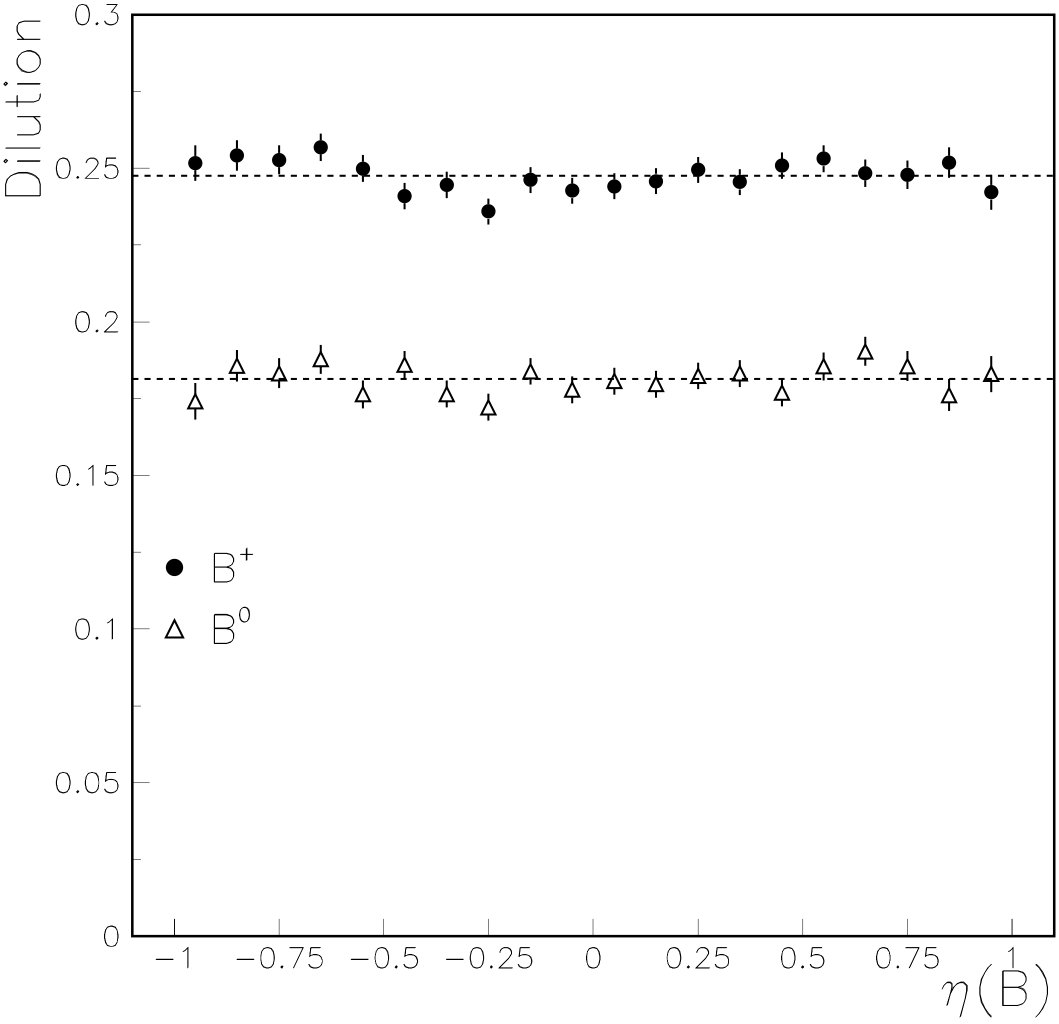

Figure A-4: Dilution vs $\eta(B)$ for MC. Horizontal lines indicate average charged and neutral dilutions.

Appendix B

Checks of Likelihood Fit: Toy Monte Carlo

The unbinned maximum-likelihood fit is complicated, and it is difficult to interpret the results of the fit visually. The figures in Chapters 6 and 10 compare binned distributions of the data to shapes that would result from binned fits with parameters identical to the results of the likelihood fits. These provide gross comparisons of the fit results to the data, but do not provide for clear estimation of the existence or magnitude of any fit biases.

To test the likelihood fit for biases, a Toy Monte Carlo (TMC) is used. It is designed to create a sample of events whose distribution is exactly (within statistical fluctuations) described by the likelihood function. This way, the expected values of the fit parameters are known exactly, and one can thus determine directly whether the values returned by the likelihood fit are unbiased estimators of the fit parameters.

The generation of the TMC is discussed in Section B.1 and tests of the TMC generation are discussed in Section B.2. Section B.3 describes the tests of the likelihood fit, and Section B.4 discusses the biases found in these tests. To exclude the possibility of a fit bias which depends on the value of $\sin 2\beta$, the likelihood tests are reperformed for different values of $\sin 2\beta$; the results of these tests are described in Section B.5. Section B.6 describes how the value of $\mathcal{L}_T$ can be used to judge the quality of the likelihood fit.

B.1 Toy Monte Carlo Generation

To ensure that the generated samples are exactly described by the likelihood function, the function is used directly in the TMC generation. The input parameters are set to the nominal values used in the fits to the data, and the fit parameters are set to the values returned by the fits to the data. Thus, the TMC samples should be exactly described (within statistical fluctuations) by the likelihood functions matching the fits to the data.

To achieve the proper statistical fluctuations, the number of events simulated for each sample is taken from the number collected in the data (N_{evt}). This number is then smeared by Poisson statistics,[1] a step necessary to give proper statistical fluctuations. Each event is assigned to be either signal or background, based on the probability f_B. The background events are further assigned to be prompt or long-lived, based on the probability f_L. The "long-lived" background are assigned to the negative tail, the short positive tail, or the long positive tail, based on the probabilities f_N and $f_{\tau 2}$.

The uncertainty on the B mass (σ_{FIT}) is taken from the distribution in the data (figure B-1a). For the signal events, the normalized mass (M_N) is generated according to $G(M_N; 0, X)$, except for $J/\psi K^{*0}$, where a fraction P_S of the events are labeled "swapped", and have M_N generated according to $G(M_N; \mu_S, X_S)$. The background events have M_N generated randomly, according to the relevant linear distribution. With the mass uncertainty and the normalized mass, the B mass is calculated: $M_{FIT} = M_N \times \sigma_{FIT} + M_0$.

For prompt events, the "true" (unsmeared) decay time (t') is always zero. For other events, the true decay time is taken from the relevant exponential distribution. The uncertainty on the decay time (σ_t) is taken from the distribution from the data (figure B-1b). The true decay time is smeared by this uncertainty to get the "measured" decay time t: $P(t) = G(t; t', Y\sigma_t)$, where Y is the decay-time uncertainty

---

[1] Actually the variation is Gaussian, with width $\sqrt{N_{evt}}$. This is equivalent to the Poisson distribution when N_{evt} is large, and it is simpler to calculate.

scale-factor. This smearing is identical for both signal and background.

Then the "reconstructed flavor" (r) is determined randomly, according to equation 8.19. Then the "produced flavor" (p) is determined randomly, according to equation 8.20. For $J/\psi K_S^0$ and $J/\psi K^{*0}$ signal events, this includes the time-dependent asymmetry, based on the "true" decay time (t'), not the "reconstructed" decay time (t). For the $J/\psi K^+$ signal and all backgrounds, p is always equal to r. For the swapped $J/\psi K^{*0}$ events, the sign of r (but not p) is swapped.

The dilution $\mathcal{D}$ and efficiency ϵ are TMC input parameters. For signal events, $p_T(tag)$ and n_{PI} are taken from the histograms from the data (figures B-1c and B-1d). These are used to determine α, according to equation 9.10. The fit uses $\gamma/\alpha = 1$, so γ is set to α. For the background, α and $\delta \equiv \mathcal{D}\gamma$ are TMC input parameters.[2]

The event is then tagged with probability $\epsilon(1 + p\mathcal{D}\gamma)$. If the event is tagged, it is assigned positive charge with probability:

$$\frac{P(+)}{P(+) + P(-)} = \frac{1 + p\mathcal{D} + \alpha + p\mathcal{D}\gamma}{1 + p\mathcal{D}\gamma}. \tag{B.1}$$

These probabilities are derived from equation 8.30.

For each event, the variables M_{FIT}, σ_{FIT}, t, σ_t, r, s, $p_T(tag)$, and n_{PI} are saved, as for the data. In addition, three more variables are saved, p, t', and ID, where p and t' are defined above, and ID is a label for the type of event: unswapped signal, swapped signal, prompt background, negative-tail background, short-tail background, or long-tail background. These three additional variables are used in tests of the TMC.

B.2 Testing the TMC

Before testing the likelihood fit, tests are first performed on the Toy Monte Carlo, to make certain that it works as expected (*i.e.* that it generates all the relevant variables with the proper distributions). To perform these tests, the TMC is used to generate many samples, and each of these samples is scanned to determine how well it matches

---

[2] For the $J/\psi K_S^0$ background, $\mathcal{D} = \delta = 0$, as always.

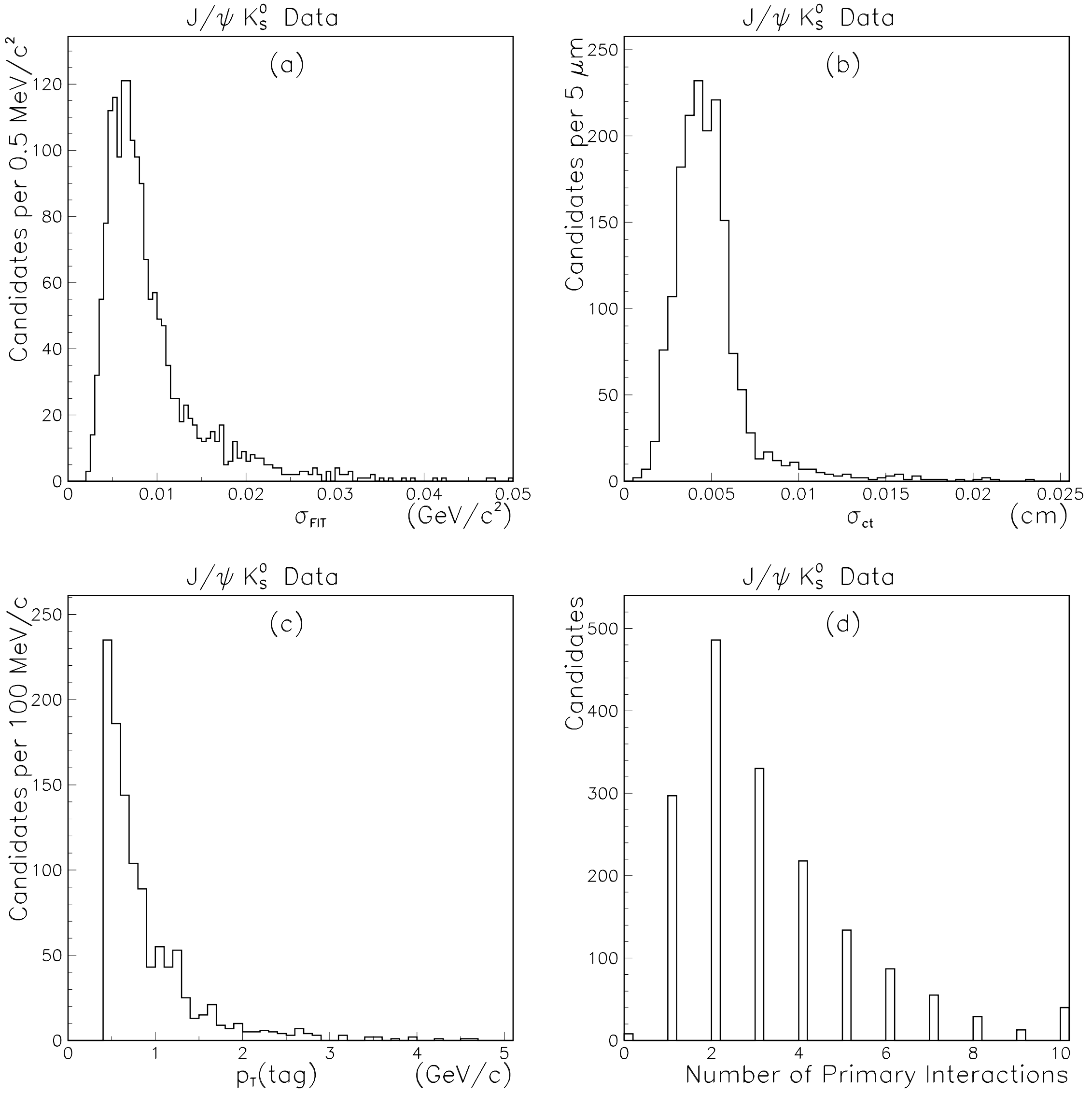

Figure B-1: Distributions of variables from the data used as input to the Toy Monte Carlo. Plot (a) shows the distribution of σ_{FIT}, (b) shows the distribution of $\sigma_{ct} \equiv c \times \sigma_t$, (c) shows the distribution of $p_T(tag)$, and (d) shows the distribution of n_{PI}, the number of primary interactions identified. These distributions are from the $J/\psi K^0_S$ data, and the TMC uses them only when generating $J/\psi K^0_S$. When generating $J/\psi K^{*0}$ or $J/\psi K^+$, the distributions used are the ones from the relevant data sample. These distributions are similar to the above distributions for $J/\psi K^0_S$.

the input specifications. Each of the parameters is calculated directly, thus excluding correlations with other parameters.

MC Check Procedure

The calculated value of the signal fraction f_B is simply the fraction of generated events which have ID indicating they are signal events. Because the generation is a random process, this fraction is not expected to be exactly the same for each sample, but the mean should match the input value (which is taken from the fit to the data). Similarly, using the variable ID, the other fractions are calculated: f_L, f_N and $f_{\tau 2}$.

The calculated background "lifetimes" τ_1 and τ_2 are the average values of $|t'|$ for the events with the relevant ID. The effects of smearing are not considered for the lifetime calculations, but they are used to calculate the decay-time resolution scale-factor Y, which is the RMS of $(t - t')/\sigma_t$, for all events.

Similarly, the calculated mass error scale-factor X is the RMS of the M_N values for those events identified as unswapped signal. The mass slopes for the background (ζ_P and ζ_L) are related to the mean values of M_N for each set of events:

$$\langle M_N \rangle = \frac{1}{2}\int_{-20}^{20} M_N(1 + \zeta M_N)dM_N = \zeta \times (400/3) \tag{B.2}$$

Since the TMC includes both s and p, the six probabilities $\mathcal{P}_3(s|p)$ can each be calculated directly for each type of event. From these probabilities, the values of ϵ, α, $\mathcal{D}$ and γ are calculated via equations 8.22 to 8.25.

TMC Tests

To test the TMC, 2500 $J/\psi K_S^0$ samples, 1000 $J/\psi K^{*0}$ samples, and 1000 $J/\psi K^+$ samples are generated. The $J/\psi K_S^0$ samples are all generated with $\sin 2\beta = 1$; tests with varying values of $\sin 2\beta$ will be discussed in Section B.5. Each of these samples is run through the "MC Check" procedure described above. Tables B.1 to B.6 display the results of these checks. The first column lists the parameter name, and the second column lists the input value used for that parameter. The third column lists the average value returned by the MC Check, and the fourth column lists the significance of the difference between the input and output values. All of the deviations are consistent with the expected statistical fluctuations, indicating that the TMC generation works as intended.

As an additional check a very large sample of $J/\psi K_S^0$, and another of $J/\psi K^{*0}$, are generated, and time-dependent asymmetries are measured in each sample. Figure B-2 shows these distributions. For $J/\psi K_S^0$, the asymmetry measured is the number of $\overline{B}^0$ versus B^0 generated (based on p), for each bin in t'. For $J/\psi K^{*0}$, the asymmetry measured is the number of unmixed ($r = p$) versus mixed ($r \neq p$) signal events, for each bin in t'. The curves in the figures indicate the expected asymmetries, and the measured values follow the expectations very well.

B.3 Testing the Likelihood Fit

The same samples used to test the Toy Monte Carlo generator are used to test the likelihood fit. Each of these samples is fit in exactly the same way as the data. Tables B.1 to B.6 display the results of these tests, in addition to the MC Checks described above. The fifth column ("Full fit") in each table shows the average fit value, and the sixth column shows the significance of the deviation from the "Input" value. While a few parameters do show significant deviations, most are quite consistent with the expected statistical fluctuations. In particular, the parameters of interest (the signal dilution measurements $\mathcal{D}_B$) show no significant deviations. The significant deviations indicate *fit biases*, which will be discussed in Section B.4.

These checks test whether the likelihood fit returns biased results, but they do not test the statistical uncertainties returned by the fits. To test these, the uncertainties ("errors") are compared with the spreads (RMS's) of fit values returned by the TMC tests. The variations in the fit values are due to statistical fluctuations and are therefore true measures of the statistical uncertainties on the fit parameters. Tables B.7 to B.12 display the comparisons between the uncertainties and the spreads for the fit parameters. The agreement is very good.

The RMS of the TMC fit results is 1% higher, on average, than the mean TMC

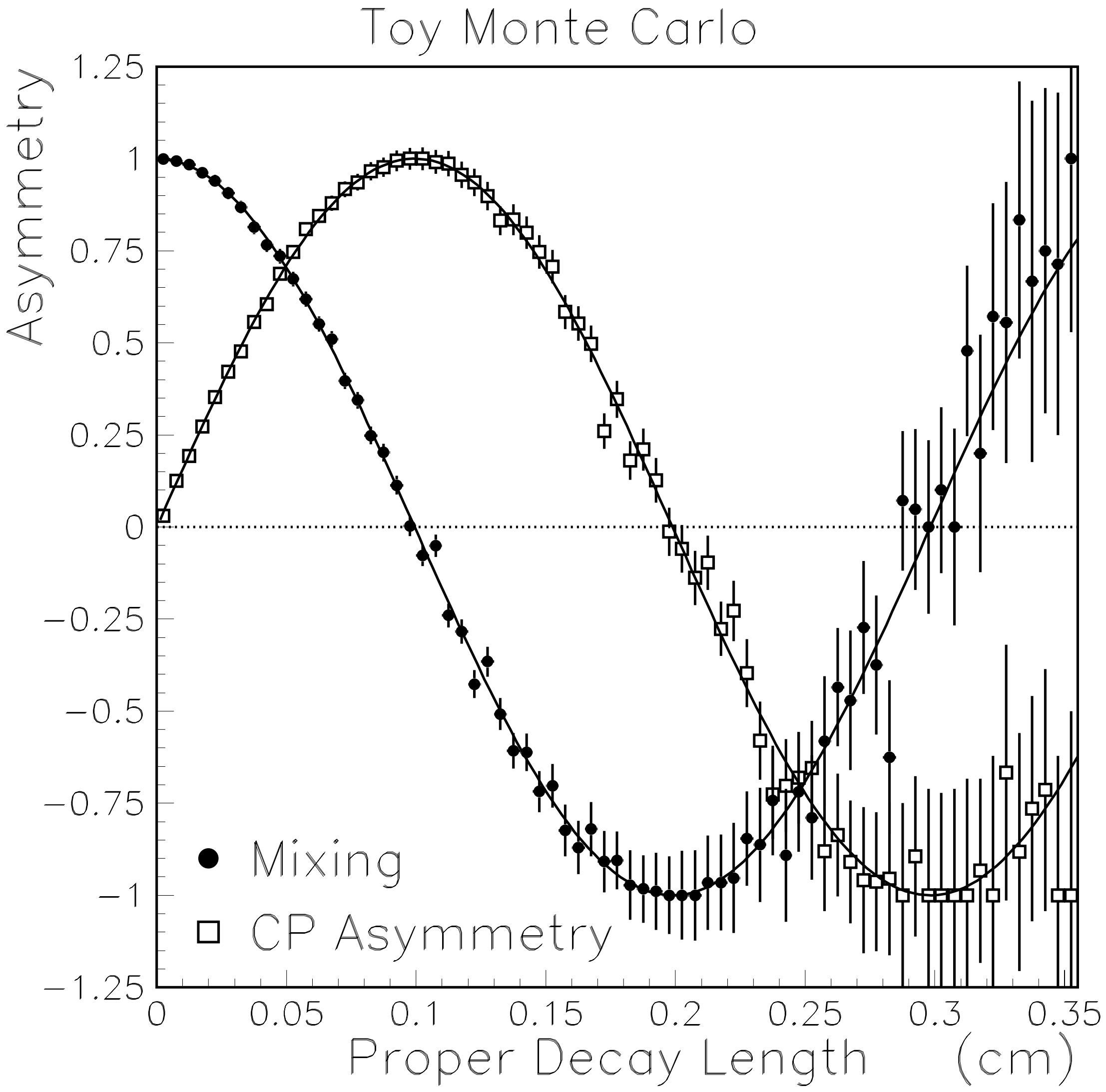

Figure B-2: Asymmetries in the Toy Monte Carlo. The mixing asymmetry is in the number of unmixed versus mixed signal events in $J/\psi K^{*0}$. The CP asymmetry is the in the number of $\overline{B}^0$ versus B^0 in $J/\psi K_S^0$. The horizontal scale is the "true" (unsmeared) decay length. The curves indicate the expected asymmetries. The statistics are several hundred times higher than those in the data.

error. The largest differences are in the background lifetimes and fractions, terms which also tend to indicate fit biases. The mechanism that causes the fit biases might also cause the uncertainties to be mis-estimated by a few percent.

The uncertainties from the data agree very well with the average uncertainties from the TMC. The agreement is not expected to be perfect, as the uncertainties in the data are each the results of one experiment, while the Monte Carlo numbers are from many experiments. While the average of the TMC *fit* values should match the value from the data very accurately, the correlations between the *uncertainties* are not so strong. Different samples with identical fit parameters will generally have slightly different uncertainties on those parameters. The uncertainties in the data are all close to those in the TMC, and are well within the range of variation of the TMC uncertainties. Thus, the uncertainties returned by the data should be accurate estimates of the uncertainties on the fit parameters.

B.4 Fit Biases

A fit bias can result when two fit parameters are correlated with one another. For example, the top plot in figure B-3 shows the distribution of the RMS values calculated for the *swapped* $J/\psi K^{*0}$ events. Each entry in the histogram represents one TMC sample. The input RMS is 5.0, and the average output value is 4.91 ± 0.02, which is low by several standard-deviations. The mean (μ_1) and the RMS (R_1) for each of these samples are calculated according to the following formulae:

$$\mu_1 = \frac{\sum M_N}{N_S} \tag{B.3}$$

$$R_1 = \sqrt{\frac{\sum (M_N - \mu_1)^2}{N_S}}, \tag{B.4}$$

where N_S is the number of swapped events.

Unfortunately, fluctuations in μ_1 will affect the value of R_1. If the RMS is calculated with the mean fixed to the correct value:

$$\mu_2 = \mu_S \tag{B.5}$$

$$R_2 \;=\; \sqrt{\frac{\sum(M_N-\mu_2)^2}{N_S}} \tag{B.6}$$

(where $\mu_S = -0.5$ is the input value for the mean), then the bias is removed. The bottom plot in figure B-3 shows the distribution of R_2, which has a mean value of 4.99 ± 0.02, quite consistent with the input value of 5.

Using the correspondence: $R_2^2 - R_1^2 \equiv (\mu_1 - \mu_2)^2$, one can calculate the expected difference between these RMS's:

$$\langle R_2 - R_1 \rangle \;=\; \left\langle \frac{(\mu_1-\mu_2)^2}{R_1+R_2} \right\rangle \simeq \frac{\sigma_\mu^2}{2R} = \frac{R}{2N_S} \tag{B.7}$$

where $R = 5$ is the input value for the RMS, and $\sigma_\mu = R/\sqrt{N_S}$ is the spread (RMS) of the measured values of μ_1. The average value of N_S is 36.5 (there are 365 $J/\psi K^{*0}$ signal events and the fraction that are swapped is $P_S = 0.1$), so the average value of $R_2 - R_1 \simeq 0.07$, which is almost exactly the difference between the above values.

Thus, when μ and R are fit for simultaneously, there will be a small bias in the value of R, but when μ is fixed to the input value, no bias results. The bias goes to zero as $N \to \infty$, so it is simply an effect of limited statistics.

B.4.1 Fit Biases in Long-Lived Background

A similar effect occurs when one attempts to fit the decay-time distribution of a sample of events to a prompt peak and a long-lived tail, letting both the fraction of events in the tail and the "lifetime" of the tail float as fit parameters. Further studies of this effect have indicated that the bias decreases with $1/N$ and disappears when one of the two parameters is fixed. This is exactly the same behavior as for the bias on R, above, indicating the effect is due to correlations in a sample with limited statistics. When fitting is done as in the data (with three fractions and two lifetimes; all five parameters floating), some bias is guaranteed to result.

In the TMC tests, the background fractions and "lifetimes" often show significant deviations from the input value, as indicated above. However, the deviations only become noticeably significant after many simulations are run, and the largest of them is ≈ 1/7th the size of the relevant statistical uncertainty. Furthermore, none of these parameters is highly correlated with the parameters of interest. Therefore, these biases are ignored.

B.5 Fits With Varying $\sin 2\beta$

The results above indicate that the likelihood fit works when $\sin 2\beta = 1$, and that the expected statistical fluctuations are equal to the uncertainty returned by the data. In order to rule out a bias whose effect varies with $\sin 2\beta$, additional TMC samples are generated for several values of $\sin 2\beta$. Figure B-4 shows the distributions of measured raw CP asymmetries, for four different values of $\sin 2\beta$: 0, 0.25, 0.5, and 0.75. Table B.13 displays the means and RMS's of the fit values of $\mathcal{D} \sin 2\beta$ for each of the TMC generations. The means are all consistent with the input values, and the distributions in figure B-4 are all Gaussian. The widths of these distributions are all very similar, and show no particular dependence on the input value of $\sin 2\beta$. Therefore, it is concluded that the effectiveness of the likelihood fit does not depend on the actual value of $\sin 2\beta$.

B.6 Fit Quality Test

The quantity $F \equiv -2\ln(\mathcal{L}_T)/N_{evt}$ is the average contribution to the χ^2 for each event. This is a measure of the fit quality, as a high average χ^2 indicates that the shape of the likelihood function does not describe the sample well. As the TMC samples are generated using the likelihood function one would expect the likelihood fit to describe them well. If the data is described well by the likelihood function, then it should have a value of F similar to that of the TMC. If the data is not well described by the likelihood function, then the average χ^2 per event should be high, and F should be higher than the value from the TMC.

Figure B-5 shows the distributions of F for the TMC samples used in the above fit tests. The average value of F for the $J/\psi K^0_S$ TMC is 5.45, the value for the data is 5.46, and the fraction of TMC samples with higher F than the data is 44%. The

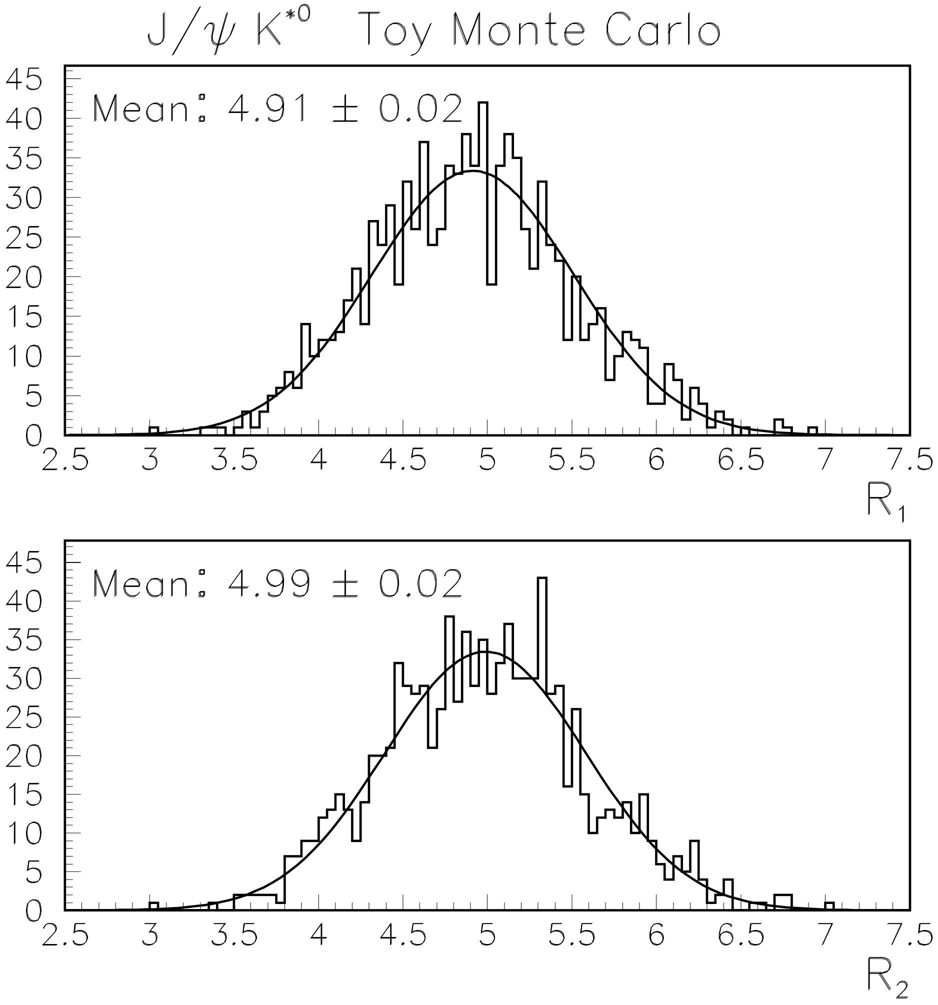

Figure B-3: Calculations of the RMS of the M_N distributions for $J/\psi K^{*0}$ Toy Monte Carlo samples. The first calculation (R_1, top plot) allows the mean to float and be determined from the sample. The second calculation (R_2, bottom plot) fixes the mean to the input value. The samples are generated with RMS of 5, so R_2 is an unbiased estimator of the RMS, and R_1 is a biased estimator.

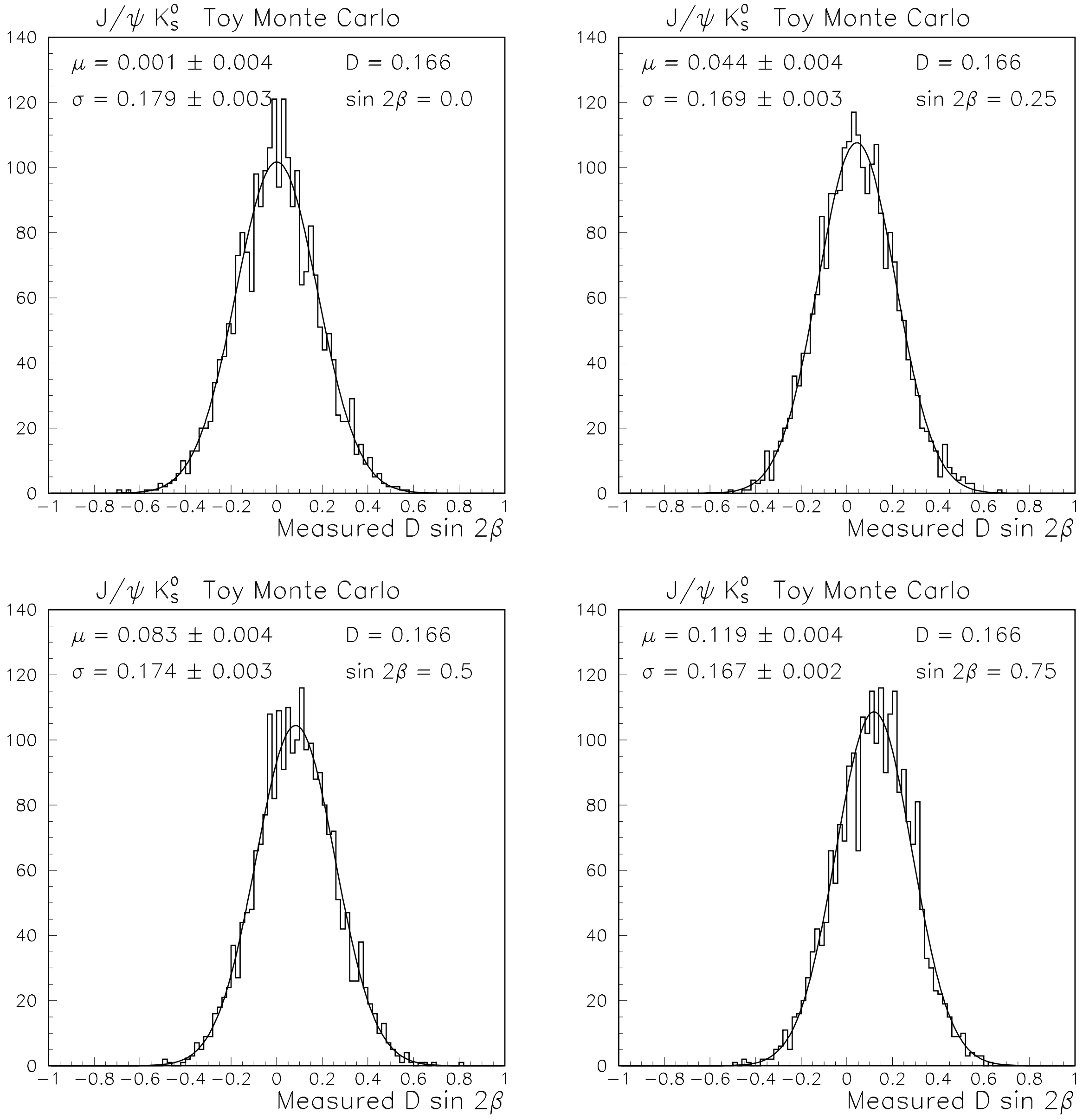

Figure B-4: Distributions of raw CP asymmetries fit for in TMC samples with varying input values for $\sin 2\beta$. The top-left plot is generated with $\sin 2\beta = 0$, top-right with $\sin 2\beta = 0.25$, bottom-left with $\sin 2\beta = 0.5$, and bottom-right with $\sin 2\beta = 0.75$. The curves represent Gaussian fits to the distributions, with the indicated mean (μ) and width (σ). Each mean is consistent with the respective value of $\mathcal{D} \sin 2\beta$.

average value of F for $J/\psi K^{*0}$ is 5.54, the value for the data is 5.61, and the fraction of TMC samples with higher F than the data is 13%. The average value of F for the $J/\psi K^{+}$ is 5.39, the value for the data is 5.42, and the fraction of TMC samples with higher F than the data is 6.3%.

All three decay modes have average values of F for the TMC near 5.5, and very few of the TMC samples have $F < 5.25$ or $F > 5.75$. If the likelihood function is not a perfect description of the distributions in the data, then one would expect the data to have higher values of F than the average for the TMC, and the difference should be more pronounced in the higher-statistics samples. While this is the case, the three data samples do have values of F within the ranges spanned by the TMC. If the likelihood does not describe the data perfectly, it at least describes the data very well. Thus, it is concluded that the quality of the likelihood fit is adequate for the analysis.

B.7 Conclusions from Toy Monte Carlo Studies

A Toy Monte Carlo has been constructed to test the likelihood function. The TMC generates samples of events using the likelihood function, with the fit parameters set to the results of fitting the data. The TMC has been tested, and the tests indicate that the TMC works as intended.

Each of the TMC samples is fit exactly like the data. The results of the fits indicate that the likelihood fit returns unbiased estimators for most of the fit parameters. The parameters that do show bias are the background fractions and "lifetimes." These biases appear to be due to statistical correlations between these fit parameters, and are small ($\leq 1/7$th of the statistical uncertainty). The biases do not affect the parameters of interest nor any parameters significantly correlated with the parameters of interest, and are therefore ignored.

Tests have been performed for 5 different values of $\sin 2\beta$ ranging from 0 to 1, and each test indicates that the returned raw CP asymmetry is unbiased and has statistical uncertainty consistent with the uncertainty indicated by the data.

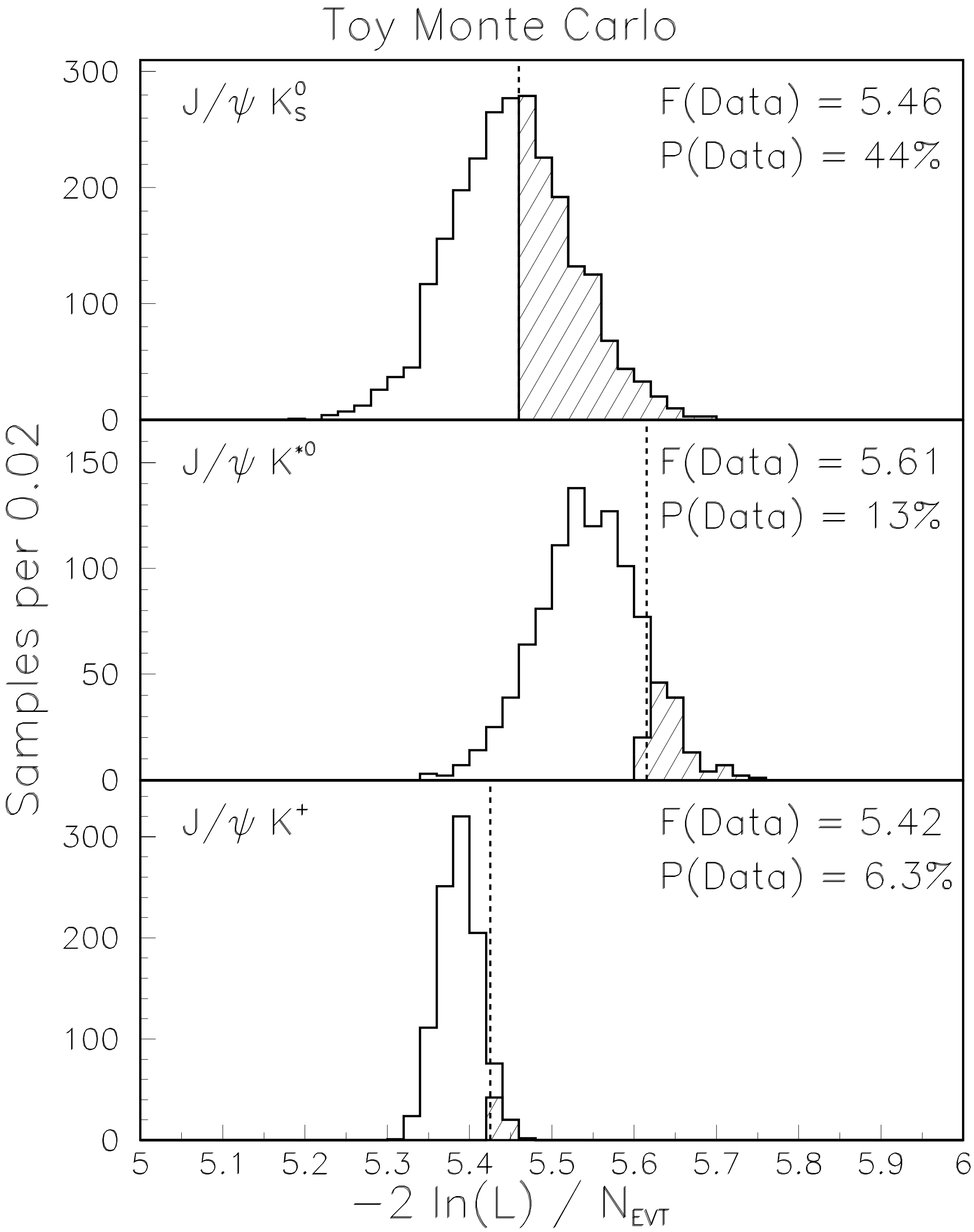

Figure B-5: Distributions of $F \equiv -2\ln(\mathcal{L})/N_{evt}$ for the $J/\psi K^0_S$, $J/\psi K^{*0}$, and $J/\psi K^+$ Toy Monte Carlo samples. The vertical dashed lines indicate the values of F for the data, and the hashed histograms indicate the TMC samples which have F larger than that in the data.

The fit-quality number $F \equiv -2\ln(\mathcal{L})/N_{evt}$ is compared between the data and the TMC. The comparisons indicate that, while the likelihood function might not describe the data perfectly, the description is very good.

All of these comparisons indicate that the likelihood function performs as intended. The returned fit values (especially for the parameters of interest) are unbiased estimators of the fit parameters, and the statistical uncertainties returned by the fit do reflect the actual statistical fluctuations expected.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Background Mass Slopes					
ζ_P	0.0094	0.0094	-0.28	0.0095	1.40
ζ_L	-0.0066	-0.0066	-0.18	-0.0067	-1.08
Event Fractions					
f_B	0.117	0.117	-0.04	0.117	0.86
f_L	0.262	0.262	2.16	0.265	5.63
f_N	0.201	0.201	-0.07	0.204	4.88
$f_{\tau 2}$	0.550	0.549	-0.83	0.537	-5.19
Long-Lived Background "Lifetimes"					
τ_1 (cm)	0.042	0.042	0.42	0.043	2.65
τ_2 (cm)	0.011	0.011	-0.65	0.011	-0.22
Error Scale-Factors					
X	1.392	1.389	-2.13	1.395	0.96
Y	0.971	0.971	-0.08	0.969	-2.13

Table B.1: Table of kinematic parameters for the $J/\psi K_S^0$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column. Some of the "full fit" results for the long-lived background do indicate significant biases; these biases are discussed in Section B.4.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Tagging Efficiencies					
ϵ_B	0.615	0.615	0.19	0.615	0.30
ϵ_P	0.626	0.626	0.21	0.625	-1.55
ϵ_L	0.751	0.751	0.55	0.752	0.91
Background Tagging Charge Biases					
α_P	-0.006	-0.006	0.35	-0.006	0.03
α_L	0.167	0.166	-0.35	0.165	-1.18
Dilutions					
$\mathcal{D}_B$	0.166	0.166	0.22	0.170	1.14

Table B.2: Table of tagging parameters for the $J/\psi K_S^0$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column. The samples used in this study all have $\sin 2\beta = 1$, so $\mathcal{D}_B \sin 2\beta = \mathcal{D}_B$.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Background Mass Slopes					
ζ_P	0.0110	0.0110	0.47	0.0110	0.45
ζ_L	-0.0123	-0.0123	-0.16	-0.0123	0.58
Event Fractions					
f_B	0.067	0.067	1.16	0.067	0.37
f_L	0.160	0.160	1.56	0.161	3.66
f_N	0.137	0.137	1.12	0.138	2.00
$f_{\tau 2}$	0.781	0.781	-0.56	0.776	-4.52
Long-Lived Background "Lifetimes"					
τ_1 (cm)	0.059	0.059	-0.25	0.059	-1.32
τ_2 (cm)	0.013	0.014	0.60	0.013	-3.28
Error Scale-Factors					
X	1.343	1.342	-0.47	1.341	-1.30
Y	0.987	0.988	1.86	0.987	-0.47
Reconstruction Asymmetries					
R_B	0.077	0.077	-0.14	0.078	0.60
R_P	0.003	0.003	0.59	0.003	-0.16
R_L	0.030	0.030	-0.41	0.030	0.00

Table B.3: Table of kinematic parameters for the $J/\psi K^+$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column. Some of the "full fit" results for the long-lived background do indicate significant biases; these biases are discussed in Section B.4.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Tagging Efficiencies					
ϵ_B	0.624	0.623	-0.87	0.623	-0.32
ϵ_P	0.703	0.703	0.29	0.703	0.15
ϵ_L	0.771	0.771	-1.00	0.770	-1.58
Background Tagging Charge Biases					
α_P	0.033	0.033	0.55	0.033	1.20
α_L	0.015	0.014	-0.35	0.014	-0.79
Background Tagging Efficiency Asymmetries					
δ_P	-0.002	-0.001	0.94	-0.001	1.37
δ_L	-0.026	-0.026	1.35	-0.026	0.61
Dilutions					
$\mathcal{D}_B$	0.185	0.185	-0.03	0.185	0.13
$\mathcal{D}_P$	-0.069	-0.069	0.99	-0.069	0.93
$\mathcal{D}_L$	-0.089	-0.090	-1.05	-0.089	-0.15

Table B.4: Table of tagging parameters for the $J/\psi K^+$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Background Mass Slopes					
ζ_P	0.0044	0.0043	-0.71	0.0044	-0.04
ζ_L	-0.0176	-0.0176	0.00	-0.0176	0.15
Event Fractions					
f_B	0.156	0.156	0.77	0.156	1.07
f_L	0.222	0.222	-0.71	0.224	3.15
f_N	0.096	0.096	-0.47	0.097	1.45
$f_{\tau 2}$	0.626	0.628	1.95	0.624	-0.46
Long-Lived Background "Lifetimes"					
τ_1 (cm)	0.037	0.037	-0.26	0.038	4.07
τ_2 (cm)	0.010	0.010	0.30	0.010	0.62
Error Scale-Factors					
X	1.540	1.542	1.00	1.545	1.65
Y	1.057	1.057	0.15	1.055	-2.13
Reconstruction Asymmetries					
R_B	-0.086	-0.087	-0.57	-0.090	-1.88
R_P	0.036	0.037	1.33	0.037	1.18
R_L	0.095	0.098	1.44	0.099	1.44

Table B.5: Table of kinematic parameters for the $J/\psi K^{*0}$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column. Some of the "full fit" results for the long-lived background do indicate significant biases; these biases are discussed in Section B.4.

Parameter	Input	MC Check	Δ/σ	Full fit	Δ/σ
Tagging Efficiencies					
ϵ_B	0.635	0.635	-0.80	0.634	-1.40
ϵ_P	0.830	0.830	-0.45	0.830	-0.26
ϵ_L	0.778	0.778	0.89	0.779	1.85
Background Tagging Charge Biases					
α_P	0.092	0.092	0.84	0.092	0.36
α_L	-0.044	-0.045	-0.34	-0.043	0.75
Background Tagging Efficiency Asymmetries					
δ_P	0.012	0.013	0.77	0.013	0.04
δ_L	-0.029	-0.028	0.73	-0.027	1.75
Dilutions					
$\mathcal{D}_B$	0.165	0.164	-0.32	0.169	1.07
$\mathcal{D}_P$	-0.003	-0.002	0.67	-0.002	1.12
$\mathcal{D}_L$	-0.050	-0.051	-0.76	-0.053	-1.37

Table B.6: Table of tagging parameters for the $J/\psi K^{*0}$ TMC. The "Input" column indicates the value returned by fitting the data, which is used as TMC input. The "MC Check" column indicates the results of the MC Checks, described in Section B.2, and the "Full fit" column indicates the results of the likelihood fit tests described in Section B.3. The "Δ/σ" columns indicate the significances of the differences between the previous columns and the "Input" column.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Background Mass Slopes			
ζ_P	0.0030	0.0030	0.0030
ζ_L	0.0059	0.0059	0.0060
Event Fractions			
f_B	0.010	0.010	0.010
f_L	0.032	0.030	0.031
f_N	0.037	0.039	0.040
$f_{\tau 2}$	0.108	0.119	0.121
Long-Lived Background "Lifetimes"			
τ_1 (cm)	0.006	0.007	0.007
τ_2 (cm)	0.002	0.002	0.002
Error Scale-Factors			
X	0.108	0.111	0.112
Y	0.032	0.032	0.032

Table B.7: Table of uncertainties on kinematic parameters for the $J/\psi K^0_S$ TMC. The "Data Error" column represents the uncertainty returned by the fit to the data. The "Mean MC Error" column indicates the average uncertainty returned by the TMC. The "MC Fit RMS" column indicates the spread of the fit values returned by the TMC.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Tagging Efficiencies			
ϵ_B	0.041	0.041	0.041
ϵ_P	0.017	0.017	0.017
ϵ_L	0.034	0.032	0.033
Background Tagging Charge Biases			
α_P	0.045	0.046	0.048
α_L	0.082	0.083	0.083
Dilutions			
$\mathcal{D}_B$	0.179	0.168	0.168

Table B.8: Table of uncertainties on tagging parameters for the $J/\psi K_S^0$ TMC. The "Data Error" column represents the uncertainty returned by the fit to the data. The "Mean MC Error" column indicates the average uncertainty returned by the TMC. The "MC Fit RMS" column indicates the spread of the fit values returned by the TMC.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Background Mass Slopes			
ζ_P	0.0009	0.0009	0.0009
ζ_L	0.0027	0.0027	0.0027
Event Fractions			
f_B	0.003	0.003	0.003
f_L	0.008	0.007	0.007
f_N	0.014	0.013	0.013
$f_{\tau 2}$	0.030	0.036	0.036
Long-Lived Background “Lifetimes”			
τ_1 (cm)	0.005	0.006	0.006
τ_2 (cm)	0.001	0.001	0.001
Error Scale-Factors			
X	0.053	0.052	0.053
Y	0.010	0.009	0.009
Reconstruction Asymmetries			
R_B	0.041	0.041	0.041
R_P	0.011	0.011	0.011
R_L	0.034	0.033	0.034

Table B.9: Table of uncertainties on kinematic parameters for the $J/\psi K^+$ TMC. The “Data Error” column represents the uncertainty returned by the fit to the data. The “Mean MC Error” column indicates the average uncertainty returned by the TMC. The “MC Fit RMS” column indicates the spread of the fit values returned by the TMC.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Tagging Efficiencies			
ϵ_B	0.020	0.020	0.019
ϵ_P	0.005	0.005	0.005
ϵ_L	0.014	0.014	0.014
Background Tagging Charge Biases			
α_P	0.013	0.013	0.013
α_L	0.037	0.037	0.038
Background Tagging Efficiency Asymmetries			
δ_P	0.007	0.007	0.007
δ_L	0.018	0.018	0.018
Dilutions			
$\mathcal{D}_B$	0.052	0.053	0.053
$\mathcal{D}_P$	0.013	0.013	0.013
$\mathcal{D}_L$	0.038	0.037	0.038

Table B.10: Table of uncertainties on tagging parameters for the $J/\psi K^+$ TMC. The "Data Error" column represents the uncertainty returned by the fit to the data. The "Mean MC Error" column indicates the average uncertainty returned by the TMC. The "MC Fit RMS" column indicates the spread of the fit values returned by the TMC.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Background Mass Slopes			
ζ_P	0.0024	0.0025	0.0025
ζ_L	0.0053	0.0053	0.0054
Event Fractions			
f_B	0.009	0.009	0.009
f_L	0.023	0.022	0.022
f_N	0.029	0.027	0.028
$f_{\tau 2}$	0.104	0.112	0.118
Long-Lived Background "Lifetimes"			
τ_1 (cm)	0.006	0.007	0.008
τ_2 (cm)	0.002	0.002	0.002
Error Scale-Factors			
X	0.101	0.094	0.091
Y	0.028	0.026	0.025
Reconstruction Asymmetries			
R_B	0.068	0.068	0.069
R_P	0.029	0.029	0.028
R_L	0.070	0.069	0.068

Table B.11: Table of uncertainties on kinematic parameters for the $J/\psi K^{*0}$ TMC. The "Data Error" column represents the uncertainty returned by the fit to the data. The "Mean MC Error" column indicates the average uncertainty returned by the TMC. The "MC Fit RMS" column indicates the spread of the fit values returned by the TMC.

Parameter	Data Error	Mean MC Error	MC Fit RMS
Tagging Efficiencies			
ϵ_B	0.030	0.029	0.029
ϵ_P	0.011	0.011	0.011
ϵ_L	0.031	0.030	0.030
Background Tagging Charge Biases			
α_P	0.031	0.031	0.032
α_L	0.079	0.079	0.080
Background Tagging Efficiency Asymmetries			
δ_P	0.013	0.013	0.013
δ_L	0.038	0.038	0.038
Dilutions			
$\mathcal{D}_B$	0.112	0.113	0.114
$\mathcal{D}_P$	0.031	0.031	0.031
$\mathcal{D}_L$	0.079	0.080	0.081

Table B.12: Table of uncertainties on tagging parameters for the $J/\psi K^{*0}$ TMC. The "Data Error" column represents the uncertainty returned by the fit to the data. The "Mean MC Error" column indicates the average uncertainty returned by the TMC. The "MC Fit RMS" column indicates the spread of the fit values returned by the TMC.

Input Value		Mean fit value	RMS of fit values
$\sin 2\beta$	$\mathcal{D}\sin 2\beta$		
0.0	0.0	0.001 ± 0.004	0.179 ± 0.003
0.25	0.042	0.044 ± 0.004	0.169 ± 0.003
0.5	0.083	0.083 ± 0.004	0.174 ± 0.003
0.75	0.125	0.119 ± 0.004	0.167 ± 0.002
1.0	0.166	0.170 ± 0.003	0.169 ± 0.002

Table B.13: Means and RMS's of the returned values of $\mathcal{D}\sin 2\beta$ for TMC fits. All five samples were generated with $\mathcal{D} = 0.166$.

Bibliography

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

[2] J. H. Christenson *et al.*, *Phys. Rev. Lett.* **13**, 138 (1964).

[3] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev.***D2**, 1585 (1970).

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

[5] J. J. Aubert *et al.*, *Phys. Rev. Lett.* **33**, 1404 (1974)

[6] J. E. Augustin *et al.*, *Phys. Rev. Lett.* **33**, 1406 (1974)

[7] S. W. Herb *et al.*, *Phys. Rev. Lett.* **39**, 252 (1977)

[8] F. Abe *et al.*, *Phys. Rev. Lett.* **74**, 2626 (1995)

[9] K. Ackerstaff *et al.*. CERN-EP-98-001, Jan 1998

[10] A. Pais, *Inward Bound*, Ch. 20, p. 511–549, Oxford University Press (1986).

[11] E. Noether, "Invariante Variationsprobleme", *Nach. v. d. Ges. d. Wiss. zu Göttingen*, 253-257 (1918).
English translation: M. A. Tavel, *Transport Theory and Statistical Physics* **1**(3), 183 (1971).

[12] F. Gursey, *Emmy Noether, Collected Papers*, Springer–Verlag (1983).

[13] E. P. Wigner, *Goett. Nachr.*, 375 (1927).

[14] E. P. Wigner, *Goett. Nachr.*, 375 (1927);
English Translation: A. Pais, *Inward Bound*, Ch. 20, p. 526, Oxford University Press (1986).

[15] O. Laporte, *Zeitschr. f. Phys.* **23**, 135 (1924).

[16] E. P. Wigner, *Goett. Nachr.*, 546 (1932).

[17] A. Pais, *Inward Bound*, Ch. 20, p. 527, Oxford University Press (1986).
Original source: H. A. Kramers, *Proc.. Ak. v. Wet. Amsterdam* **33**, 959 (1930); repr. in *H. A. Kramers, collected scientific papers*, p. 522, North Holland, Amsterdam (1956).

[18] C. D. Anderson, "The Positive Electron," *Phys. Rev.* **43**, 491 (1933).

[19] M. Gell-Mann and A. Pais, *Phys. Rev.* **97**, 1387 (1955).

[20] K. Lande *et al.*, *Phys. Rev.* **103**, 1901 (1956).

[21] F. Eisler *et al.*, *Nuovo Cim.* **5**, 1700 (1957).

[22] E. Boldt *et al.*, *Phys. Rev.* **112**, 1746 (1958).

[23] R. Jost, *Helv. Phys. Act.* **30**, 409 (1957).

[24] F. J. Dyson, *Phys. Rev.* **110**, 579 (1958).

[25] G. Lüders, *Kong Dansk. Vid. Selskab, Mat.-Fys. Medd.*, **28**, 5 (1954)

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

[27] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Phys. Rev.* **105**, 1415 (1957).

[28] J. Friedman and V. Telegdi, *Phys. Rev.* **105**, 1681 (1957).

[29] M. Bardon *et al.Ann. Phys.* **5**, 156 (1958).

[30] D. Neagu *et al.Phys. Rev. Lett.* **6**, 552 (1961).

[31] G. Kane *Modern Elementary Particle Physics*, Addison-Wesley (1987)

[32] Particle Data Group *Phys. Rev. D* **54** (1996)

[33] M. Gell-Mann, "A Schematic Model of Baryons and Mesons", *Phys. Lett.* **8**, 214-215 (1964).

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

[35] L. Wolfenstein, Phys. Rev. Lett. **51**, 1915 (1983).

[36] Original paper: P. Paganini, F. Parodi, P. Roudeau, and A. Stocchi "Measurements of the ρ and η parameters of the V_{CKM} matrix and perspectives." SLAC Preprint: hep-ph/9711261, submitted to *Phys. Scripta* (1997).

Updated measurements: F. Parodi, P. Roudeau, and A. Stocchi "Constraints on the parameters of the V_{CKM} matrix at the end of 1997" SLAC Preprint: hep-ph/9802289, submitted to *Phys. Scripta* (1998)

[37] L. L. Chau and W. Y. Keung, Phys. Rev. Lett.**53**, 1802 (1984).

[38] I. Dunietz, Personal communication.

[39] T. Sjöstrand, *Int. Jour. of Mod. Phys. A* **3**, 715 (1988)

[40] F. Abe *et al.*, "Observation of the B_c Meson in $b\bar{b}$ Collisions at $\sqrt{s} = 1.8$TeV", FERMILAB-Pub-98/157-E, Submitted to *Phys. Rev. Lett.*

[41] P. Nason, S. Dawson, R. K. Ellis, *Nucl. Phys. B* **327**, 49 (1989)

[42] B.Todd Huffman for the CDF collaboration. FERMILAB-CONF-92-337-E, Nov 1992.
Published in *DPF Conf. 1992*, 771 (1992)

[43] F. Abe *et al.*, *Phys. Rev. Lett.* **79**, 572 (1997)

[44] F. Abe *et al.*, *Phys. Rev. Lett.* **71**, 3421 (1993)

[45] F. Abe *et al.*, *Phys. Rev. Lett.* **80**, 2057 (1998).
See also P. Maksimovic, Ph.D. dissertation, Massachusetts Institute of Technology, 1998 (unpublished), available at:
`http://www-cdf.fnal.gov/grads/thesis_complete.html`.

[46] F. Abe *et al.*, FERMILAB-Pub-98/188-E (1998)

[47] C. Gay, G. Michail, R. Stroehmer, J. de Troconiz, *Measurement of the Time-Dependence of* B^0-$\overline{B}^0$ *Oscillations Using the Low-*p_T *Electron-Muon Data*, CDF internal note 3791 (1996).

[48] M. Peters, University of California at Berkeley, Ph.D. dissertation, *A Measurement of B Meson Oscillations Using Inclusive Leptons in Proton-Antiproton Collisions at 1.8 TeV*, May 1997.

[49] D. Buskulie *et al.*, *Phys. Lett. B* **356**, 409 (1995).

[50] O. Long, University of Pennsylvania, Ph.D. dissertation, *A Proper Time Dependent Measurement of* Δm_d *Using Jet Charge and Soft Lepton Flavor Tagging* (unpublished), available at:
`http://www-cdf.fnal.gov/grads/thesis_complete.html`.

[51] F. Abe *et al.*, NIM Preprint 271A3871988;
Fermilab-Pub-94/024-E, submitted to Nucl. Instrum. Methods Phys. Res.

[52] M. Gronau, A. Nippe, J. Rosner, "Method for Flavor Tagging in Neutral *B* Meson Decays," *Phys. Rev. D* **47** (1988) 1993

[53] M. Gronau, J. Rosner, "Identification of Neutral *B* Mesons using Correlated Hadrons,", *Phys. Rev. D* **49** (254)1994

[54] D. Vucinic, Ph.D. dissertation, Massachusetts Institute of Technology, 1998 (unpublished), available at:
`http://www-cdf.fnal.gov/grads/thesis_complete.html`.

[55] F. James and M. Roos, " 'MINUIT' A SYSTEM FOR FUNCTION MINIMIZATION AND ANALYSIS OF THE PARAMETER ERRORS AND CORRELATIONS," *Comput. Phys. Commun.* **10** (343) (1975)

[56] T. Daniels, Ph.D. dissertation, Massachusetts Institute of Technology, 1998 (unpublished), available at:
`http://www-cdf.fnal.gov/grads/thesis_complete.html`.

[57] F. Abe *et al.*, *Phys. Rev. D* **40** 3791 (1989)

[58] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).

[59] Particle Data Group *Eur. Phys. Jour. C* **3**, 175 (19968)

[60] see F. Abe *et al.*, "The CDF II Detector," Fermilab-Pub-96/390-E (1996)

[61] F. Abe *et al.*, "The CDF II Detector," Fermilab-Pub-96/390-E (1996)

[62] P. Sphicas, **CDF** Internal Note 2655 (1994)

[63] P. Nason, S. Dawson, R. K. Ellis, *Nucl. Phys.* **B327**, 49 (1989)

[64] C. Peterson *et al.*, *Phys. Rev. D* **27**, 105 (1985).

[65] P. Avery *et al.*, *QQ: A Monte Carlo Generator*, CLEO Software Note CSN-212 (1985).

[66] P. Avery *et al.*, *QQ: A Monte Carlo Generator*, CLEO Software Note CSN-212 (1985).

[67] M. Shapiro *et al.*, *A User's Guide to QFL*, **CDF** Internal Note 18180 (1992).

[68] H. Bengtsson, T. Sjöstrand *PYTHIA: The LUND Monte Carlo for Hadronic Processes*, UCLA SSC Workshop, 228 (1986)

[69] W. Trishchuk and A. Warburton, **CDF** Internal Note 4423 (1997)