MEASUREMENT OF EXCLUSIVE B LIFETIMES IN THE MODES
$B^+ \rightarrow J/\psi\, K^+$, $B^+ \rightarrow J/\psi\, K^{*+}$,
$B^0 \rightarrow J/\psi\, K^*$, AND $B^0 \rightarrow J/\psi\, K^0_s$,
AND MEASUREMENT OF THE B^+/B^0 LIFETIME RATIO

by

Luis Roberto Flores-Castillo

B.S. Universidad Nacional Autónoma de México, 1994

M.Eng. Universidad Nacional Autónoma de México, 1997

M.S. University of Pittsburgh, 1999

Submitted to the Graduate Faculty of

the Department of Physics and Astronomy in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2004

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF PHYSICS AND ASTRONOMY

This dissertation was presented

by

Luis Roberto Flores-Castillo

It was defended on

July 13th 2004

and approved by

Dr. Joseph Francis Boudreau, Dept. of Physics and Astronomy

Dr. Anthony Duncan, Dept. of Physics and Astronomy

Dr. James Muller, Dept. of Physics and Astronomy

Dr. David W. Pratt, Dept. of Chemistry

Dr. Paul Shepard, Dept. of Physics and Astronomy

Dissertation Director: Dr. Joseph Francis Boudreau, Dept. of Physics and Astronomy

MEASUREMENT OF EXCLUSIVE B LIFETIMES IN THE MODES
$B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$,
$B^0 \to J/\psi\, K^*$, AND $B^0 \to J/\psi\, K^0_s$,
AND MEASUREMENT OF THE B^+/B^0 LIFETIME RATIO

Luis Roberto Flores-Castillo, PhD

University of Pittsburgh, 2004

This thesis presents the results from the measurement of the lifetimes of B^+ and B^0 mesons in four final states in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$ using $195\,\mathrm{pb}^{-1}$ of data collected with the CDF detector at the Fermilab Tevatron Collider. The proper decay length is estimated using the J/ψ vertex in all cases to reduce the systematics in the ratio. For the B^+ the mean life $c\tau$ is determined to be $499 \pm 12 \pm 6$ μm; for the B^0, $446 \pm 15 \pm 8$ μm. In both measurements the first error is statistical and the second is systematic. The ratio between $B^{\pm}$ and B^0 meson lifetimes is found to be $1.119 \pm 0.046(stat) \pm 0.014(syst)$

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ix

PREFACE

In my time at the University of Pittsburgh and at Fermilab, I have had the opportunity of learning from and working with many people, both inside and outside of the realm of physics. I deeply thank all of them, regardless of their inclusion in this small list, necessarily incomplete.

I owe a huge debt of gratitude to Joe, my advisor. Besides your help and guidance, I particularly thank your contagious enthusiasm, and your teaching me, in passing, that working in High Energy Physics does allow one to fully enjoy much more than lab life.

I want to thank especially my wife, Via. Thank you for your love and support, for your being always there for me, even when my mind was so often lost somewhere among CAF jobs and badly behaved B_0 mesons; thank you for making our life together such a beautiful journey.

Thank you, Juan Pablo, for all your invaluable help along the way and, first and foremost, for being such a fantastic friend. Un abrazo, hermano.

Thank you Dad for so many things... as time passes, I keep finding out more and more ways in which we are alike. Thank you for your example and your trust. Thank you Mom for your love and your warm laughter, for being always so positive, and for showing me that life can sometimes be told what to do. Thank you Sandra, Mauricio, Eduardo for your friendship, for making me feel at home even in the distance. I am also deeply thankful to my grandparents Felipa, Josefina and Domingo. I owe them all a lot of what I am. This thesis is, in good part, a result of their work and love.

The results reported in this document would not have been obtained without the dedication and work of hundreds of people from CDF and Fermilab. I want to express my gratitude to the B and the J/ψ groups, and especially to their conveners, Michael Schmidth, Mar-

jorie Shapiro, Christoph Paus, Stefano Giagu, Ting Miao, Mary Bishai, Tom LeCompte and Matthew Herndon, for providing direction, as well as a lot of work, to our joint effort. Also, I want to thank the people from the Accelerator Division, the Computing Division and the Technical Division. Our work in CDF would not be possible without all the infrastructure they help providing. I am very grateful to many people from the CDF trailers. Thank you Yuri for your help and company, and for being such a good example on so many aspects of life. Thank you Ting for your continual support and good advice. Thank you Tom for your stories and drive. Thank you Vaia, Mario and Jack for those shared nights of hard work. Thank you Robyn and Konstantin for being always willing to lend a hand, even when it meant more work for you.

I want to thank many people from the Department of Physics and Astronomy at Pitt. To name but a few, professors Paul Shepard, Carlo Rovelli, Martin Vincent, Dan Boyanovski and Jeremy Levy, from which I so much enjoyed learning what physics is about. Thank you Suresh, for your company and friendship during so many good and bad times during the years at Pitt.

Thank you Francisco, for helping me so much at the beginning of this enterprise, even when you were not sure if it was the best for me.

1.0 INTRODUCTION

Particle physics has covered a lot of ground since the second half of the twentieth century. What we now call the *Standard Model* (SM) of particle physics [1, 2, 3] has been able to accurately describe an immense majority of experimental data, and it does so starting from a small set of assumptions about the nature of fundamental particles and their interactions.

In the SM, all matter is assumed to be formed by spin-1/2 particles organized in three "generations", each containing two quarks and two leptons. The first generation is formed by the *up* (denoted u) and *down* (d) quarks, the electron (e) (the lightest charged lepton) and the electron-neutrino (ν_e) (a neutral lepton). The second generation consists of the *charm* (c) and *strange* (s) quarks, the muon (μ) and a corresponding neutrino (ν_μ). Finally, the heaviest quarks (and, consequently, the ones most recently discovered) are called *top* (t) and *bottom* (b); together with the tau lepton (τ) and the tau neutrino (μ_τ), they form the third generation. The interactions among these particles is described as the exchange of gauge bosons with integer spins: the vector bosons $W^\pm$, Z^0 and the photon mediate the electroweak interaction, and eight gluons mediate the strong interaction. Gravity (the only other fundamental force we know about, and the weakest of them all) is not described in the SM framework.

Due to the nature of the strong interaction, quarks can only be found in bound states with other quarks, forming particles called *hadrons*. In the current terminology, hadrons formed by the combination of three quarks are called *baryons*, and those formed by one quark and one antiquark are called *mesons*. There is currently some evidence for the existence of bound states of five quarks [4], but the issue of their quark content has not been completely settled at this point.

Hadrons are named according to their quark content. Mesons formed by quarks of different flavor (like $b\bar{u}$ or $\bar{c}d$) are given a name based on their heaviest quark. B mesons are those that contain a b quark, D mesons contain a c, and those containing an s quark are called K mesons (or $Kaons$). Special names are given to same-flavor combinations; of interest for us is the combination $c\bar{c}$, called a J/ψ meson.

Besides the particle content mentioned, the SM can be characterized in terms of the symmetries of its Lagrangian (which themselves determine the particle content; each particle corresponds to an irreducible representation of the symmetry group). The SM Lagrangian is invariant under transformations of the gauge group $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ component corresponds to the color charge (mediated by and coupled to gluons and quarks, respectively). $SU(2)$ corresponds to the weak isospin and $U(1)$ to hypercharge; these two mix to give rise to the weak and electromagnetic interactions. Without further refinements, all particles would need to be massless, since mass terms would break gauge invariance. In order to allow mass terms for quarks, leptons, and the electroweak massive gauge bosons, an extra spin zero neutral gauge field is introduced, which allows for a spontaneous symmetry breaking mechanism (the Higgs mechanism), and implies the existence of a neutral scalar particle called the Higgs boson, which remains undiscovered.

Other relevant transformations, that give their names to corresponding symmetries, are charge conjugation, spatial inversion (or $parity$), and time reversal, denoted as C, P and T respectively. Until 1957, it was believed that particle interactions were invariant under parity transformations (which is to say, that no physics results could show a difference between left-handedness and right-handedness). It came as a surprise that such is not the case (as first shown by C.S.Wu and collaborators [5]) [1]. Still, it was possible that interactions were invariant under the combined action of C and P (i.e., under CP transformations). However, in 1964, Christenson, Cronin, Fitch and Turlay observed that K meson decays do violate CP [6]. This violation was described, in the SM, via a complex phase in the quark mixing matrix (which will be described in section 1.2.1). Such description, motivated by findings in Kaon decays, provides testable predictions for the behavior of CP violation in the decay of

---

[1] When a nucleus Co^{60} was placed in a magnetic field, electrons from the beta decay were preferentially emitted in the direction opposite that of the angular momentum of the nucleus.

B mesons [2]. In the last few years, it has been possible to accurately measure this and other properties of particles containing b and c quarks. This influx of experimental information has motivated the development of a set of techniques for their analysis, sometimes encompassed under the name *Heavy Quark Theory* (see, for example, [7]), which will also be more tightly tested as more data becomes available. Accurate measurements of the B^0 and B^+ lifetimes, for example, will eventually be able to determine the validity of some of these techniques.

In this thesis work, we present a measurement of the lifetimes of the B^+ and B^0 mesons and of the corresponding lifetime ratio. The measurement was made using exclusive decays to states containing a J/ψ in the CDF detector at Fermilab's Tevatron (the highest energy proton-antiproton collider available today). The measurements of lifetimes for these heavy mesons, and of the ratio, are important on their own right, but also because they establish experimental procedures necessary in the study of CP violation and in the validation of the *Heavy Quark Expansion.*

To help the reader situate the measurement within the context of B physics, this chapter presents a summary of the measurement, a comparison with other results (from CDF and other experiments), and a brief description of CP violation and the Heavy Quark Expansion technique. The CDF detector is described in chapter 2. The data sample and the procedures used to reconstruct our final states are described in chapter 3. The analysis and fitting methods are detailed in chapter 4; chapter 5 describes the systematic uncertainties, and chapter 6 summarizes the results.

1.1   THE MEASUREMENT AND ITS EXPERIMENTAL CONTEXT

The Fermilab accelerator complex underwent a major upgrade after the period of data taking that spanned from 1993 to 1996, and it restarted physics data taking on April of 2001. To distinguish the two periods (and the corresponding results), the former is referred to as Run I, and the current one as Run II. The work reported on this thesis is based on Run II data available up to November of 2003.

---

[2]In particular, one of the channels reconstructed as part of this thesis work ($B^0 \to J/\psi\, K^0_s$) is especially useful for the extraction of the CP violation parameter $\sin 2\beta$.

We reconstruct and measure the lifetimes of two b-flavored mesons: B^+ and B^0. Each meson is reconstructed in two decay channels. The B^+ is reconstructed in the modes: $B^+ \to J/\psi\, K^+$ (with the J/ψ decaying into $\mu^+\mu^-$) and $B^+ \to J/\psi\, K^{*+}$ ($J/\psi \to \mu^+\mu^-$, $K^{*+} \to K_s^0 \pi^+$, $K_s^0 \to \pi^+\pi^-$). Over 2100 B^+ events with precise vertex information are observed in the channel $B^+ \to J/\psi\, K^+$, and about 200 more are observed in the mode $B^+ \to J/\psi\, K^{*+}$ (to be compared with 900 and 80 observed events respectively in Run I [8]). The statistical error of the lifetime measurements in Run II is now smaller than that of Run I.

The B^0 is also reconstructed in two decay modes: $B^0 \to J/\psi\, K^*$ ($J/\psi \to \mu^+\mu^-$, $K^{*0} \to K\pi$) and $B^0 \to J/\psi\, K_s^0$ ($J/\psi \to \mu^+\mu^-$, $K_s^0 \to \pi^+\pi^-$). About 950 signal events, with silicon information, are found in the first channel and other 600 events in the second channel (to be compared with 430 and 135 events respectively in Run I). Figure 1 illustrates the topologies of these decays. It is convenient at this point to comment on the order of magnitude of the quantities measured, and the units used throughout this study. B mesons have a mean lifetime τ of about 1.5×10^{-12} seconds, or 1.5 ps (1 ps, or picosecond, is one thousand of a billionth of a second, or 1.0×10^{-12}s). When reconstructing the decays, we usually do not make a measurement of time, but of distance. That being the case, instead of τ, we will most of the time refer to the quantity $c\tau$ (i.e., the speed of light c times the mean lifetime), which is approximately half a millimeter or, in a more convenient unit, 500 microns (μm). Detector resolutions, systematic uncertainties, some background parameters, etc. are also more conveniently expressed in microns.

As will be described on chapter 6, the statistical uncertainty on our measurement is of about 0.05 picoseconds for the lifetimes and 0.05 for the ratio. As can be seen in table 1, currently the BaBar and Belle experiments (the "B Factories") [9, 10] have smaller errors than the ones reported here; however, our errors are dominated by statistical uncertainties which will decrease greatly as CDF accumulates the full data sample of Run II, expected to be ten times as large as the sample used here. Also, understanding the systematic uncertainties involved in these measurements is important for the physics program at CDF because other states (in particular, the B_s meson and the Λ_b baryon, which are not produced at the B Factories) will be reconstructed with very similar procedures.
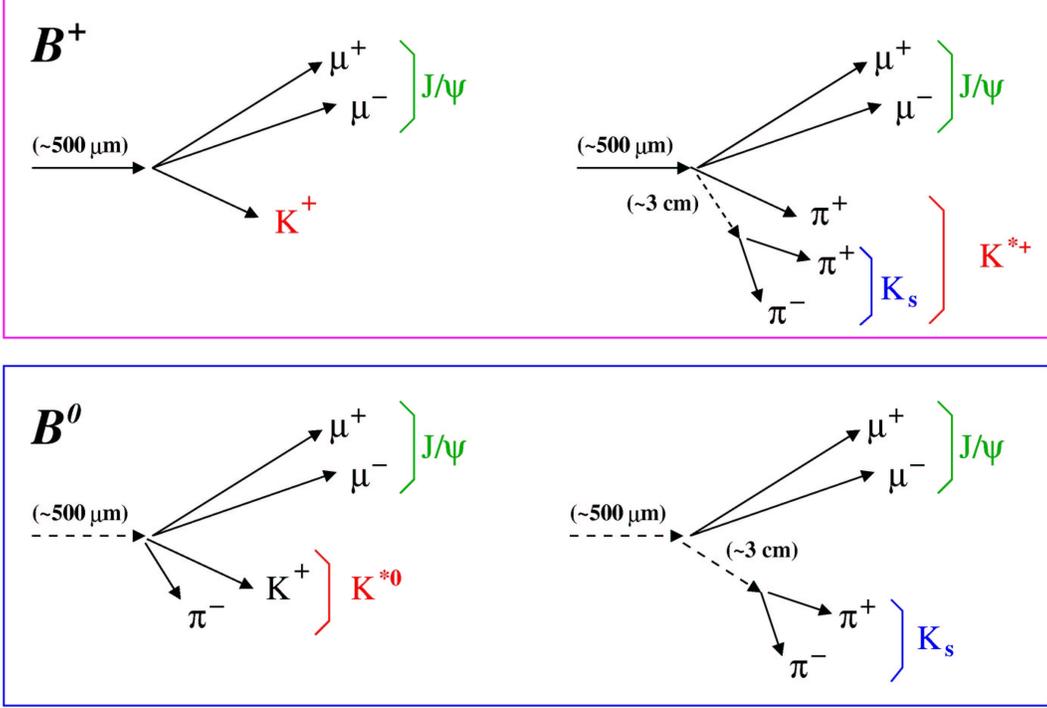
Figure 1: *Topologies of B meson decay channels considered.*

Table 1: Results from CDF Run I, BaBar and Belle. The first errors are statistical, the second are systematic.

CDF Run I	τ_{B^+}	$=$	1.68 $\pm$0.07 $\pm$0.02 ps,
	τ_{B^0}	$=$	1.58 $\pm$0.09 $\pm$0.02 ps,
	τ_{B^+}/τ_{B^0}	$=$	1.06 $\pm$0.07 $\pm$0.02
BaBar	τ_{B^+}	$=$	1.673 $\pm$0.032 $\pm$0.023 ps,
	τ_{B^0}	$=$	1.546 $\pm$0.032 $\pm$0.022 ps,
	τ_{B^+}/τ_{B^0}	$=$	1.082 $\pm$0.026 $\pm$0.012
Belle	τ_{B^-}	$=$	1.695 $\pm$0.026 $\pm$0.015 ps,
	$\tau_{\bar{B}^0}$	$=$	1.554 $\pm$0.030 $\pm$0.019 ps,
	$\tau_{B^-}/\tau_{\bar{B}^0}$	$=$	1.091 $\pm$0.023 $\pm$0.014

1.2 THEORETICAL CONTEXT

The decay of most mesons, including those considered in this thesis, is the result of the charged weak interaction, in which a W boson is exchanged, as in nuclear beta decay. The charged weak interaction is also (in the SM) considered to be responsible for other phenomena besides weak decays; among them, particle-antiparticle oscillations and the violation of CP. The couplings of quarks to the weak interaction is governed by coupling constants whose relative amplitudes can be written as a 3×3 matrix. Measurements of lifetimes, mixings, and CP violation, are, generally speaking, measurements of this matrix and therefore of the charged weak interaction with quarks.

The measurements reported in this thesis correspond to properties of B mesons (i.e., two-quark states that include a b quark), as measured in exclusive modes including a J/ψ particle (with quark content $c\bar{c}$). Besides the characterization of the mesons considered, these measurements are important because they are necessary components in the study of CP violation in the b sector, and because they will ultimately allow us to test the adequacy of a theoretical approach to the study of heavy flavor physics (the Heavy Quark Expansion). Both issues are briefly described in what follows.

1.2.1 CP violation

In the SM, the couplings of u-type quarks (u, c, t) to d-type quarks (d, s, b) are encoded in the so called *Cabbibo-Kobayashi-Maskawa* (or *CKM*, for short) matrix, which is usually written as:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.1}$$

so that entries are labeled by the quark flavors. The unitarity of V implies constraints that give information on its unmeasured (or poorly measured) entries. For instance, $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$ (and similarly for all other rows and columns). Also, $|V_{cb}|$ and $|V_{ub}|$ are known to be small; as a result, if only three generations exist, $|V_{tb}|$ should be close to 1.
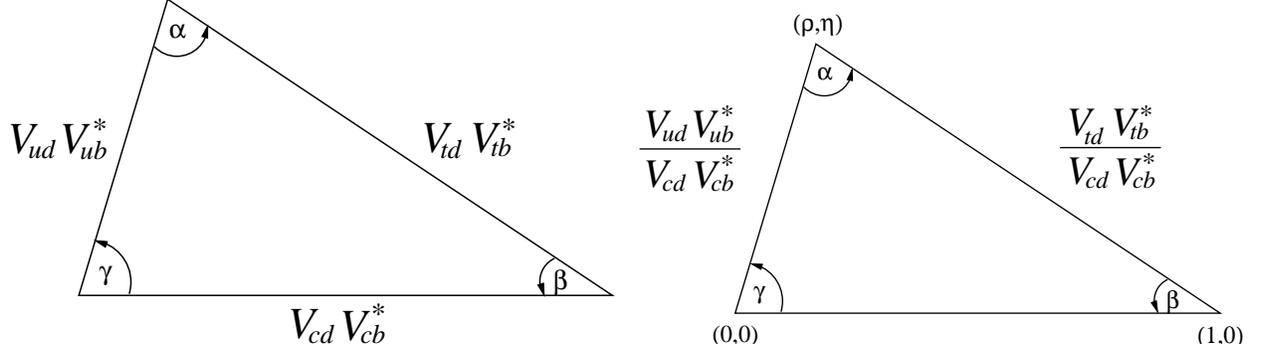
Figure 2: *The unitarity triangle. The version on the left directly expresses Eq. (1.2). The re-scaled version shows the definition of (ρ, η).*

Other constraints arise from the orthogonality of columns (or rows); in particular,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.2}$$

which can be depicted as a triangle on the complex plane, which is hence called the "unitarity triangle," shown in figure 2. The lengths of the sides are simply $|V_{ud}V_{ub}^*|$, etc., and the angles are

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \qquad \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \qquad \gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{1.3}$$

This representation provides a simple summary of the CKM mechanism. Separate measurements of lengths, through decay and mixing rates, and angles, through CP asymmetries, should fit together, unless there are non-CKM contributions to flavor or CP violation.

The Particle Data Book [11] advocates for a "standard" parametrization of V that utilizes angles θ_{ij} and a phase δ_{13}:

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

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for the "generation" labels $i, j = 1, 2, 3$.

A convenient parameterization, due to Wolfenstein [12], stems from the observation that, in the measured matrix, diagonal elements are close to 1, and others are progressively smaller

as they get farther away from the diagonal. This can be formalized by defining λ, A, ρ, and η via

$$\lambda \equiv s_{12}\,, \qquad A \equiv s_{23}/\lambda^2\,, \qquad \rho + i\eta \equiv s_{13}e^{i\delta_{13}}/A\lambda^3\,. \tag{1.5}$$

We learn from experimental data that $\lambda \approx 0.22$, $A \approx 0.8$, and $\sqrt{\rho^2 + \eta^2} \approx 0.4$, so it is useful to expand V in powers of λ:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{1.6}$$

The unitarity triangle described by Eq. (1.2) is specially well suited for measurement because its three sides are of the same order ($A\lambda^3$). This also happens for the triangle corresponding to the first and third rows, but the fast decay of the top quark does not allow the formation of the mesons that would be necessary to measure the angles. The remaining triangles are long and thin, with sides $(\lambda, \lambda, A\lambda^5)$ (e.g., for the Kaon) and $(\lambda^2, \lambda^2, A\lambda^4)$ (e.g., for the B_s meson).

Special attention has been given to the decay $B^0 \to J/\psi\, K_s^0$ due to the fact that this channel allows a relatively clean extraction of a parameter from the unitarity triangle; namely, $\sin 2\beta$. This measurement requires the reconstruction of the signal, extraction of lifetime information, and the identification of the initial flavor of the B meson as either B^0 or $\bar{B}^0$ (this identification is called "flavor tagging"). In this context, the present thesis work is relevant because it includes the reconstruction of this exclusive channel, as well as the extraction of the corresponding lifetime. Also, reconstruction of the channel $B^+ \to J/\psi\, K^+$, which we also perform in this study, is a necessary step for the calibration of the algorithms used for flavor tagging.

Besides its use as a component in the measurement of $\sin 2\beta$, the lifetime of B mesons has an important role in the determination of the CMK matrix element V_{cb} (see, for example, [13]). This is because V_{cb} is closely related to the weak decay of the b quark into a c. If there were no other interactions within a meson, (i.e., if the other quark in the meson only played the role of an "spectator"), all mesons with a b quark would have the same lifetime as the b quark; this simple picture is usually called the "spectator model". This is, of course,

not the case. Interactions between the b and the remaining quark within the meson are expected to change the lifetime of the meson in a species-dependent way by a few percent. The framework for calculating this is the Heavy Quark Expansion. Precise measurements of the ratio of lifetimes of particular species of B mesons can help establish this technique.

1.2.2   The Heavy Quark Expansion

The availability of experimental data on b decays (both present and expected in the near future), from the B Factories and from hadron colliders, has strongly motivated the development of systematic theoretical approaches for the study of b-flavored hadrons; one of them, the *Heavy Quark Expansion* (HQE) [7], allows the calculation of expected B-hadron lifetimes within the SM.

As mentioned before, different species of heavy hadrons would all have the same lifetime as their heaviest quark Q if there were no interactions between Q and the other quarks. For both theoretical and experimental reasons, we need to refine the description. On the theoretical side, we do know that there are other interactions. The spectator model prediction has to be corrected due to the effects of the strong force (described by Quantum Chromodynamics, or QCD) and weak interactions among the heavy quark and the others (which are not "spectators" anymore), and even because of the quantum mechanical interference among different possible ways for the hadron to decay. From the experimental viewpoint, the spectator model prediction fails clearly in the case of the c hadrons since, for them, the ratio between the shortest and the longest measured lifetimes is roughly 16.

This section describes briefly the application of the HQE to the calculation of B lifetime ratios. Roughly speaking, the basic idea is to isolate the physics in which we are interested in two steps: first, by factoring out the non-perturbative behavior of the strong interaction and, second, by expressing the quantity of interest as a power series in the (inverse of the) mass of the heavy quark. For heavy hadrons, the leading terms of such expansion should provide a good agreement with experiment.

Going into some more detail, these two steps are as follows. First, an operator product expansion (OPE) is used to separate the non-perturbative QCD contributions and the short-
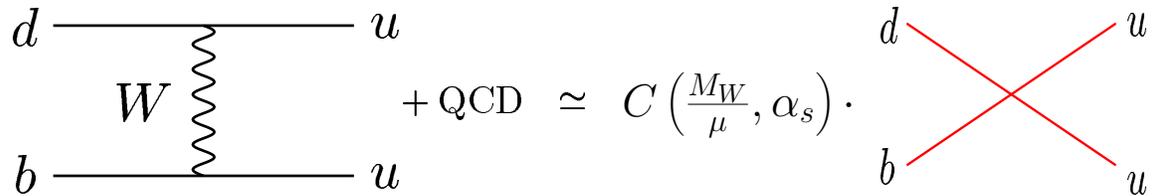
Figure 3: *Illustration of the OPE for weak decays (From [7]).*

distance contributions coming from b quark dynamics. Second, the fact that $m_b \gg \Lambda_{QCD}$ (where m_b is the mass of the b quark and Λ_{QCD} is a constant that indicates the onset of non-perturbative QCD) is exploited by expressing the total width as an expansion in terms of Λ_{QCD}/m_b. In fact, B hadrons are the best candidates for application of the HQE, since the b quark is heavy enough to confidently use the expansion in Λ_{QCD}/m_b, yet it is not heavy enough to make the leading corrections negligible. In contrast, the HQE can be applied to c hadrons only with some reservations; in some cases it gives a reasonable approximation, but m_c is often too low for the $1/m_c$ expansion to be reliable [14]. The following gives a brief description of the method and its results; for a more comprehensive explanation, we refer the reader to [7, 15, 16].

The basic idea of the OPE is that interactions at higher (energy) scales can be modeled as local operators at lower scales. This is illustrated in figure 3; the exchange of the heavy W boson can be approximated by a point-like four-quark interaction, multiplied by a *Wilson coefficient C*, which absorbs all effects of QCD interactions above some factorization scale μ (short-distance effects), while the low-energy contributions below μ (long distances) are collected in the matrix elements of local operators. This results in an effective Hamiltonian describing the flavour-changing weak interactions up to corrections m_b^2/M_W^2, of the form [15]:

$$H = \frac{G_F}{\sqrt{2}} V_{cb}^* \sum_{\substack{d'=d,s \\ u'=u,c}} V_{u'd'} \left[C_1(\mu_1) Q_1^{u'd'}(\mu_1) + C_2(\mu_1) Q_2^{u'd'}(\mu_1) \right] + \text{h.c.} \qquad (1.7)$$

where G_F is the Fermi constant and the V_{ij}'s are elements of the CKM matrix. The Wilson coefficients $C_i(\mu_1)$ contain the short-distance physics associated with scales above the renormalization scale μ_1. The weak interaction is encoded in the four-quark operators $Q_1^{u'd'}$ and

$Q_2^{u'd'}$. Penguin operators and doubly Cabibbo-suppressed terms have a negligible effect on the B^+–B_d^0 lifetime difference, and have been omitted in 1.7.

The total decay rate $\Gamma(H_b)$ can then be related to H through the optical theorem:

$$\Gamma(H_b) \;\;=\;\; \frac{1}{2m_{H_b}}\langle H_b|\mathcal{T}|H_b\rangle, \tag{1.8}$$

where the transition operator $\mathcal{T}$ is defined by

$$\mathcal{T} \;\;=\;\; \operatorname{Im} i\int d^4x\, T[H(x)\,H(0)]. \tag{1.9}$$

After the OPE, in the case of hadrons composed of one heavy quark (either b or c)[3] and any of the light quarks (u, d, s), it is possible to use the fact that the mass m_Q of the heavy quark is large compared to the QCD scale parameter Λ_{QCD}, and expand $\mathcal{T}$ in terms of Λ_{QCD}/m_Q; this expansion is what motivates the name of the method (HQE), and has the form [7]:

$$\mathcal{T} = \Gamma_b\,\bar{b}b + \frac{z_G}{m_b^2}\,\bar{b}g\sigma\cdot Gb + \sum \frac{z_{qi}}{m_b^3}\,\bar{b}\Gamma_i q\;\bar{q}\,\Gamma_i b + \dots \tag{1.10}$$

where only the first few operators, of dimension three ($\bar{b}b$), five ($\bar{b}g\sigma\cdot Gb$) and six ($\bar{b}\Gamma_i q\;\bar{q}\,\Gamma_i b$), are shown. As before, the operators contain the soft, nonperturbative physics, while their Wilson coefficients (Γ_b, z_k) absorb the hard contributions, which are calculable in perturbation theory. G is the gluon field tensor, q stands for light quarks u, d, s, and Γ_i denotes spin and color structures of the four-quark operators. The Feynman diagrams for the three terms in (1.10) are shown in figure 4.

Using a similar expansion technique, from a framework known as *Heavy Quark Effective Theory* (HQET) [16, 7], the matrix element $\langle H_b|\bar{b}b|H_b\rangle$ of the leading operator $\bar{b}b$ between heavy hadron states H_b can be expressed as

$$\langle\bar{b}b\rangle = 1 + \frac{1}{2m_b^2}\,\langle\bar{b}(iD)^2 b\rangle + \frac{1}{4m_b^2}\,\langle\bar{b}g\sigma Gb\rangle \tag{1.11}$$

where $\langle\dots\rangle \equiv \langle H_b|\dots|H_b\rangle/(2m_{H_b})$, and D is the QCD covariant derivative. Together, equations (1.8), (1.10) and (1.11) imply that, to leading order in m_b^{-1}, $\Gamma_H = \Gamma_b$; i.e., that the total decay rate of all b-flavored hadrons is equal to the rate of free b-quark decay. This is

---

[3] Although the top quark is heavier than b and c, its decay time is too short to allow the formation of bound states.
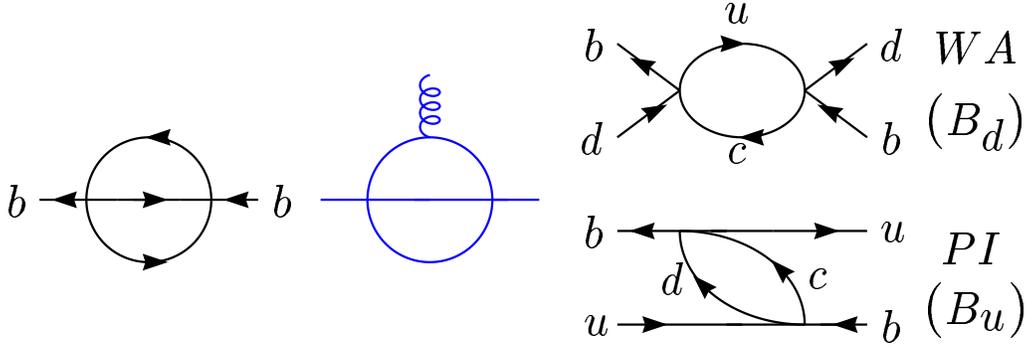
Figure 4: *Heavy quark expansion for the total decay rate of b-hadrons. For the transition operator, only the absorptive part of the diagrams is used. The third term in 1.10 gives rise to the diagrams on the right, usually referred to as Weak Annihilation (WA) and Pauli Interference (PI).*

the "spectator model" estimate of the decay rate; it is called this way because it corresponds to a hadronic decay in which the light quarks do not intervene at all in the process, which is then determined only by the b quark decay. In b hadrons, the relative contribution from higher-order terms in the $1/m_b$ expansion is at the level of a few percent, so this is already a good estimate. In this approximation, all b hadrons are predicted to have the same lifetime, regardless of the flavor of the light quarks. However, experiments have been precise enough to clearly show that this is not the case. Measured ratios of b hadrons' lifetimes are closer to 1 than those of charm hadrons. This is expected, since the spectator model should become a better approximation as the mass of the heavy quark increases. However, the longest and the shortest lifetimes for c hadrons differ by a factor of about 16.

Corrections to the spectator model start at order $(1/m_b)^2$. There is no linear term because no gauge-invariant operator of dimension four exists that could appear in the HQE. The first correction term in (1.11) depends on the expectation value of the momentum squared $\langle \mathbf{p}^2 \rangle$ of the heavy quark inside the hadron. This term corresponds to a correction factor $1 - \langle \vec{p}^2 \rangle/(2m_b^2) = 1 - \langle \vec{v}_b^2 \rangle/2$, and can be interpreted as the reduction of the free decay rate from time dilation due to the recoil motion of the heavy quark.

Another correction of order $1/m_b^2$ comes from interaction of the light hadronic cloud with the heavy-quark spin. This interaction (represented by the term $\langle \bar{b} g\sigma G b \rangle$) enters twice in the calculation of the decay rate; explicitly with coefficient z_G in (1.10) and via the expansion of $\langle \bar{b}b \rangle$ shown in (1.11). Its value

$$\langle \bar{b} g\sigma G b \rangle = \begin{cases} \frac{3}{2}(m_{B^*}^2 - m_B^2) & H_b = B \\ 0 & H_b = \Lambda_b \end{cases} \tag{1.12}$$

is zero for the Λ_b baryon because the light degrees of freedom are in a state of zero total angular momentum. As a result, there is a difference among B mesons and Λ_b at this order in the expansion.

In the next order, $1/m_b^3$, contributions appear where the light quark participates directly in the weak interactions. For b-mesons, this contributions have been historically depicted as the effect of *weak annihilation* (WA) of the b-quark with the valence $\bar{d}$-quark (for $\bar{B}_d$) and as the effect of *Pauli Interference* (PI) (for $\bar{B}_u$). The latter occurs because in the nonleptonic decay of a $\bar{B}_u$, $b(\bar{u}) \rightarrow c\bar{u}d(\bar{u})$, two identical $\bar{u}$-quarks are present in the final state, hence allowing for interference among the two ways of choosing them. These effects (PI and WA) are especially important in the case of b mesons since, although the third term in (1.10) is suppressed by an extra power of $1/m_b$, it is also enhanced by an extra factor $16\pi^2$. This factor comes from the fact that the diagrams determining the third term contain a two-body phase space, while the second-order terms involve a three-body phase space (see figure 4).

The main potential caveat in the application of the OPE to decays of b hadrons is that it relies on the assumption of *quark-hadron duality* at the energies considered; i.e., the assumption that sufficiently many exclusive hadronic channels contribute to the inclusive rate, so that the details of the low-energy resonance structure do not affect the total rates of the inclusive processes. Theoretical understanding of the onset of quark-hadron duality is still not detailed enough for a quantitative evaluation of possible departure from duality in b and c decays. So far, the adequacy of this assumption can only be assessed through comparison with experiment.

A second, already mentioned, potential problem in the application of the HQE to heavy hadron decays is related to its use for c decays; the mass of the c quark may not be large enough to justify discarding higher terms in Λ_{QCD}/m_c.

Yet another source of uncertainty arises from poor knowledge of matrix elements of certain quark operators. They can, however, be estimated within theoretical models (in which case the result is model dependent), and sometimes extracted from experimental data.

Table 2 shows a comparison between predictions based on the HQE and experimental results. Despite the relatively low value of the mass of the c quark, the HQE does achieve a good agreement with data in the case of charm hadrons. On the other hand, as expected, the ratios for b hadrons (both expected and measured) are closer to unity, and at this time the experimental values are compatible with the predictions, given the size of the uncertainties in both, for $\tau(B^+)/\tau(B^0)$ and $\tau(B_s)/\tau(B^0)$. The case of $\tau(\Lambda_B)/\tau(B^0)$ shows a possible discrepancy, but it is still marginal.

Table 2: Lifetime ratios in the charm and bottom sectors. HQE predictions vs. measurements. Adapted from [17] and [7].

Ratio	HQE prediction	Measured value
$\tau(D^+)/\tau(D^0)$	~ 2	2.55 ± 0.034
$\tau(D_s^+)/\tau(D^0)$	0.9 - 1.3	1.125 ± 0.042 PDG '98
		1.211 ± 0.017 E791, CLEO, FOCUS
$\tau(\Lambda_c^+)/\tau(D^0)$	~ 0.5	0.489 ± 0.008
$\tau(\Xi_c^+)/\tau(\Lambda_c^+)$	~ 1.3	1.75 ± 0.36 PDG '98
$\tau(\Xi_c^+)/\tau(\Xi_c^0)$	~ 2.8	3.57 ± 0.91 PDG '98
$\tau(\Xi_c^+)/\tau(\Omega_c)$	~ 4	3.9 ± 1.7 PDG '98
$\tau(B^+)/\tau(B^0)$	1 - 1.1	1.083 ± 0.017
$\tau(B_s)/\tau(B^0)$	0.99 - 1.1	0.947 ± 0.038
$\tau(\Lambda_B)/\tau(B^0)$	0.9 - 1.0	0.797 ± 0.053

2.0 THE CDF EXPERIMENT AT THE TEVATRON COLLIDER

The *Collider Detector at Fermilab*, or CDF, is a general purpose experiment for the study of $p\bar{p}$ collisions at the Tevatron, the highest-energy particle collider currently operational in the world. Using the Tevatron accelerator, CDF discovered top-quark events in 1995. The Tevatron also produces all flavors of B hadrons, which has allowed CDF to carry out a broad B physics program in Run I.

To exploit this unique tool fully, and to meet the goals of the high energy physics research program through the 1990's and into the twenty-first century, a phased upgrade of the accelerator complexes and the detector has taken place. The upgrade is expected to give an opportunity not only to discover new physics but also to perform precision measurements. An extension of the B physics capabilities of the the detector has been part of the reason for the upgrade; tens of thousands of fully reconstructed B mesons, measured precisely with the CDF detector, will ultimately make the Tevatron an excellent place to study the properties of B mesons. In this chapter, various complexes and functions of the Tevatron accelerator and CDF detector are briefly described. For a complete review see [18] and [19].

2.1 THE PARTICLE ACCELERATOR

The Tevatron is located in the high energy physics laboratory *Fermilab*, (short name for the Fermi National Accelerator Laboratory) in Batavia, approximately 40 miles west of Chicago in the State of Illinois. In the Tevatron, it is possible to study the hard scattering events resulting from collisions between protons and antiprotons of nearly 1 TeV.

The energy available in the center of mass on these collisions is 1.96 TeV. The Tevatron ring has a 6.3 km perimeter and is located inside an underground tunnel. An aerial shot of Fermilab is shown in figure 5.

2.1.1 Accelerator chain

In order to reach the final colliding energy of 980 GeV, five systems are used: Cockcroft-Walton accelerator, Linac, Booster accelerator, Main Injector and Tevatron. The schematic view of the accelerator complex at Fermilab is illustrated in figure 6.

The accelerating process starts with electrical discharges in a hydrogen gas bottle, where H ions are produced. These H's are pulled out from the gas bottle by a Cockcroft-Walton accelerator, which is a 750KV DC voltage source, and H's are accelerated to 750 keV.

In a second step, these ions enter a linear accelerator (Linac) 145 m long, where they reach the energy of 400 MeV. The Linac has two parts: a 116 MeV drift-tube Linac operating at 201.25 MHz and a 400 MeV side-coupled cavity Linac operating at 805 MHz [20]. The H's are accelerated through radiofrequency electromagnetic waves generated by a series of cavities in the Linac; as a result, the continuous beams are separated into several bunches.

At the end of the Linac, the electrons from the H ions are stripped off by a thin carbon foil, and the resulting protons are passed into the Booster ring, which is a circular synchrotron ring with a 75 m radius. The protons are circulated until they acquire an energy of 8 GeV before being collected in bunches, each of 6×10^{10} particles, to be used in the Tevatron. The acceleration in the Booster is accomplished by a series of electromagnetic kicks applied by RF cavities: about 500 kV per turn.

The next step is further acceleration in the Main Injector, where protons are accelerated from 8 GeV to 150 GeV for injection into the Tevatron or to 120 GeV for $\bar{p}$ production. The Main Injector is a proton synchrotron ring with a circumference of 3319.4 m which operates at 53 MHz. For the Tevatron Run I operation, the 8 GeV proton bunches were injected into the Main Ring synchrotron, where they were accelerated to 150 GeV. In the Run II phase, the existing Main Ring has been replaced by a new accelerator, the Main Injector, for the purpose of reducing the inefficiency of the antiproton production. The 120 GeV proton
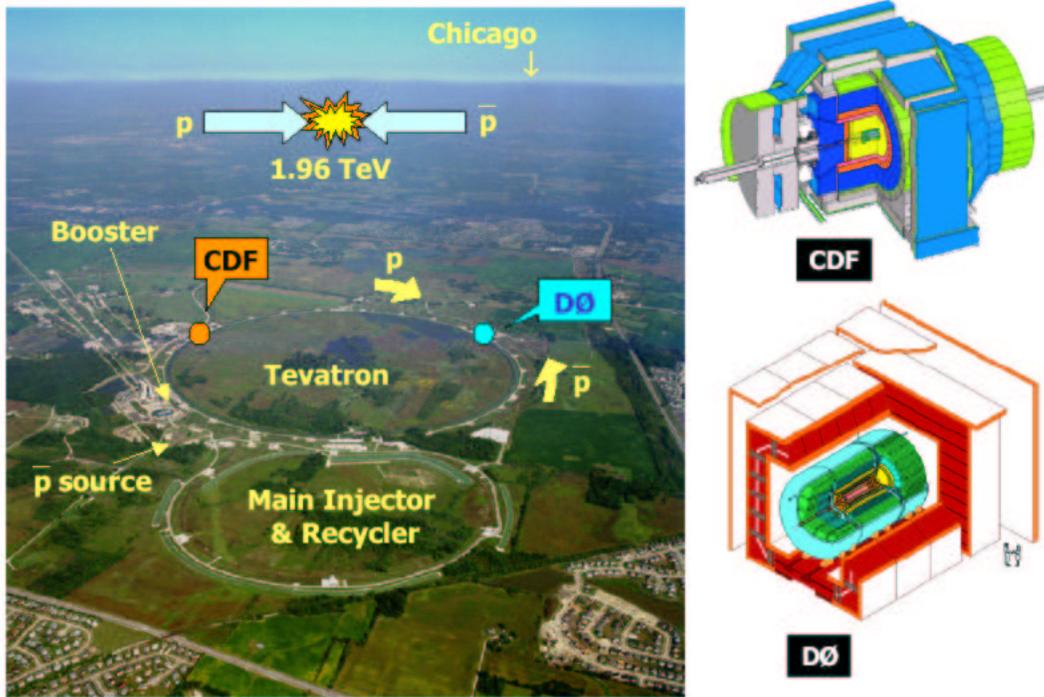
Figure 5: *Aerial shot of Fermilab. The CDF detector is located at one of the collision points, called B0, on the Tevatron Ring.*

bunches are finally injected into the evacuated beam pipe of the Tevatron ring and brought to an energy of 980 GeV.

2.1.2 Antiprotons' production

Antiprotons are created at a target station. On its way towards the Tevatron, a fraction of the proton beam is extracted to bombard a nickel target. The 120 GeV protons extracted from the Main Injector hit the target and produce antiprotons over large spread angles and energies in the forward direction. On average, 20 antiprotons per million protons are produced, with a mean kinetic energy of 8 GeV. These antiprotons then undergo a process called stochastic cooling to reduce random motion [21]. A cylindrical lithium lens is used to focus the antiprotons into a parallel beam and a pulsed magnet separates them from other particle species. The resulting antiproton beam is then directed to the Debuncher, a rounded triangular-shaped synchrotron with a mean radius of 90 m. Its primary purpose is to reduce the momentum spread of the antiprotons by rotating the bunches. There are also beam cooling systems that act to reduce the oscillations in the plane perpendicular to the orbit (transverse plane) apart from reducing the momentum spread of the antiprotons. Then, the antiprotons are transferred to the Accumulator, which is the storage ring for the antiprotons. Antiprotons are stored there at 8 GeV and cooled until needed. After a period of 10 to 20 hours, when the stack is large enough, bunches of antiprotons are transferred into the Main Injector and accelerated to 150 GeV, and finally into the Tevatron ring and accelerated to 980 GeV. The total number of antiprotons in the collider is determined by the product of the antiproton production rate, the typical store duration, and the transmission efficiency from Accumulator to storage in the Tevatron.

As part of the Tevatron Run II upgrade, a new Antiproton Recycler Ring has been proposed. The role of the Recycler Ring would be to provide more antiprotons for the Tevatron, by recycling 2/3 of the 75% of the antiprotons left at the end of a store. This proportionally doubles the luminosity. If the Recycler Ring becomes available, the antiprotons in the Tevatron at the end of the store will not be dumped but collected into the Recycler Ring, and then sent to the Main Injector and to the Tevatron again.
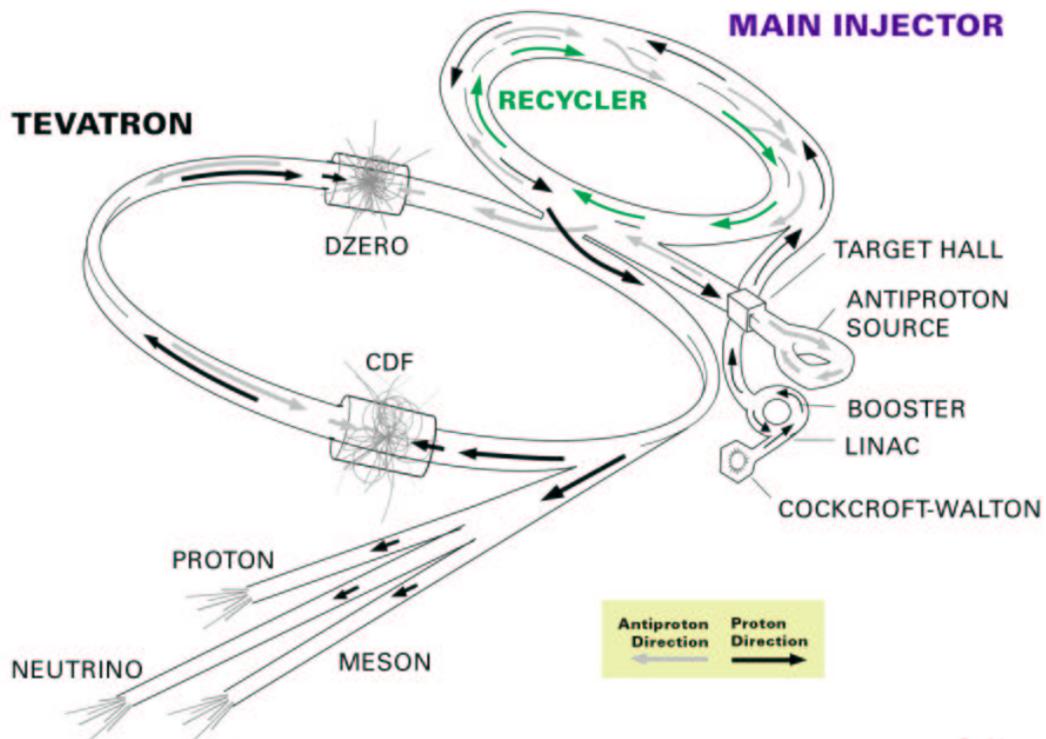
Figure 6: *A schematic view of the Run II Tevatron accelerator complex at Fermilab.*

2.1.3 The Tevatron

The Tevatron is the only cryogenically cooled accelerator at Fermilab. It is a circular synchrotron of radius 1 km that has 8 accelerating cavities in the RF section of the accelerator. At the start of a store, about once per day, the Tevatron receives currently 36 bunches of about 3×10^{11} protons and 36 bunches of about 3×10^{10} antiprotons from the Main Injector and accelerates them from 150 GeV to 980 GeV. Protons and antiprotons share the same ring but move in opposite directions[1]. Two systems of independent superconducting magnets bend the protons and antiprotons so that they do not collide inside the ring except in two points. In the Tevatron collider mode [2], the beam is stored in the Tevatron for several hours. The Tevatron can accelerate the beam every 120-200 seconds. Once 36 bunches of protons and 36 bunches of antiprotons are obtained, the two beams are focused using quadrupole magnets. There are two collision halls on the Tevatron ring, B0 and D0, located at the two collision points of the Tevatron. The former is the collision hall for the CDF experiment, and the later is for the D0 experiment.

Collisions continue typically for 8 hours. At the collision point at the CDF detector, the typical beam size is 120 cm in the direction of the beam axis, and 30 μm in directions perpendicular to it.

2.1.4 Tevatron luminosity

A quantity that characterizes a collider is *luminosity*. Luminosity (L) is the product of incident beam flux with the mean target density. In the absence of a crossing angle or position offset it can be obtained as

$$L = \frac{f B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F(\sigma_l/\beta^*)(cm^{-2}s^{-1}), \tag{2.1}$$

where f is the revolution frequency, B is the number of bunches in each beam, $N_p(N_{\bar{p}})$ is the number of protons (antiprotons) in a bunch, $\sigma_{p(\bar{p})}$ is the RMS proton (antiproton) beam

---

[1]An advantage of a proton-antiproton collider is that protons and antiprotons can share the same ring and therefore greatly reduce the cost of the facility, but the luminosity is limited by the possible intensity of the antiproton beam. On the other hand, in a proton-proton collider, the two proton beams have to be accelerated and stored in separate rings, but the luminosity can reach a much higher value.

[2]There is also another mode: the Fixed Target mode, wherein the beam is sent to Switchyard over a 20 second period. Then, Tevatron can accelerate beam every 60 seconds for the Fixed Target mode.

size at the interaction point, and F is a form factor that depends on the ratio of the bunch length σ_l, to the beta function at the interaction point β^*. The β function is a measure of the beam width, and is proportional to the beam's x and y extent in phase space. The luminosity can be rewritten in a form that more directly displays its dependencies on the limiting factors within the Tevatron complex,

$$L = \frac{3\gamma f}{\beta^*}(BN_{\bar{p}})(\frac{N_p}{\epsilon_{Np}})\frac{F(\sigma_l/\beta^*)}{(1 + \frac{\epsilon_{N\bar{p}}}{\epsilon_{Np}})}, \tag{2.2}$$

where $\epsilon_{Np}(\epsilon_{N\bar{p}})$ is the normalized transverse emittance[3] containing 95% of the proton (antiproton) beam (see [22] for the detailed definition and explanation of accelerator concepts). The major luminosity limitations are dominated by the number of antiprotons $(BN_{\bar{p}})$, and the proton beam brightness N_p/ϵ_p. Note that for a given total number of antiprotons, the luminosity does not depend explicitly on the number of bunches. The luminosity in the Tevatron is proportional to the total antiproton intensity.

The typical luminosity for Run Ib was $L = 1.6 \times 10^{31}$ cm^{-2} s^{-1}, with a peak of about 2.5 $\times 10^{31}$ cm^{-2} s^{-1}. During Run Ia and Ib from 1992 to 1996, CDF collected data corresponding to an *integrated luminosity* of approximately 130 pb^{-1}.

Run II started in June 2001. Unfortunately, the instantaneous luminosities achieved by the Tevatron have not yet met the design goals, but are steadily increasing. This can be seen in figure 7 that presents the peak instantaneous luminosities of every store of Run II. Figure 8 displays the increase of the integrated luminosity delivered by the Tevatron since the start of Run II.

Antiproton availability is the most important limiting factor for attaining high luminosities [18]. The improvements on the accelerator (the replacement of the Main Ring by the Main Injector and, if built, a new antiproton storage ring, the Recycler) should allow to increase the instantaneous luminosity up to $L \approx 20 \times 10^{31}$ cm^{-2} s^{-1}. The plan is to deliver at the end of the Run II era (by 2009) an integrated luminosity ≥ 4 fb^{-1}. Table 3 summarizes the collider characteristics in Run Ib and the goals for Run II.

---

[3]Emittance is, roughly, the density of the tiny clouds of particles and is akin to the cross-sectional area occupied by the beam.
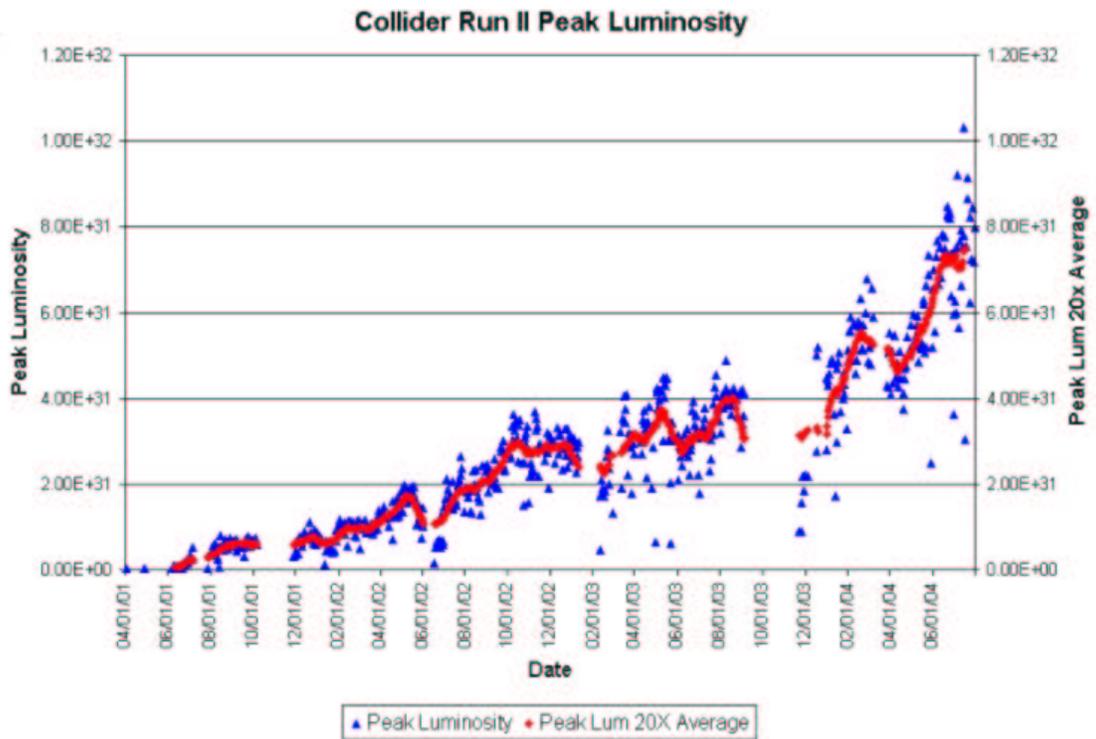
Figure 7: *Initial luminosity per store in $cm^{-2} s^{-1}$.*
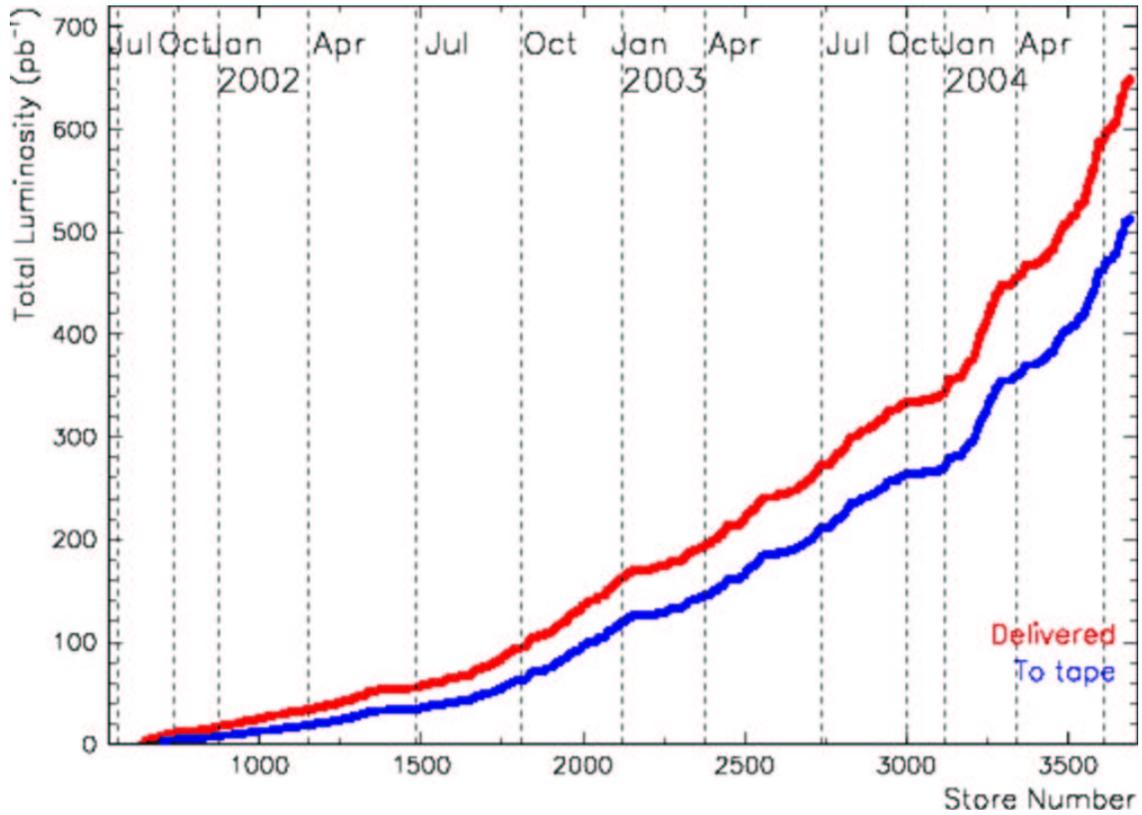
Figure 8: *Total integrated luminosity delivered by the Tevatron (upper curve) and recorded by CDF (lower curve) since the start of Run II. In the analysis presented in this thesis, we use the data taken from March 2002 to August 2003 corresponding to 195 pb^{-1} (only runs with valid SVX information).*

Table 3:   *Parameters describing the accelerator configuration in Run I and Run II. The Run Ib column shows typical operating parameters during Run Ib. The Run II column shows projections for Run II in 36 × 36 operation [18].*

Run	Ib(1993-1995)	II(goals)
Energy	900	998
Protons/bunch	2.3×10^{11}	2.7×10^{11}
Antiprotons/bunch	5.5×10^{10}	3.0×10^{10}
Antiproton Bunches	6	36
$\bar{p}$ production rate [hr^{-1}]	6.0×10^{10}	2.1×10^{11}
total $\bar{p}$	3.3×10^{11}	1.1×10^{12}
Proton emittance [mm-mrad]	23 π	20 π
Antiproton emittance [mm-mrad]	13 π	15 π
β^* [cm]	35	35
Bunch Length(rms) [cm]	60	37
Bunch Spacing [ns]	$\approx$3500	396
Interactions / Crossing	2.5	2.3
Typical Luminosity [cm^{-2} s^{-1}]	0.16×10^{31}	16×10^{31}

2.2   THE CDF DETECTOR

2.2.1   Overview

The Collider Detector at Fermilab (CDF) is a general purpose solenoidal detector with the goal to collect data on a broad range of high energy physics studies. It is capable of studying top, QCD, electroweak and heavy flavor physics, as well as searching for Higgs, SUSY and exotic particles. It combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection. The detector has both azimuthal and forward-backward symmetry with respect to the nominal interaction point. The CDF is built and maintained by a collaboration of more than 50 institutions and eleven countries.

A schematic view of the CDF II detector is shown in figure 9. The tracking system is contained in a superconducting solenoid, 1.5 m in radius and 3.5 m in length, which generates a 1.4 T magnetic field parallel to the beam axis (and opposite to the protons direction) provided by a superconducting solenoid. Calorimetry and muon systems are all outside the solenoid.

To deal with the higher luminosities of the Tevatron in Run II, some parts of the Run I detector and most of the data acquisition system have been replaced. Also, there have been improvements to extend the coverage and capabilities of the existing subdetectors. The calorimetry systems are now exclusively scintillator-based. The electronics and trigger systems are fully compliant with the new pipelined configuration, and all the software has been re-written using C++ and an Object Oriented architecture. A more detailed description of the CDF II detector can be found in its technical design report [19]. The Run I detector is described in detail elsewhere [23, 24].

Figure 10 shows CDF from another perspective, an elevation view of one half of the detector. At CDF, the proton beam moves from west to east; this is used as the z direction, the north horizontal direction is defined as the x direction, and the direction parallel to gravity, with positive being "up", is defined as the y direction, forming a right-handed coordinate system. The polar angle θ in cylindrical coordinates is measured from the proton beam axis; the azimuthal angle ϕ, from the plane of the Tevatron, and around the beam

direction (see figure 11). Throughout this thesis, longitudinal means parallel to the proton beam and transverse means perpendicular to the proton beam. The rapidity, which is defined as :

$$y = \frac{1}{2} \ln \frac{E + P_z}{E - P_z} \tag{2.3}$$

is often used instead of the polar angle θ in the laboratory coordinate frame. The advantage of the rapidity is that a change of rapidity is a constant $\delta y = \tanh^{-1} \beta$ under a boost in the z direction with velocity β. For the case where $E \gg m$, the rapidity can be approximated by the pseudorapidity, defined by $\eta = -\ln(\tan(\theta/2))$. It has to be noted that, for tracks, the pseudo-rapidity is obtained from the direction of the track, and not from the angle θ in the coordinate system defined above. When the track comes from the origin of coordinates, both values coincide, but this is a rather rare case, since the interaction point at CDF is not at the coordinate (0,0,0). The quantity obtained from the θ coordinate is sometimes referred to as detector pseudo-rapidity.

At the most basic level, the CDF detector can be separated into 4 subsystems (that can be seen in the photograph of figure 12): tracking, calorimetry, muon identification, and particle identification. When a particle is created in the $p\bar{p}$ collision, its momentum and charge can only be measured by the tracking system if the particle is charged. Right outside the tracking system, there is the Time-of-Flight detector which measures the time-of-flight of particles. Besides the TOF system, there is the calorimeter in which hadrons, electrons and photons deposit most of their energy. Unlike hadrons and electrons, muons only deposit minimum ionization energy in the calorimeters, so a muon system outside the calorimeter is used to identify muons. Further details of these systems, plus the Cherenkov Luminosity Counters (CLC) and trigger methods are described below, with special emphasis on those relevant to this analysis.

2.2.2 Tracking systems

There are two primary tracking systems in CDF: the inner tracking system, consisting of the silicon-based detectors (L00, SVX II and ISL), and the Central Outer Tracker (COT), a wire-based drift chamber. A schematic view of the tracking systems is illustrated in
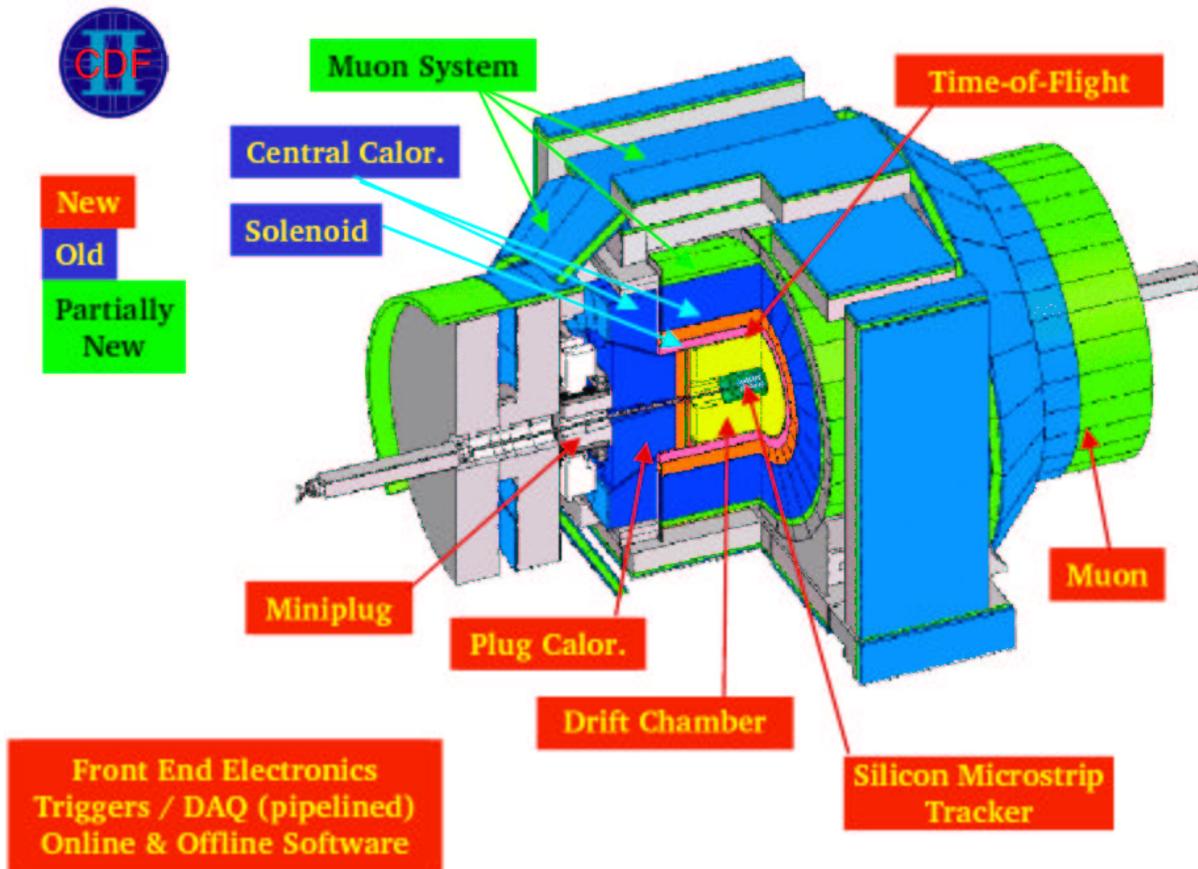
Figure 9: *A schematic view of the CDF II detector.*
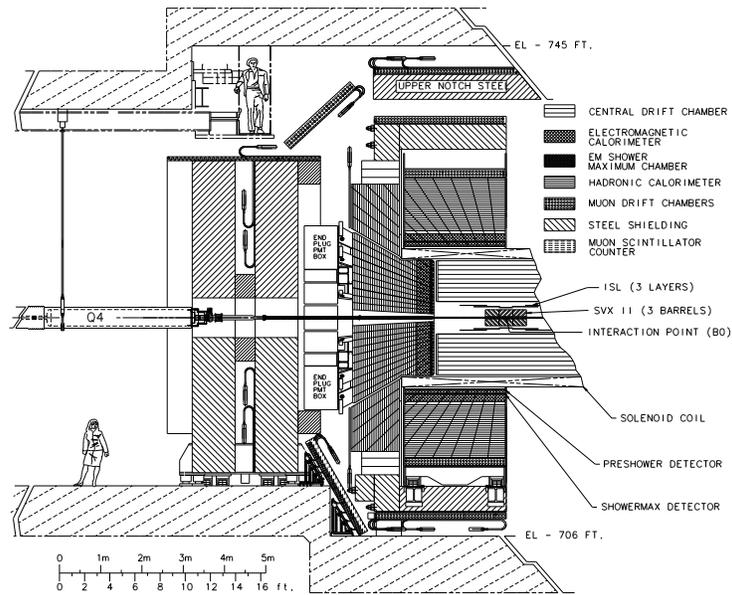
Figure 10: *Elevation view of one half of the CDF II detector.*
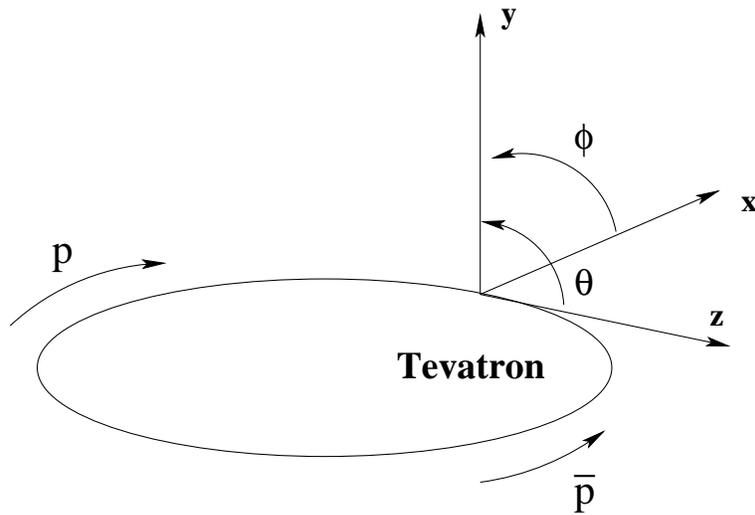
Figure 11: *Definition of the coordinate system in CDF.*
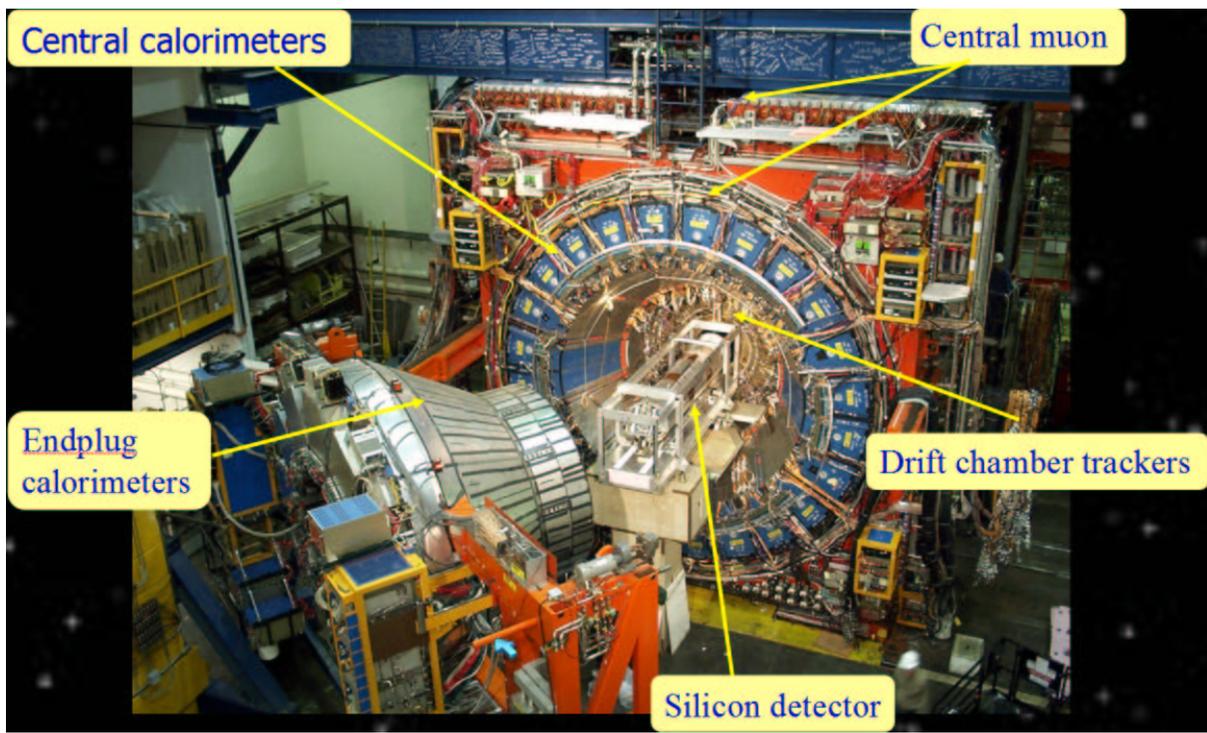
Figure 12: *Photograph of the Collider Detector at Fermilab.*

figure 13. Starting from the innermost, the subsystems are the *Layer00* (L00), the *Silicon Vertex Detector* (SVX II), and the *Intermediate Silicon Layers* (ISL) detectors constitute the inner tracking system, and the COT constitutes the outer tracking system. The COT covers the central region, in the range $|\eta| \leq 1$, with high momentum resolution and reconstruction efficiency. The silicon tracker is used not only to precisely reconstruct the track impact parameter or interaction point, but also to be able to perform the silicon stand-alone tracking for the region $(1 < |\eta| < 2)$ which is not covered by the COT.

The tracking system provides precise momentum measurement of charged particles. Also, the spatial resolution of COT-based tracks is greatly improved by adding hits from the silicon detectors to them. The resulting tracks, with silicon information, can then be used to obtain a precise estimation of the luminous region and of the position of the primary vertex, as well as to reconstruct displaced vertices and other quantities (e.g. information on jets). These capabilities allow a rich *b*-physics program at CDF and, consequently, contribute to Higgs searches and top physics because both Higgs and top decay mainly via the *b* quark. For *b* physics, the precise momentum and position resolution of the combined tracking system provides an excellent platform for mass and lifetime measurements. In mass measurements, for instance, requiring a displaced vertex significantly reduces the huge QCD background in the $p\bar{p}$ collision environment. In terms of this analysis, besides the necessary determination of event-by-event decay lengths, the tracking system is used in combination with the muon systems to identify muon tracks and to reconstruct J/ψ candidates, which are required in all the four decay channels we consider.

2.2.2.1 Tracks. We reconstruct the trajectories of charged particles ("tracks"), in the COT and SVX II. In a homogeneous magnetic field, tracks bend to form helices. Their transverse momenta can be related to half-curvature C by $p_T = B/2cC$, where c is the speed of light and B the magnitude of the magnetic field. The half-curvature is generally referred to as curvature. The tracking algorithms use 5 parameters to describe the helices: impact parameter (d_0), curvature(C), ϕ_0, z_0, and $\lambda = \cot\theta$. These are illustrated in figure 14. The curvature is a signed quantity, positive for counterclockwise trajectories and negative for clockwise trajectories, and its sign is the charge of the corresponding particle.
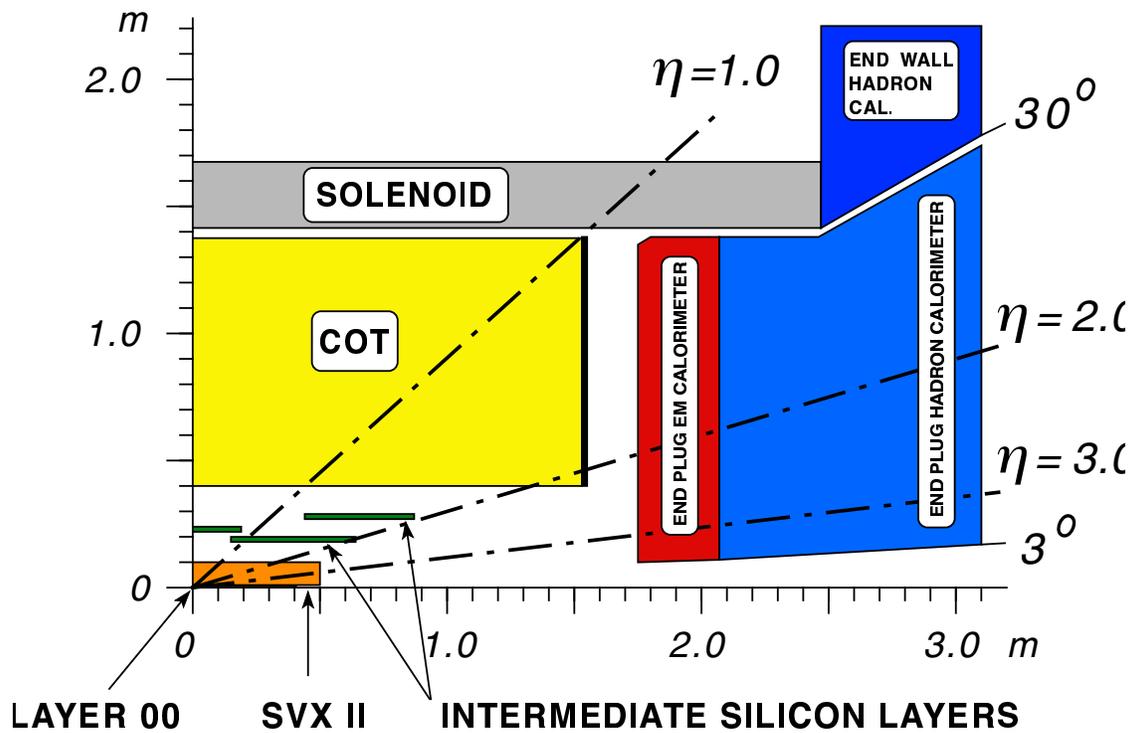
Figure 13: *A cut-away view of one quadrant of the inner portion of the CDF II detector showing the tracking region surrounded by the solenoid and end-cap calorimeters.*
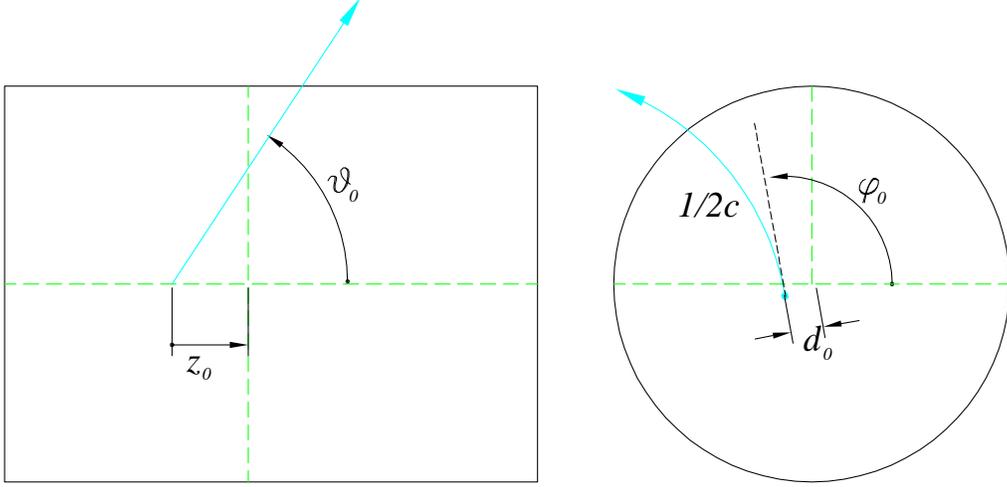
Figure 14: *Illustration of the tracking parameters: the figure on the left shows a charged track in the $r - z$ view of the tracking volume. The figure on the right shows an $r - \phi$ view of the track. d_0 is the track's impact parameter, or point of closest approach to the origin. c is the curvature.*

The helix is a circle in the xy plane, whose curvature is related inversely to the transverse momentum, as described above. The circle in the xy plane has a well-defined point of closest approach to the origin, $\vec{P}$. ϕ_0 is the angle between the x axis and a line tangent to the track at $\vec{P}$. The impact parameter of a track has a sign that is defined by the following formula:

$$d_0 = \frac{\hat{z} \cdot (\vec{r} \times \vec{P_T})}{\left| \vec{P_T} \right|},\tag{2.4}$$

where $\vec{P_T}$ is the transverse momentum vector of the particle, $\vec{r}$ is the vector pointing from the primary vertex to the reconstructed particle trajectory at the point of closest approach to the primary in the $r - \phi$ plane and $\hat{z}$ is the unit vector along the z axis. In other words, the signed impact parameter d_0 is the y-intercept of the track, after rotating the coordinate system so that $\phi_0 = 0$ [4]. z_0 is the position of the track along the z-axis at P, and λ is defined as $\lambda \equiv \cot \theta = p_z/p_T$, being θ the polar angle at the minimum approach. The axial

---
[4]If a track has a positive(negative) charge and the reference point is outside(inside) the circle of the track, then the impact parameter has a positive sign, see figure 15.
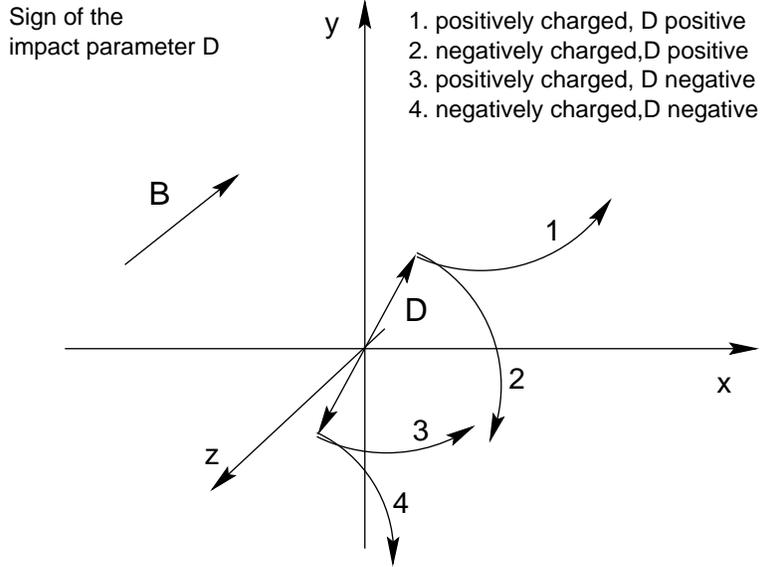
Figure 15: *Track of a particle with positive/negative charge and positive/negative impact parameter.*

parameters, which provide information in the xy plane only, are C, d_0 and ϕ_0. λ and z_0 are called stereo parameters, since they provide z information.

2.2.2.2 Central outer tracking system: COT. The COT, or Central Outer Tracker, is the main tracking chamber in CDF [19, 25]. It is an open-cell cylindrical drift chamber segmented into 8 concentric superlayers for particle reconstruction in the central region $|\eta| \leq 1$ with transverse momenta as low as 400 MeV/c. The active volume extends 310 cm in z, and from 40 to 137 cm from the beam line in radius. The entire azimuth, ϕ, is covered. Each superlayer is sectioned in ϕ into separate cells. A cell is defined as one sense plane (active and read-out) with two adjacent field planes, which are grounded. A diagram of a section of the endplate, with slots for the field and sense planes, is shown in figure 16. Figure 17 shows a diagram of 3 cells in the $r - \phi$ plane. The rows of small circles and crosses represent high voltage wireplanes. There are a total of 29 wires in each cell, 12 of which, called sense wires, are read out. The sense wires have approximately the same maximum drift distance; therefore, the number of cells in a given superlayer is roughly proportional

34

to the radius of the superlayer. The remaining wires are needed to shape the electric field, adjusting for the taper of the cell with decreasing radius. The lines adjacent to the sense planes represent the grounded field planes.

The sense plane wires are composed of 40 μm gold-plated tungsten wire. The main body of the field sheets is 0.25 mil gold-coated mylar. The field sheets are separated by $\simeq$ 2 cm. Mylar is stretched and supported by two 12 mil stainless steel wires, which are epoxied in a parabolic shape along each side of the field sheet. In axial superlayers, they are approximately parallel to the z-axis. The field sheets are much closer to a true grounded plane than arrays of wire, which have often been used in wire chambers, including the predecessor to the COT. Use of field sheets results in a smaller total radiation length, and allows the COT to operate at much higher drift field than with an array of wire. This is an important factor in maximum drift time. In addition to this, the total endplate load is less, because a single field plane requires less tension than an array of field wires.

The eight superlayers of the COT alternate between stereo and axial, beginning with superlayer 1, which is a stereo layer. In an axial layer, the wires and field plates are parallel to the z axis, and thus provide only $r - \phi$ information. In stereo layers, a given wireplane or field sheet which starts at a slot in one endplate does not end at the mirror-image slot in the other. Instead, it is offset by 6 cells. This generates a stereo angle of $\pm 3^o$, depending upon the direction, which corresponds to a rotation about an axis in the radial direction.

The COT is filled with Argon/Ethane(50-50) with a drift velocity of $\simeq$ 100 μm/ns. This will give a maximum drift time of 180 ns. When a charged particle passes through, the gas is ionized. Electrons drift toward the sense wires, resulting in an avalanche at the wire surface, which provides a gain of $\approx 10^4$. For a charged particle traveling through the entire COT radially, the 4 axial and stereo superlayers will provide 96 (8×12) measurements. The drift time of ionization electrons in the gas is used to measure the charged particle's spatial position and the pulse height can be used to measure the amount of ionization.
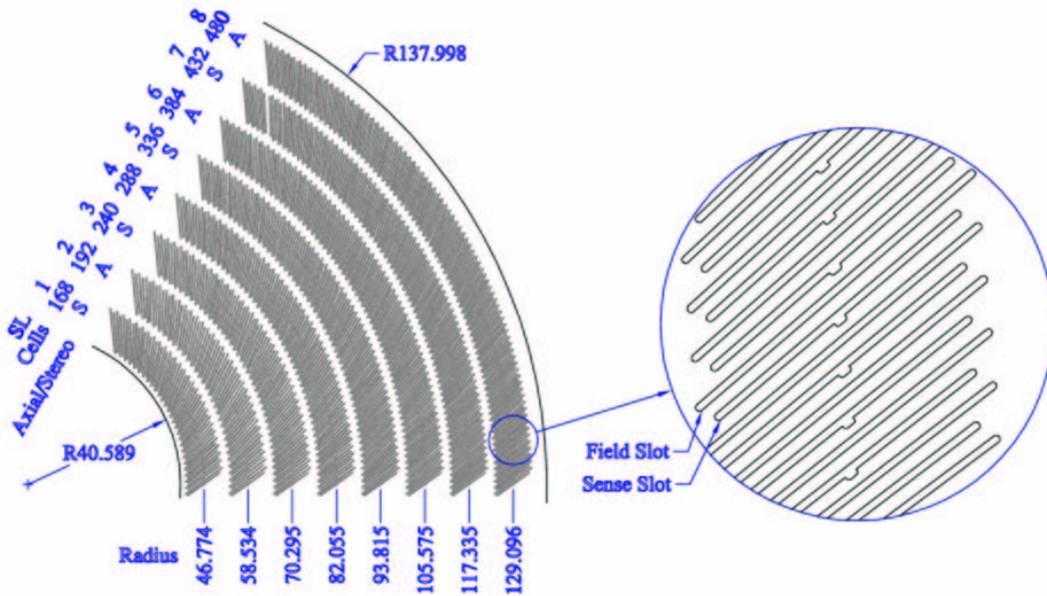
Figure 16: *1/6 section of the COT endplate, showing the cell counts in superlayers 1-8. The radii at the center of each superlayer are shown in cm. The endplate has inner and outer radii of 40.589 cm and 137.998 cm.*
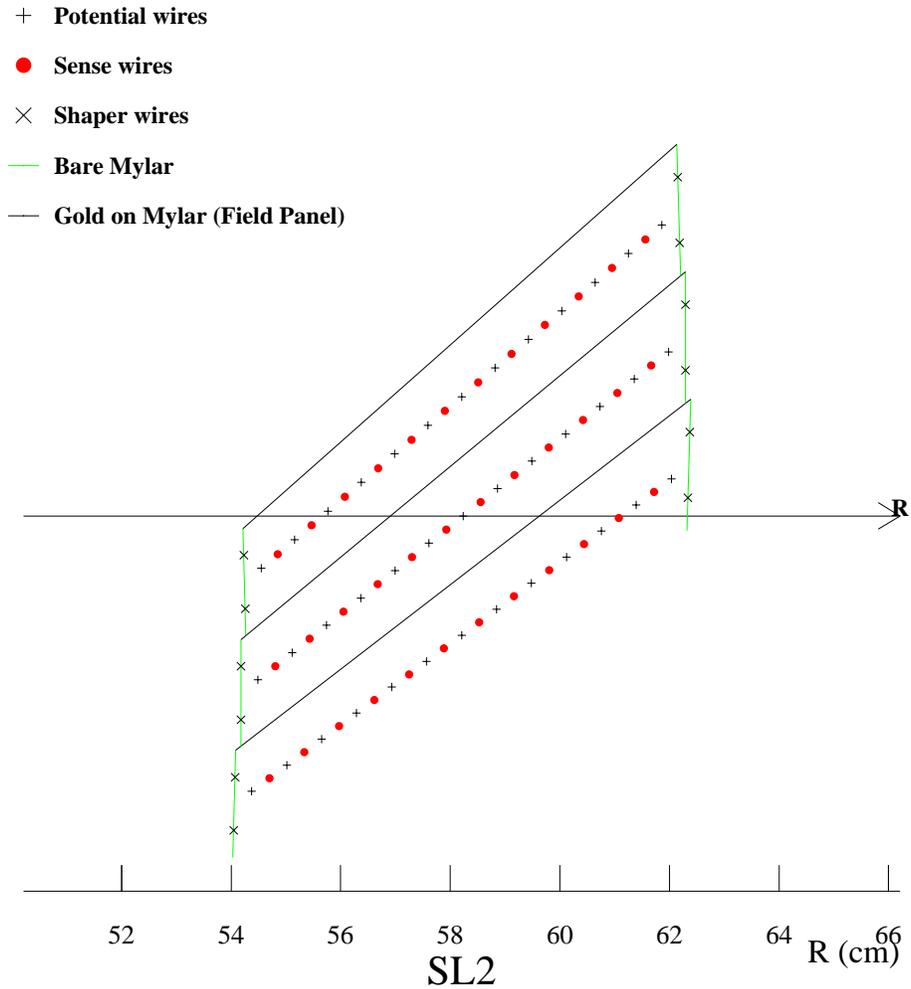
Figure 17: *A transverse view of three cells in superlayer 2 of the COT. The continuous lines represent field sheets, which are grounded. The 11.8 cm length arrays of circles and crosses represent high-voltage wireplanes. The circles are the sense wires, which are read out. The crosses are the field wires.*

Due to the magnetic field, the electrons drift with a Lorentz angle of $\approx 35^o$. It is for this reason that the cells are tilted with respect to the radial direction. A simulation of the electron drift using is shown in figure 18 [26]. The voltage on the wire planes is set in order to insure the maximum drift time to be smaller than the time between beam crossings, which is 396 ns. Table 4 summarizes the main features of the COT.

Signals on the sense wires are processed by an ASDQ Chip (Amplifier Shaper Discriminator with Charge Encoding), which provides input protection, amplification, pulse shaping, baseline restoration, discrimination, and charge measurement [25]. This charge measurement is encoded in the width of the discriminator output pulse, which also allows the measurement of the energy loss through ionization of the gas per unit length (dE/dx)[5]. The pulse is sent through ≈ 35 ft. of micro-coaxial cable, through repeater cards, and finally to the TDC's (Time to Digital Converters), which reside in the collision hall. Hit times are then processed by pattern recognition and fitting algorithms to form helices. These algorithms are collectively referred to as "tracking". Figure 19 shows the COT hit resolution vs. drift distance, measured by an online monitoring program. The single hit resolution is about 150 μm in the center of the cell.

2.2.2.3 Inner Tracker: L00 + SVX II + ISL.
One the most important components of CDF is the silicon detector [27, 19]. Among its major features, we find a large geometric coverage of the interaction region, fast data acquisition and trigger systems, and radiation-hard sensors to survive the radiation fields created by the expected Run II luminosity (capable of withstanding several Mrads of integrated dose). The CDF silicon detector is essential for enhancing not only the overall tracking capabilities but also the heavy flavor tagging of the experiment. The inner tracking system consists of three silicon detectors [27]: The Silicon Vertex Detector (SVX II), the Intermediate Silicon Layer (ISL) and Layer 00 (L00), described below. Comprising a total of 7-8 silicon layers arranged in cylinders, these detectors allow one to achieve, among other things, a good impact parameter resolution and a silicon stand-alone tracking.

---

[5]The dE/dx of a charged particle is a function of particle velocity, which can be used to infer the particle mass when combined with the information on the particle momentum.
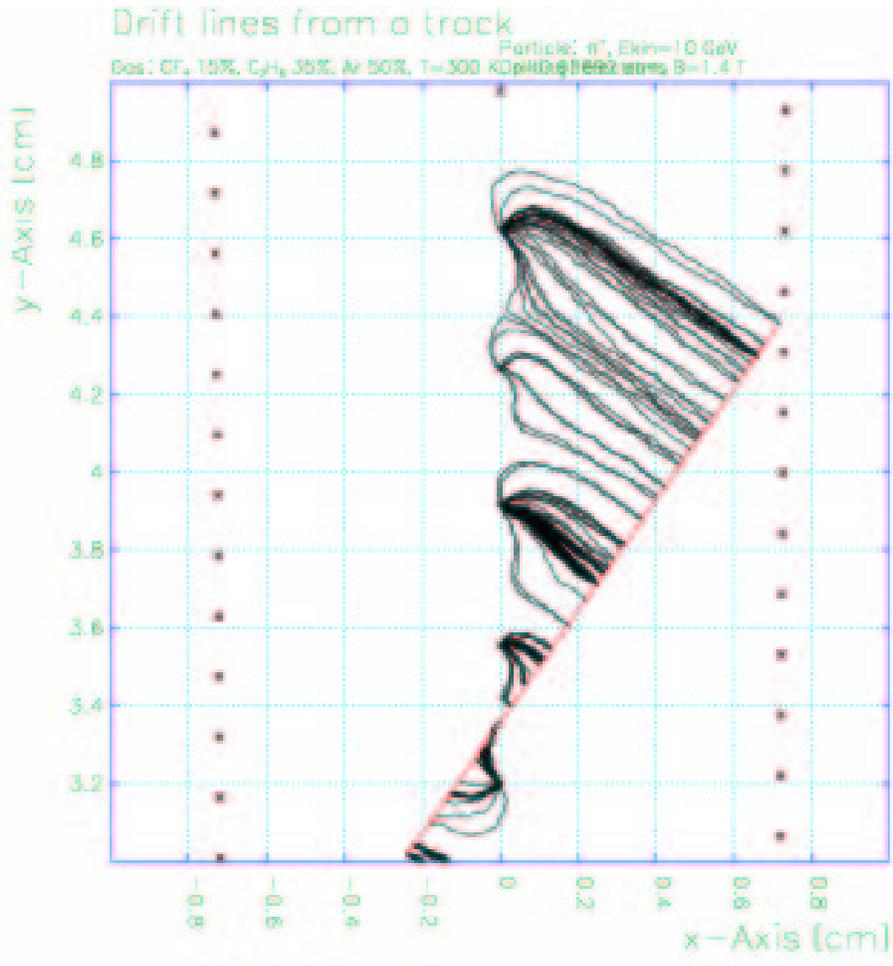
Figure 18: *Simulated drift of electrons toward the sense wires [26]. The straight line represents a charged particle passing through the detector. Along the track, ionizations occur, and the liberated electrons, or clusters of electrons, drift toward the sense wires. The direction of drift is determined by the electric field (due to the sense wires, potential wires, and field sheets) and the magnetic fields, which is required for the COT to function as a spectrometer.*

Table 4: Summary of the main COT features.

COT	
Number of superlayers	8
Measurements per superlayer	12
Stereo Angle [degrees]	+3 0 -3 +3 0 -3 0
Cell/Layer	168 192 240 288 336 384 432 480
Radius at Center of SL [cm]	46 58 70 82 94 106 119 131
Tilt angle [degrees]	35
Length of Active Region	310 cm
Number of channels	30,240
Material thickness	1.3 % X_0

The Silicon Vertex Detector (SVX II) is built in three cylindrical barrels with beryllium "bulkheads" at each end for support and cooling of the silicon units. It is positioned end-to-end along the beam axis and centered longitudinally with the detector with a total length of 96 cm and a coverage in pseudorapidity of $|\eta| \leq 2$. Each barrel is divided in azimuth into 30° wedges and each wedge consists of five radial layers of double sided silicon microstrip detectors between radii of 2.4 and 10.7 cm. One layer consists of two wire-bonded pairs of double-sided silicon microstrips sensors of 7.28 cm length. The bulk of the sensor is n-doped. On the side facing the beamline, the strips are spaced in $r\phi$ by approximately $60\,\mu$m, and have p$^+$-implant widths of 14-15 μm(resulting in a high density array of pn diodes). On the other side, both 90° and small angle stereo sensors are used, in the pattern (90 90 -1.2 90 +1.2) degrees from the innermost to the outermost SVX II layer. They are spaced by (141, 125.5, 60, 141, 65) μm, and have n$^+$-implant widths of $2\,\mu$mfor the 90° strips and $15\,\mu$mfor the small angle stereo layers. This is designed to allow good resolution in locating the z-position of secondary vertices and to enhance the 3-D pattern recognition capability of the silicon tracker.

COT Resolution vs Hit Distance
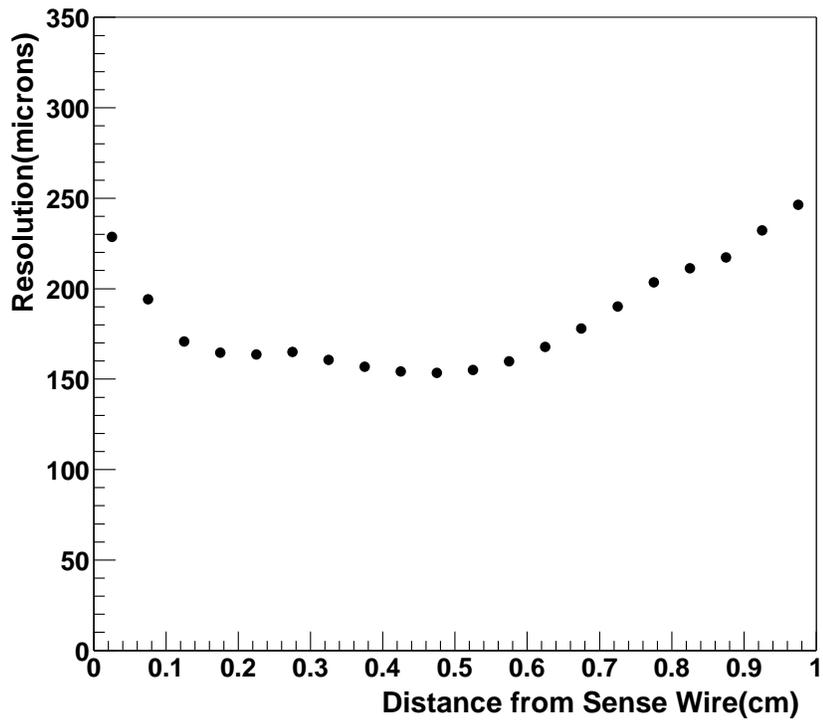
Figure 19: *Single hit resolution vs. drift distance measured in data. The measurements are an average over all superlayers.*

When the detector is reverse-biased, passage of a charged particle through the material results in the creation of electron-hole pairs. Due to the electric field, the charge then drifts toward the readout strips, providing position measurements. Figure 20 shows a detailed end view of the SVX II system and table 5 shows a summary of some of the SVX II features.

An additional double-sided silicon layer at a radius of 22 cm covering the interval $|\eta| < 1$ and two layers of silicon placed at the radii of 20 and 28 cm in the forward and backward region covering $1 < |\eta| < 2$ [28]. The ISL incorporates many features of the SVX II design. The crystals are double-sided with axial strips on one side and small angle stereo at 1.2^o on the other. The readout electronics are identical to the SVX II (see below). In the forward and backward region, where the COT tracking is not efficient, the silicon stand-alone tracking is performed. The ISL thus extends tracking, lepton identification, and b-tagging capabilities over the full region $|\eta| < 2$. However, we do not use this detector for the present analysis. The main reason for this decision is that the properties of tracks without ISL information are better understood at this point than the properties of tracks with ISL hits.

The innermost Layer 00 (L00) [29] is a single-sided, radiation-hard silicon layer, placed immediately outside the beam pipe at 1.35-1.62 cm radius. The readout electronics is identical to that used for SVX II. Being so close to the interaction region, L00 significantly enhances CDF's impact parameter resolution (see the results achieved with L00 on impact parameter resolution below, subsection of *Tracking performance with silicon*). For this analysis, we do not make use of L00 because the improvement in resolution is not needed for a measurement of B lifetimes for which the error is already much smaller than the mean value. Also, noise problems in the L00 system make it difficult to use at this time.

All components of the CDF II silicon system achieve their data readout through a set of custom integrated circuit chips with the designation SVX3D [30]. It includes preamplification, a multi-cell analog storage pipeline, and simultaneous analog and digital operation capability. An optional data acquisition mode allows common-mode noise to be reduced independently for each chip by dynamic data-driven determination of pedestal levels. For SVX II, the SVX3D chips are mounted on electrical hybrids on the surface of the silicon detectors. ISL, with more space available, has the hybrids attached right after the sensors. L00 hybrids are well separated from the sensors at the end of the detector and the SVX3D
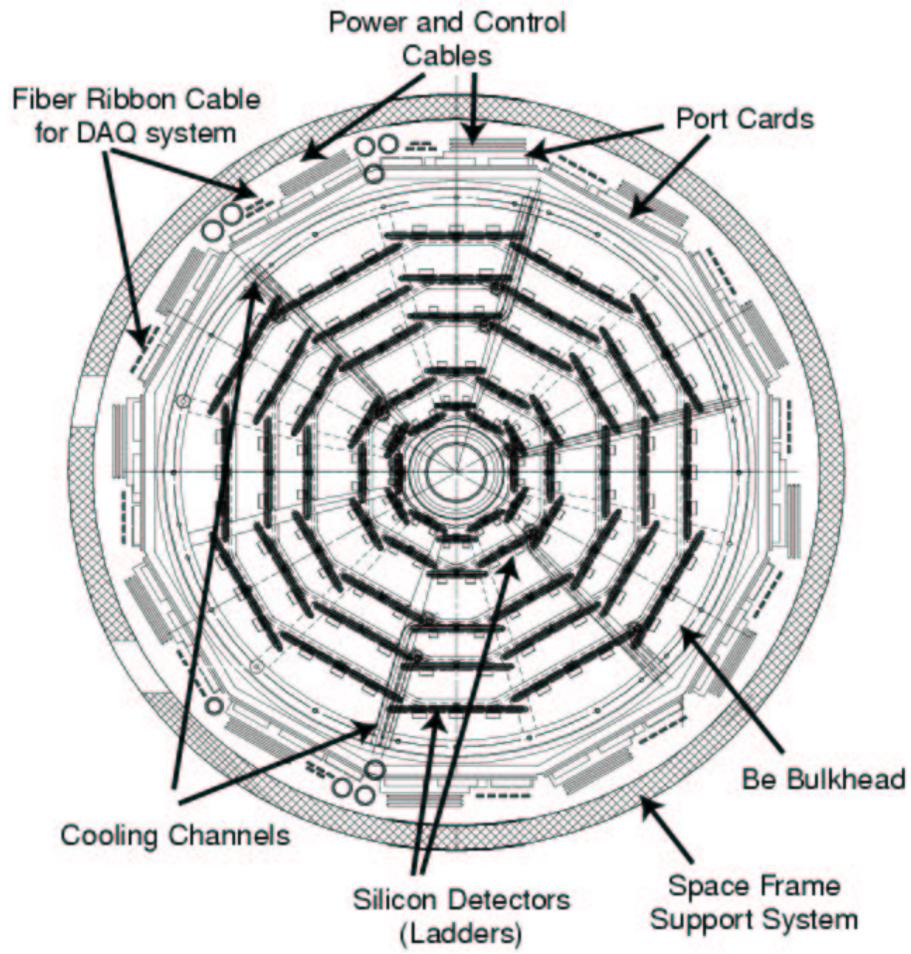
Figure 20: *An end view of SVX II including the cooling and support systems.*

Table 5: Summary of the main SVX II features.

SVX II	
Readout coordinates	$r - \phi, r - z$
Number of barrels	3
Number of layers per barrel	5
Number of wedges per barrel	12
Ladder length	29.0 cm
Combined barrel length	87.0 cm
Radius at axial layers	2.545 4.120 6.520 8.220 10.095
Radius at stereo layers	2.995 4.570 7.020 8.720 10.645
number of ϕ strips	256 384 640 768 896
number of Z strips	256 384 640 512 896
ϕ strip pitch	60 62 60 60 65 μm
Z strip pitch	141 125.5 60 141 65 μm
Number of channels	405,504
Material thickness	3.5% X_0

chips are connected via fine-pitched cables to their corresponding L00 sensor. Each readout chip has 128 channels, each with a charge-sensitive amplifier, 42-cell dual-ported pipeline with four additional cells for buffers, and an ADC.

The data acquisition system for the CDF II silicon detectors is a fully pipelined DAQ+ trigger architecture that can operate without deadtime losses at machine bunch crossing intervals as low as 132 ns. The 42-cycle "Level 1" (L1) pipeline within each SVX3 chip allows storage of the analog signals for later digitization and transmission at rates up to 50 kHz. A highly parallel fiber-based data acquisition system reads out the entire detector in approximately 10 μs. This high speed and dual porting of the readout allows the SVX II information to be used for impact parameter discrimination in the Silicon Vertex Trigger (SVT) processor of the "Level 2" (L2) trigger. This allows an online trigger to identify displaced tracks at L2, as described in references [31, 32, 33, 34]. This trigger operates with 20 μs latency at rates up to 300 Hz. It is followed in the data stream by a third level of software-based trigger processing that implements a portion of the offline analysis and reduces the final rate of logged events to less than 50 Hz. Further detail on the components and logic of the silicon DAQ system can be found in [35, 36, 37].

Here we give a brief description of the performance of the silicon detector in terms of some relevant parameters (like $\frac{Signal}{Noise}$, single hit efficiencies, intrinsic resolution, alignment, etc.) as well as the status of the combined COT-Silicon tracking.

The silicon system is running ∼92.5% of its modules and getting good quality data from ∼85% of them. On average, for any ladder type and chip setting, $\frac{Signal}{Noise}$ >= 10, with operating conditions set at 11-15 depending on ladder type. Figure 21 shows a distribution of a typical ladder's pedestal and signal ADC counts, exhibiting clear separation between the Gaussian noise and Landau distribution for the signal. With high signal to noise, the single hit efficiencies can be kept over 99%, maintaining high signal efficiency without prohibitive data volume. With tracks in the data, both the global and relative alignment continue to improve. In addition, the cluster resolution of the detector is as expected, with for example a 9 μm intrinsic resolution on 2 strip clusters for the $r-\phi$ measurement, as shown in figure 21.
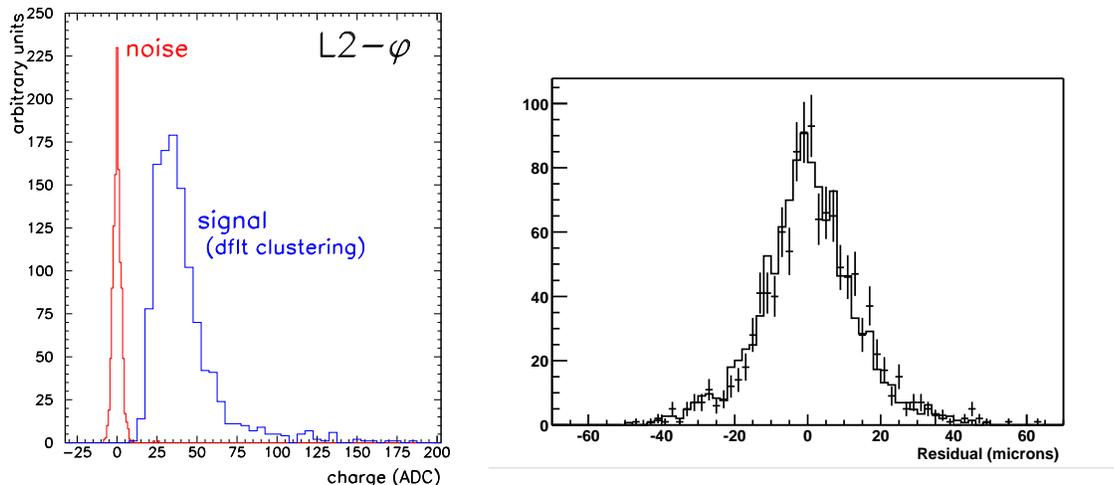
Figure 21: *Pedestal and signal for a SVXII L2 ϕ-side ladder, and Two-Strip ϕ cluster resolution ($\sim 9\,\mu$m) (points), which agrees with the Monte Carlo simulation (histogram).*

2.2.2.4 Tracking performance with silicon. The silicon efficiency has been measured for muons from J/ψ decays. All COT tracks which are extrapolated through three SVXII layers are considered in the calculation. The efficiency is defined to be the fraction of those tracks that are assigned three silicon hits from three working layers of silicon. Background under the J/ψ mass peak is estimated from the sidebands of that peak, and subtracted. The efficiency is determined to be 93%. The fake rate, defined as the fraction of events for which the signed impact parameter is negative and lying more than 3 σ away from the beam line, is determined to be 1.6%. Figure 22 shows both the efficiency and the fake rate so defined, for different selections of quality cuts.

Multiple scattering in passive material from the readout electronics and support structure of the other layers degrades the resolution for low momentum tracks ($p_T < 3$ GeV/c). L00 was designed to recover that resolution. The effect is particularly relevant in the high-mass regions (those ranges in z with many electrical components for SVX II readout) [see figure 23]. Since we do not use L00 in this analysis, the resolution we obtain is slightly worse under those regions. The region in z with less passive material is shown in the same figure (right plot). Notice that the beam spread has not been subtracted from the plots and has a contribution to the impact parameter resolution of $\simeq 33\,\mu$m. For tracks with a P_T greater
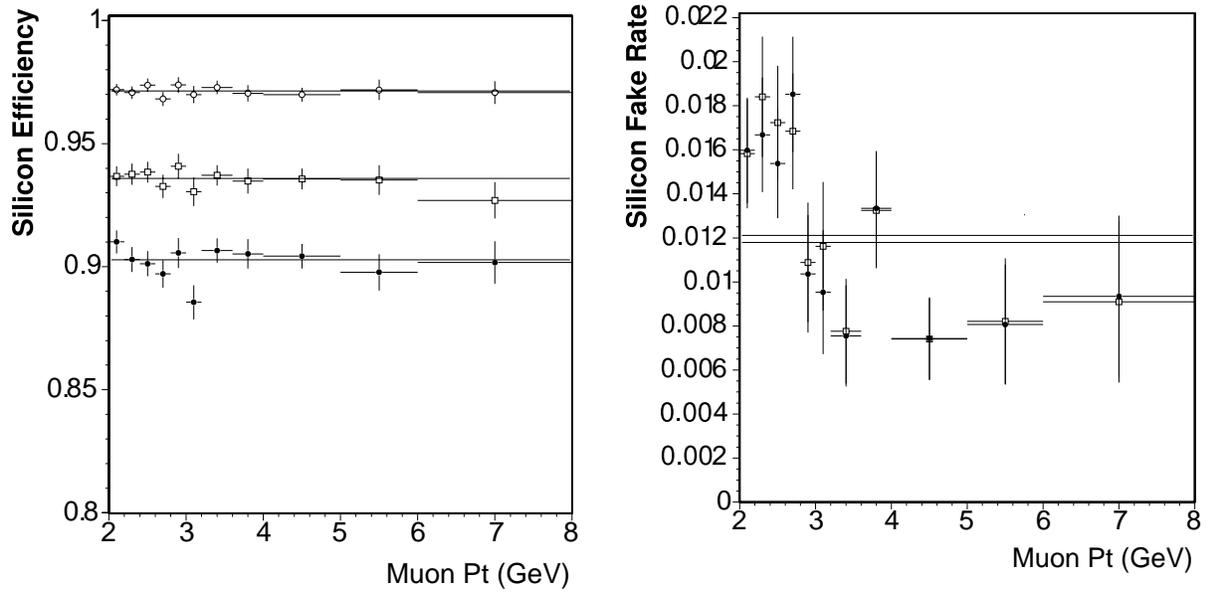
Figure 22: *Silicon efficiency for J/ψ muons with background subtraction (left), and fake rate as a function of muon P_T (right), for different selections of quality cuts.*
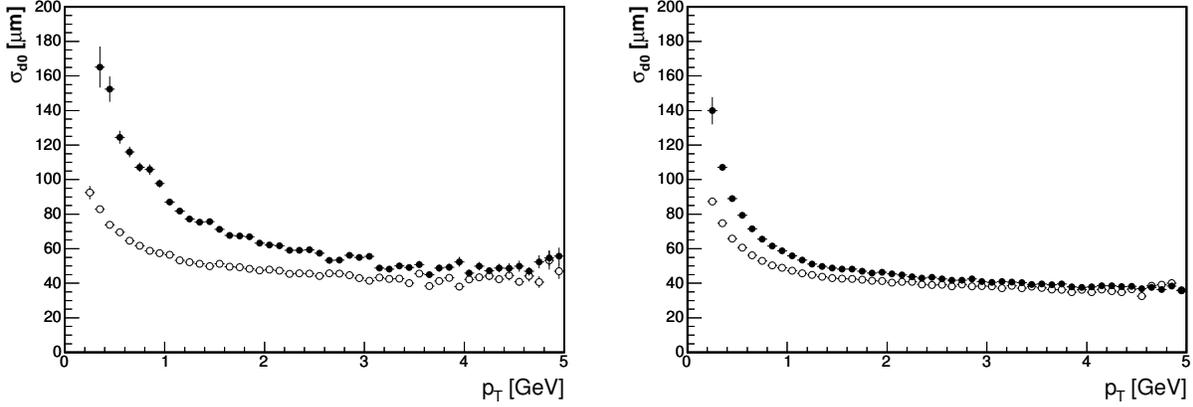
Figure 23: *Left: Impact parameter resolution with (hollow circles) and without (points) using L00 in the z region with high mass (passive material). Right: The region in z with less passive material is also shown in the right figure. As expected, The use of L00 results on a significant improvement on the impact parameter resolution.*

than $1.5\,\mathrm{GeV}$(as is the case of the J/ψ tracks we select), the impact parameter resolution (including the beam spread) is of the order of $50\,\mu\mathrm{m}$, this can be compared with the mean lifetime we are measuring, which is close to $500\,\mu\mathrm{m}$.

Using the silicon detectors and the COT, the overall momentum resolution for charged particles is $\delta P_T < 0.1\% GeV/c$.

2.2.3 Pattern recognition in tracking

Track reconstruction begins in the outer tracking chamber - the COT. The first step in the pattern recognition is to form line segments from hits in each superlayer. Line segments from the axial layers which are consistent with lying tangent to a common circle are linked together to form a track. A 2D circle fit is then performed. Line segments in stereo layers are then linked to the 2D track, and finally a helix fit is performed. At this point we have a set of tracks which have only COT hits [38]. These are referred to as COT-only tracks.

The next step is to extrapolate the COT-only track into the SVX and add hits which are consistent with lying on that track. This process starts with the outermost layer in SVX II. A road, or window, around the track is established based on the errors on the COT track

48

parameters. If hits lie within the road, they are added to the track. A new track fit is then performed, resulting in a new error matrix and a new road. This road is then used to add hits from the next SVX layer. This procedure is repeated until there are no SVX layers left. There may be multiple tracks with different combinations of SVX hits associated with one COT track. In this case, the track with the largest number of SVX hits is chosen [39].

The set of tracks resulting from this process will be henceforth referred to as "combined tracks." This set is a mix of tracks with varying numbers of SVX hits, including a certain number with zero SVX hits; i.e, COT-only tracks. Everyone of the combined tracks has a unique COT-only parent, which is stored in the event record. We point this out here because, in this analysis, we sometimes prefer the COT-only tracks. This means we use the COT-only parent of a given track from the combined tracks collection.

2.2.4 Calorimeter system

CDF has both electromagnetic and hadronic calorimetry. The central system is segmented into towers in η and ϕ. These towers span 15° in azimuth and 0.11 in pseudorapidity. The coverage is $0 < |\eta| < 1.3$. For the forward region, a system called the Plug Calorimeter, with variable tower size, extends coverage out to $|\eta| = 3.6$. This analysis uses only Tracking and Muon Identification. For a more detailed description of the calorimetry, we refer the reader to the Run I description of the CDF detector [23] and the Run II Technical Design Report [19], which discusses the Plug Upgrade Calorimeter.

2.2.5 Muon systems

The muon systems are specially relevant for this measurement. Each of the four decay channels we use includes a J/ψ particle. Although the J/ψ decays into hadrons most of the time, it has a significant branching fraction (roughly 6%) into two muons. This provides a very effective way to identify potential J/ψ events, as those in which (at least) two muons have been detected.

Because the muon is roughly 200 times heavier than the electron, its bremsstrahlung radiation is about 40000 times weaker than that of an electron, so muons deposit relatively

very little energy in the calorimeters, which totally absorb most other particles. In order to identify muons, there are several muon detectors outside the calorimeters; these detectors do not measure charge or momentum. However, since the muon is a charged particle, the tracking system can determine those quantities and reconstruct muon tracks. These tracks, extrapolated to the muon systems' region, can be used in conjunction with hits in the muon detectors to provide good muon identification capabilities. CDF uses four systems of scintillators and proportional chambers in the detection of muons over the region $|\eta| < 2$. The absorbers for these systems are the calorimeter steel, the magnet return yoke, additional steel walls, and the steel from the Run I forward muon toroids. New chambers have been added for the CMP and CMX systems (described below) to close gaps in the azimuthal coverage in Run II, while the central chambers (CMU) have almost the same configuration without major changes from Run I. Finally, the forward muon systems used in Run I was replaced with a completely new Intermediate Muon System (IMU). For this analysis we do not make use of the last one.

The Central Muon Detector (CMU) is the oldest muon detector and is fully documented in [40]. It consists of a set of single-wire drift tubes arranged in 144 modules with 16 rectangular cells per module, located behind $\simeq 5.5$ λ of absorber (embedded in the Central Hadron Calorimeter) at radius 347 cm. Each module consists of four layers of drift tubes. The CMU is divided into two halves at $\theta = 90^o$, and each half is segmented into 24 wedges. The wedge segmentation in the ϕ direction is the same as the central calorimeter (see figure 24). Each wedge has three towers, with dimensions 63.5 mm (x), 26.8 mm (y), 2262 mm (z), and each tower consists of 16 cells. Each cell has a cross section of 2.68×6.35 cm and has a 50 μm stainless steel sense wire, which runs parallel to the z-axis, located at the center of the cell (see figure 25). The drift time information recorded by the stainless steel wire is used to calculate the muon $r - \phi$ position, while the charge distribution along wires is used to determine the z position. This information will be used for the muon match between stub and the tracks from the tracking system.

The Central Muon Upgrade (CMP) consists of a second set of muon chambers behind an additional 60 cm of steel (for a total of ~ 7.8 λ of absorber) in the region $55° \leq \theta \leq 90°$, and it is located outside of the magnetic field's return yoke [40]. The pseudorapidity coverage of

y

x ←

Muon
Chambers

3470 mm

CHA

CEM

Φ = 6.31

Φ = 7.50

y

→ z

← 2260 mm →

Central
Calorimeter
Wedge

Tower   0 1 2 3 4 5 6 7 8 9

Θ = 88.5

Θ = 55.9

beam axis

interaction region
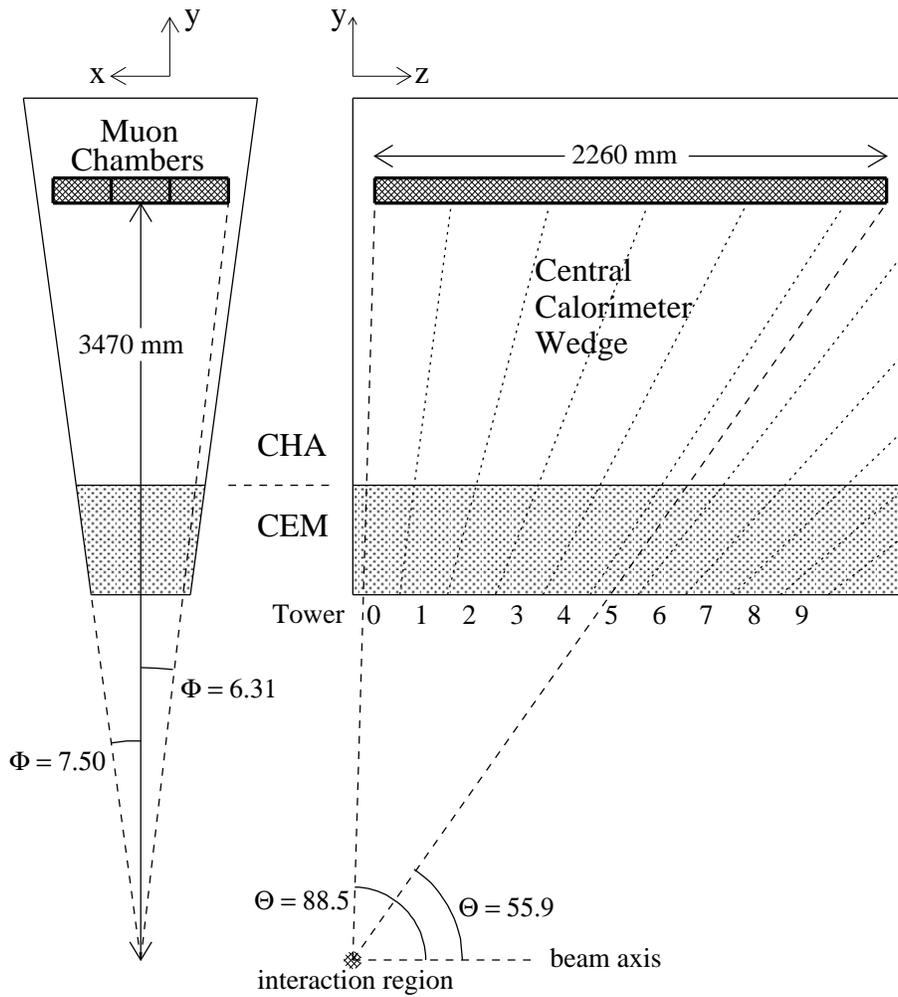
Figure 24: *A wedge of CMU.*
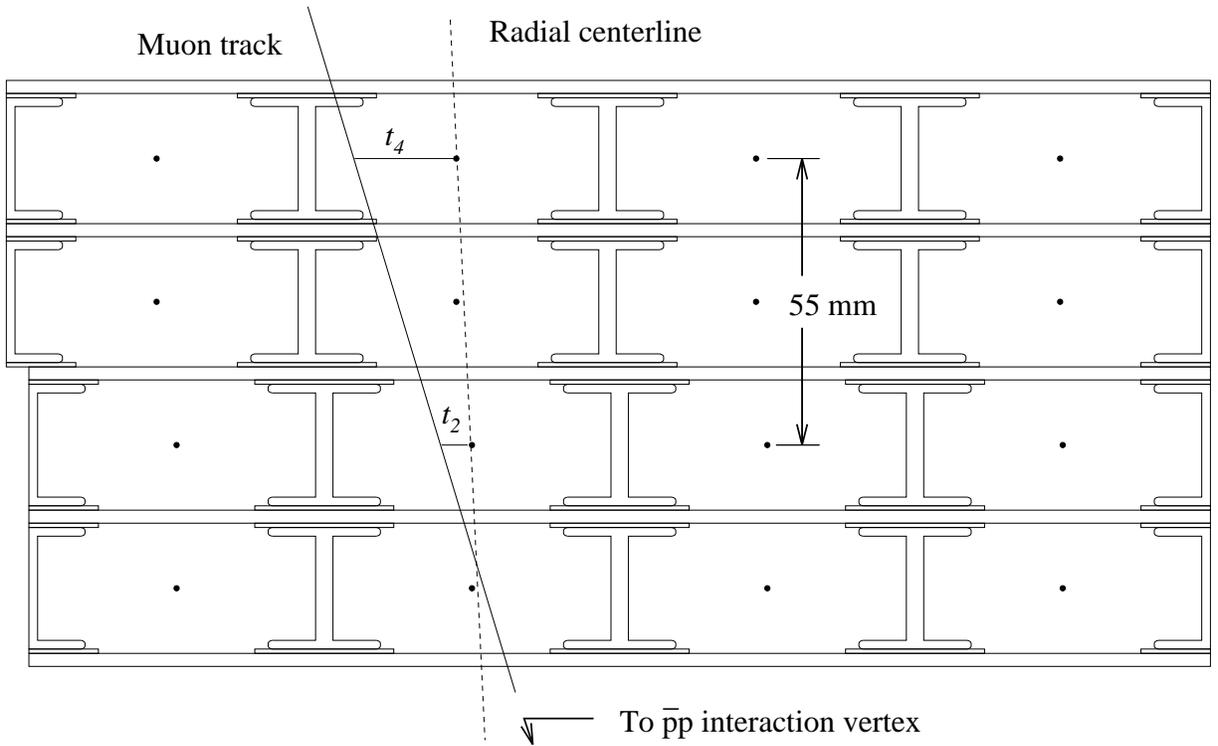
Figure 25: *End view of a CMU module, with 4 layers of 4 rectangular drift cells. Each cell has a 50 μm sense wire in the center.*

the CMP varies with azimuth angle due to the detector geometry that the chambers form around the CMU. A layer of scintillation counters (CSP) is also installed on the outside surface of the CMP. Basically, the CMP overlaps the region covered by CMU, but requiring additional CMP stub matching can further reject the background (the so-called hadron punch-through background[6]). As in the CMU, the CMP drift cells are rectangular, but with cross-sectional dimensions of 2.5 cm × 15 cm.

The central extension consists of conical sections of drift tubes (CMX) and scintillation counters (CSX), which are used to reject background based on timing information. It is located at each end of the central detector and extending in polar angle from 42° to 55° [40]. The CMX covers 240° in ϕ. There is a gap of the azimuth coverage of 30° at the top of the detector and is used for the solenoid refrigeration. The drift cells differ from those of the CMP only in length. They are arranged in azimuthal sections, as is the case in the central detector: in each of the 24 ϕ sectors, there are 4 layers of 12 drift cells. Adjacent layers are offset by one half cell in order to reduce ambiguities. The detailed design parameters of the central muon detectors are given in Table 9.

The Intermediate muon detector (IMU) is designed to trigger on muons with $|\eta| < \simeq 1.5$ and to identify off-line muons with $|\eta| < \simeq 2.0$. The IMU has almost the same configuration as the CMX/CSX, a set of drift tubes with four layers, located behind the plug calorimeter (6.2-20λ of steel). The forward muon toroidal magnet used in Run I[7] is re-used as the absorber without energizing. Additionally, a ring of steel $\simeq 60$ cm^2 is welded onto the inner face of the toroids. The detailed design parameters of the IMU are given in table 9.

A muon tower consists of 4 drift tubes, with successively larger radii. The drift tubes are filled with Argon-Ethane. As in the COT, when a particle passes through, it leaves a trail of ionization. The electrons liberated during the ionization drift toward the sense wire at high voltage, causing an avalanche. Pulses generated by the avalanches are sent to amplifier-shaper-discriminators in the collision hall, which in turn generate differential digital pulses. TDC's (Time to Digital Converters) give hit times for the pulses. The hit times are used to reconstruct short tracks, referred to as "stubs".

---

[6]Pions and kaons which survive passage through the calorimeter are a source of non-muon background, and are referred to as "punch through".

[7]The position of the toroidal magnet is moved 5.5 m closer to the interaction point.

Table 6: *Design parameters of the CDF II Muon Detectors. Pion interaction lengths and multiple scattering are computed at a reference angle of $\theta = 90°$ in CMU and CMP/CSP, at an angle of $\theta = 55°$ in CMX/CSX, and show the range of values for the IMU.*

	CMU	CMP/CSP	CMX/CSX	IMU
Pseudorapidity coverage	$\lvert\eta\rvert \leq 0.6$	$\lvert\eta\rvert \leq 0.6$	$0.6 \leq \lvert\eta\rvert \leq 1.0$	$1.0 \leq \lvert\eta\rvert \leq 1.5$
Drift tubes				
thickness	2.68 cm	2.5 cm	2.5 cm	2.5 cm
width	6.35 cm	15 cm	15 cm	8.4 cm
length	226 cm	640 cm	180 cm	363 cm
Max drift time	800 ns	1.4 μs	1.4 μs	800 ns
Total drift tubes	2304	1076	2208	1728
Scintillators				
thickness	-	2.5 cm	1.5 cm	1.5 cm
width	-	30 cm	30 - 40 cm	17 cm
length	-	320 cm	180 cm	180 cm
Total counters	-	269	324	864
Pion interaction lengths	5.5λ	7.8λ	6.2λ	6.2-20λ
Min. muon p_T	1.4 GeV	2.2 GeV	1.4 GeV	1.4-2.0 GeV
Multiple scatt. res.				
[cm/p (GeV/p)]	12	15	13	13-25

Muon stubs are matched to tracks in the drift chamber. Tracks are extrapolated to the muon chambers, and the distance between the track and stub is computed. To obtain quality muons, we place an upper limit on the value of the quantity χ^2_ϕ. This is the χ^2 of the track-muon match. It is computed based in the distance between the track and stub, the difference in direction of the track and stub, and the covariance matrix of the track.

2.2.6 Time of Flight system

Particle Identification is done both in the COT, using dE/dx, and in the Time Of Flight system (TOF) [41, 42, 43]. This detector is situated between the solenoid and the COT, and measures the time between a beam crossing and a particle incident upon the detector itself. We haven't made use of it in this particular analysis because we did not want to loose in efficiency. The TOF detector consists of scintillator panels which provide both timing and amplitude information. The timing resolution is 100 ps. The detector covers the central region out to $|\eta| < 1.1$ and will be capable of identifying Kaons from Pions by their flight time difference with at least 2σ separation up to Kaon momenta of 1.6 GeV/c. The major motivation for this system is to improve the b flavor tagging power, which is essential for B mixing and CP violation analyses, and will be used in the near future once those analyses are in a more advanced state.

2.2.7 Cherenkov Luminosity Counter

In Run II, a Cherenkov Luminosity Counter (CLC) has been built to measure the luminosity. It is designed to measure Run II luminosity with an uncertainty of 5% or less in the very high rate regime of $L \simeq 2 \times 10^{32} cm^{-2} s^{-1}$ [44]. There are CLC modules in the CDF detector, installed at small angle in the proton (East) and antiproton (West) directions with rapidity coverage between 3.75 and 4.75. Each module consists of 48 thin, long, gas-filled, Cherenkov counters. The counters consist of three concentric layers around the beam-pipe, each layer has 16 counters and points back to the collision point. In this analysis, we make use of the CLC to determine the luminosity of the data sample used.

2.2.8   Trigger and data acquisition systems

At CDF, proton and antiproton bunches cross at a rate of 2.5 MHz. As a result, the $p\bar{p}$ collision rate is much higher than the rate at which the data can be recorded. On the other hand, the rate of physics events relevant to any given analysis is, typically, only a very small fraction of the total collision rate. In our analysis, for example, we are only interested in events which contain a J/ψ particle in the final state. In order to select the most interesting events and reject the large amount of inelastic background [45], CDF has a three level architecture trigger system, which is briefly described in this section. Each trigger level provides a rate reduction sufficient to allow for processing in the next level with minimal deadtime. The data size is reduced in accordance to the triggering ability. Figure 26 shows the flow of the trigger and data acquisition (DAQ) systems.

Due to the raw collision rate, the trigger must have a large rejection factor while maintaining high efficiency for the broad range of physics topics. Figure 27 shows the block diagram of the trigger system for Run II.

The lowest level trigger, L1 uses output from the muon detectors, for muon triggers, and from all the calorimeters for electron and jet triggers.

Already at L1, COT information is used to reconstruct tracks. This is done by the XFT (eXtremely Fast Tracker) [46, 47]. Since this analysis begins with a sample of $J/\psi \to \mu^+\mu^-$, we are most concerned with muon stubs matched to tracks in the drift chamber. This is one of the handful of physics signatures which lead to L1 accepting an event. The muon triggers at L1 require then the presence of a track pointing at the muon systems. The track extrapolation to the muon detector is performed by the dedicated processor known as the XTRP [48].

It takes $\simeq 4$ μs to make a L1 decision. All front-end electronics are fully pipelined, with on-board buffering for 42 beam crossings. This allows the 41 next crossings to be stored while the L1 decision of crossing i is taken. Once the L1 decision is taken, the buffer i is released for the storage of another crossing.

The L1 trigger is a synchronous system with a decision reaching each front-end card at the end of a 42-crossing pipeline. It reduces the event rate from 2.5MHz to $\simeq$10kHz.

Dataflow of CDF "Deadtimeless" Trigger and DAQ

Detector

7.6 MHz Crossing rate
132 ns clock cycle

L1 Storage Pipeline:
42 Clock Cycles Deep

L1 trigger

Level1:
7.6 MHz Synchronous pipeline
5544ns latency
<50 kHz Accept rate

L1 Accept

L2 Buffers:
4 Events

L2 trigger

Level 2:
Asynchronous 2 stage pipeline
~20μs latency
300 Hz Accept Rate

L2 Accept

L1+L2 rejection: 20,000:1

DAQ Buffers

L3 Farm

Mass Storage

PJW 10/28/96
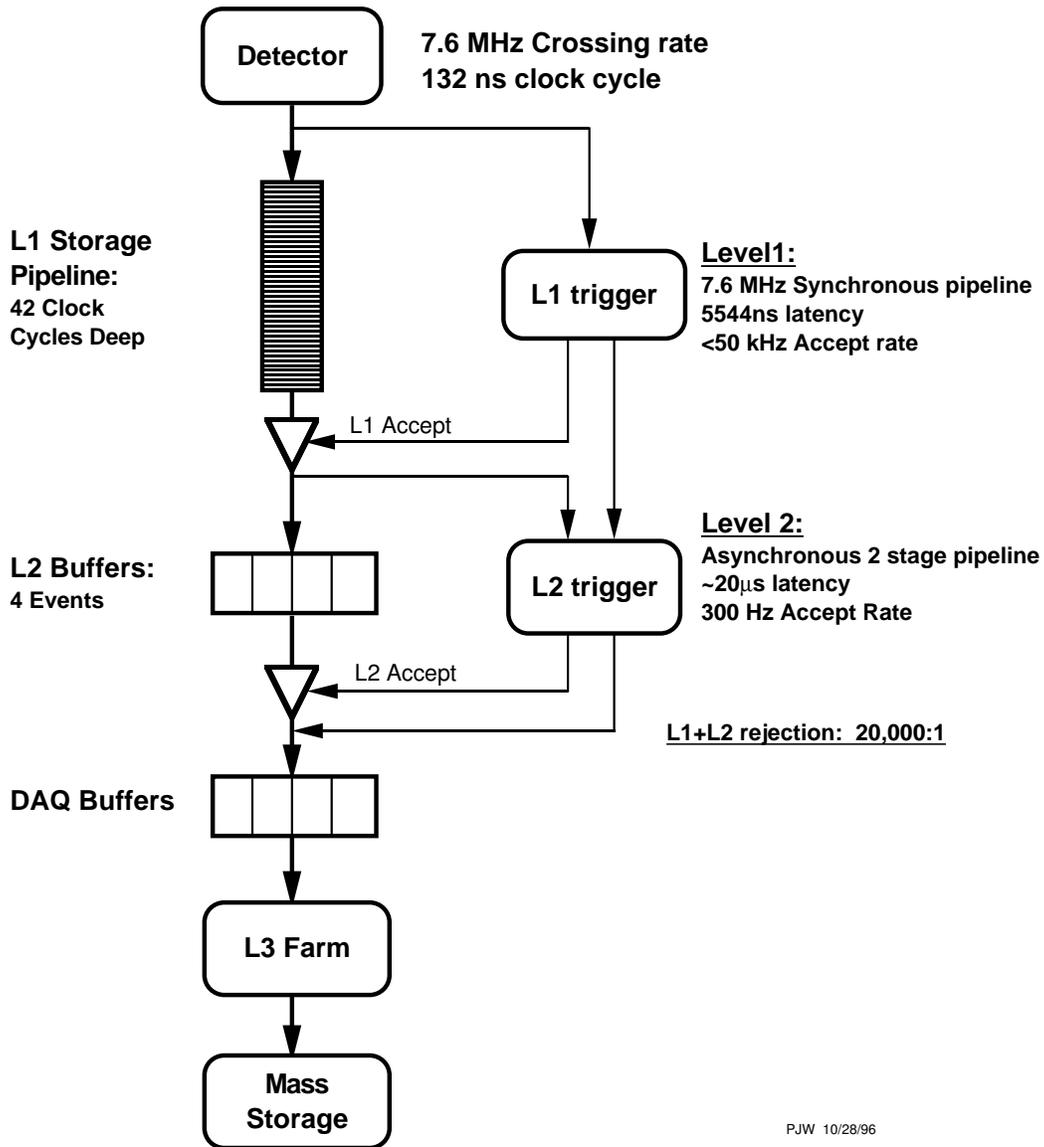
Figure 26: *The Run II readout functional block diagram of the three level pipelined and buffered trigger system of CDF.*

RUN II TRIGGER SYSTEM

Detector Elements
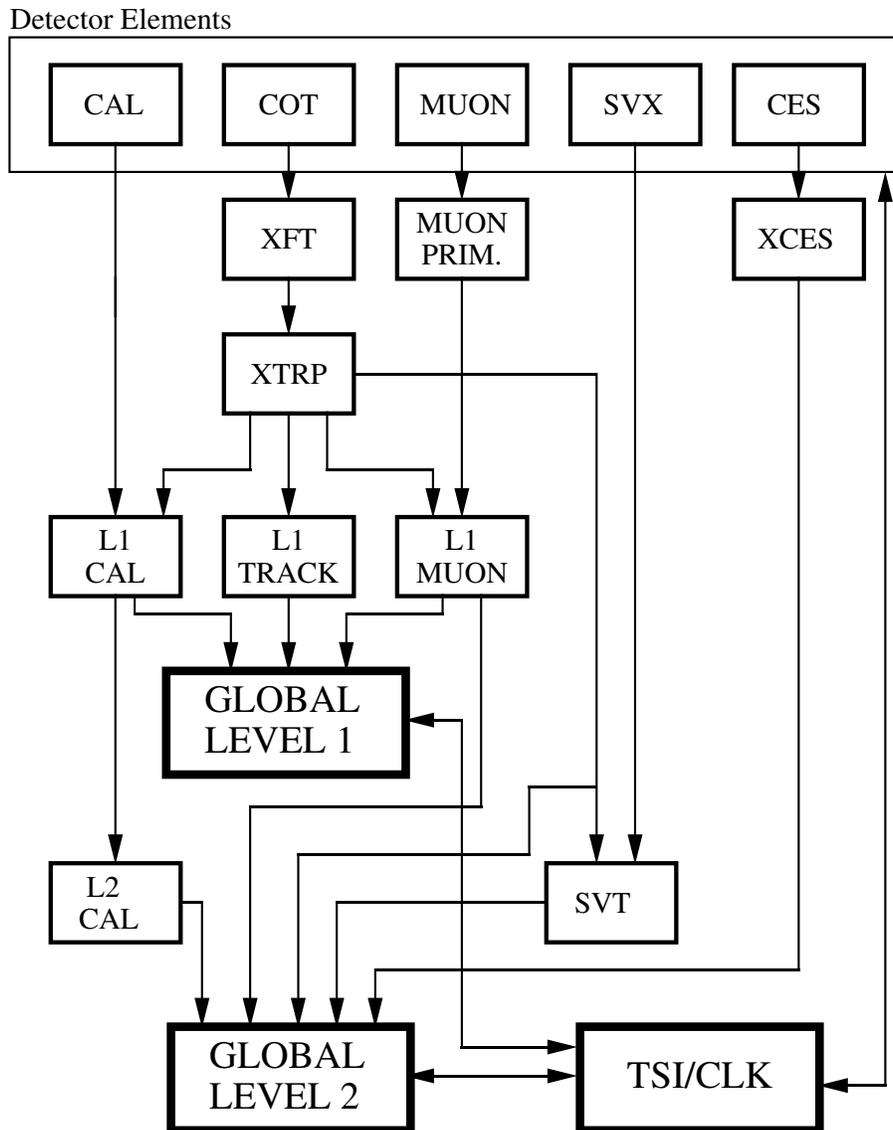
PJW 9/23/96

Figure 27: *Block diagram for the Run II trigger system.*

Upon a L1 accept, the data on each front-end card are transferred to one of four local L2 buffers. The second trigger level, L2, is an asynchronous system with an average decision time of 20 μs. The L2 trigger system has improved momentum resolution for tracks and finer angular matching between muon stubs and central tracks. It provides also the transverse mass of two muons.

The most challenging addition for L2 is the Silicon Vertex Trigger (SVT). The SVT provides the ability to trigger not only on electrons, muons, and jets, as during Run I, but to select events with tracks which have large impact parameters. This provides secondary vertex information already at L2 and opens a completely new window for physics measurements at hadron colliders, like high-statistics charm physics and the study of hadronic B decays. The L2 output rate is approximately 300 Hz. For this analysis, we do not make use of the SVT; our sample is selected strictly from the dimuon trigger.

All events accepted by the L2 trigger are collected in the dedicated processor known as the Event Builder (EVB), and then the EVB assembles those event fragments into one data block and delivers it to the third level of triggering, the L3. The L3 uses the C++ object oriented algorithms run in the "offline" reconstruction. The software is run on a PC farm. On one CPU, one event is processed in approximately 1 second. More complex quantities, such as dimuon invariant mass, are calculated here. The output rate of the L3 trigger is approximately 50 Hz at present. The accepted events are then electronically transferred to the *Feynman Computing Center* and stored on tape. To facilitate handling of the huge data volumes collected with the CDF, the data coming from L3 is currently split into eight different streams. For each event, several conditions are evaluated and encoded as "trigger bits"; these conditions determine the stream to which the event is sent; e.g., for events with hits in both CMU and CMP, which are compatible with the same track, the CMU-CMP trigger bit is set, and they are sent to "stream J", which is the data sample used in this analysis.

3.0 DATA SAMPLE AND EVENT RECONSTRUCTION

Going from raw detector data to usable lifetime information for our considered channels requires many processing steps. These include selection at the trigger level, event reconstruction, computation of relevant quantities and a final, channel-specific filtering. This chapter describes the processing of data that occurs after the trigger level (which was already covered in chapter 2).

For the sake of clarity, here we will not make use of the terminology needed for a full description of the exact procedure used to extract the lifetime ratio. Such a description would include a specification of the dataset name and version, the processing history including version numbers of the offline code used for both processing and for analysis, version numbers and/or history of corresponding constants from the conditions database, numbered ranges of data-taking run numbers, the time-dependent configuration of the trigger, and so forth. Version numbers would figure prominently in such description because the CDF offline code base is under continuous development, and database constants (*e.g.*, those describing the alignment of the silicon system components) are continually refined, and even event reconstruction is redone at certain times when significant improvements in reconstruction code have been made. Complete details on the procedures can, generally speaking, only be made precise in the jargon of the CDF experiment. They are left out of this treatment, and the reader is referred instead to the CDF internal note 6387 [49].

3.1 DATA SAMPLE

This analysis uses data collected by CDF from March 2002 to August 2003. Our particular data sample is usually referred to as "the J/ψ sample" because of the trigger requirements that events must pass in order to be included on it. We exclude the runs marked bad for either known detector problems (*e.g.*, COT high voltage) or trigger problems. Since a lifetime measurement requires that the position of the J/ψ vertex is well measured in the SVX II, we require SVX II information for the muons, which restricts us to the subsample of the data where the SVX II was fully functioning. This was not always true during the earlier running periods; as a result, the sample we used comprises approximately 195 pb^{-1} of integrated luminosity.

The J/ψ candidates will be combined with other particle candidates depending on the particular decay channel we seek for. In this analysis we study the following channels:

$B^+ \rightarrow J/\Psi K^+$ (+c.c.[1])

$B^+ \rightarrow J/\Psi K^{*+}, K^{*+} \rightarrow \pi^+ K_s^0$

$B^0 \rightarrow J/\Psi K^{*0}, K^{*0} \rightarrow K^+ \pi^-$

$B^0 \rightarrow J/\Psi K_s^0, K_s^0 \rightarrow \pi^+ \pi^-$

The topology of each decay channel is different and consequently needs a particular treatment for its reconstruction. This will be discussed later.

3.2 CDF ONLINE RECONSTRUCTION

3.2.1 Muon identification

Muon identification at CDF is based on two facts: a muon is a charged particle which leaves a track in the tracking system, and a muon can penetrate material more easily than all other charged particles. Because of the latter fact, a muon can traverse the calorimeter and reach the muon system to create hits in the muon chambers. The minimum detectable muon P_T is 1.4 GeV for CMU and 2.2 GeV for CMP (see table 9). The track left by a muon in

---

[1]The charge conjugate should be accounted for in all cases.

the tracking system can be extrapolated into the CMU muon chamber and be compared with the position of the track reconstructed in the muon chamber. A muon hit cluster in the muon chamber may match to more than one track in the tracking system. So the best matched (minimum χ^2) is selected as the muon candidate. A matching variable, the $r\phi$ position difference $|\Delta X|$, is required to be less than 15 cm at the trigger level. In order to reduce the hadron punch through background, a further requirement of the match between the track and the CMP muon chamber is applied offline to the muon candidate, $|\Delta X| < 10$ cm. Furthermore, the χ^2 of the position match between track segments in the CMU and CMP and the extrapolated track is required to be less than 9 in the $r\phi$ plane offline. In other words, the extrapolated track is required to be within 3σ of a muon hit.

3.2.2 Trigger requirements

The J/ψ sample is defined by specific trigger requirements at L1 and L3. In this data sample, any event which passes the L1 J/ψ trigger is automatically accepted by L2. To be considered a J/ψ candidate at L1, there must be either two muon stubs in the CMU, or one muon stub in the CMU and one in the CMX. The muon stubs must be matched to an XFT track. The XTRP [48] extrapolates the XFT tracks into the muon chambers. Taking into account multiple scattering and alignment corrections, a maximum $\delta\phi$ between the track and muon stub is determined. In the CMU, the XFT tracks are required to have $p_T > 1.5$ GeV. Stubs in the CMX must match an XFT track with $p_T > 2.0$ GeV.

As noted above, if an event passes the L1 J/ψ trigger, it is automatically accepted by L2, and passed on to L3. At L3, muon pairs are required to have opposite charge. Muon matching requirements are again enforced. The requirement is ΔX(track, stub) < 30 cm for CMU muons and ΔX(track, stub) < 50 cm for CMX muons. The two muon tracks are required to have $\Delta z < 5$ cm at the point of closest approach to the origin; since the error on z_0 for COT tracks is roughly half a centimeter, this is a very loose cut. Finally, for an event to be part of the J/ψ dataset, it is required that $2.7 < m_{\mu\mu} < 4.0$ GeV, where $m_{\mu\mu}$ is the invariant mass of the dimuon pair. To be part of the datasets used in this analysis, an event must have passed one of two sets of trigger conditions (encoded as trigger bits, and

usually called simply "triggers"). In the first set (that corresponds to earlier runs), L3 placed a requirement on the dimuon pair: $\Delta\phi(\mu, \mu)$ <2.25 radians. This was essentially dropped in later data by adding a new L3 trigger without this requirement. In the second set (used for, roughly, the more recent half of data used), two triggers were added that have a looser requirement on dimuon invariant mass ($2.7 < m_{\mu\mu} < 4.0$ GeV), but require that the vector sum of p_T for the dimuon pair be greater than 9 GeV. The various triggers include J/ψ's with both legs in the CMU, one in CMU and one in CMX, and one in the CMX with the other in both the CMU and CMP.

3.3   THE CDF OFFLINE RECONSTRUCTION

After data are accepted by the trigger and written to disk, they are copied to tape. A set of reconstruction routines is then run on the raw data, in a process that we will call "Basic Reconstruction" in this document, by the Farms Processing Systems [50], located in the *Feynman Computing Center*. During basic reconstruction, physics quantities required by most analyses are computed and stored; tracking, muon finding, primary vertex finding, calorimetry and jet clustering are a few examples. In this analysis we use the output of that reconstruction code as well as other reconstruction procedures (for the benefit of our particular studies).

3.3.1   Primary vertex reconstruction

Our lifetime measurements use the 2D vector difference between two points, one of them being the interaction point (IP), so we need to know this point to a good precision. Such knowledge can be obtained in two ways: using event-by-event information, or through the use of the beam position, which is obtained from an average over many events. The former can be subject to biases due to the use of displaced vertices in the computation of an event's primary vertex. In order to avoid that potential bias, we use the latter to obtain an estimate of the IP.

3.3.2   The beam line

The transverse profile of the luminous region inside the CDF detector can be described by Gaussian functions in x and y. The mean values of those Gaussians define the beam position. The width of the Gaussians varies along the length of the interaction region due to the focusing of the beams; nonetheless, it can be approximated by a Gaussian with a width of 30 cm.

The beam position is obtained on a run by run basis by fitting the event-by-event primary vertices (PV). The PV on each event is determined *via* a topological vertex finder based on a progressive vertex fit; selected tracks are added if they are consistent with the common vertex. The full process can be done using COT-only tracks or using tracks with information from the silicon detectors (L00, SVX II and ISL). Also, different versions of the software for basic track reconstruction will produce different beamline information. As a result, there are many versions of the beam positions, which are stored in a database. For this analysis, we use the beamline information obtained from the silicon tracks, and which is consistent with the version of the reconstruction code that we used.

To be used in the determination of a primary vertex based on the silicon detectors, a track needs to meet the following requirements: $p_T > 0.5$ GeV, $|d_0| < 3.0$ cm, "standard" COT track [2] with at least four ϕ silicon hits (at least five ϕ and two stereo silicon hits for stand-alone tracks).

From the distributions of the primary vertices one can directly derive the beam position and other beam parameters. Profile plots of the vertex positions in x over z and y over z are used to fit the beamline. Figure 28 shows the profile plots for a particular run.

During basic reconstruction, a special module produces profile plots with the primary vertices reconstructed using COT-only tracks and silicon tracks. Later on, a script combines the plots for the different streams, makes line fits through the data points, and enters the results of these beamline fits into the database.

Figure 29 shows the result of the above procedure; the x and y coordinates of the beam line, for all the runs used in this analysis. Notice how the beam has experienced variations

---

[2]By "standard" we mean a track with at least two stereo and two axial super layers with at least six hits.
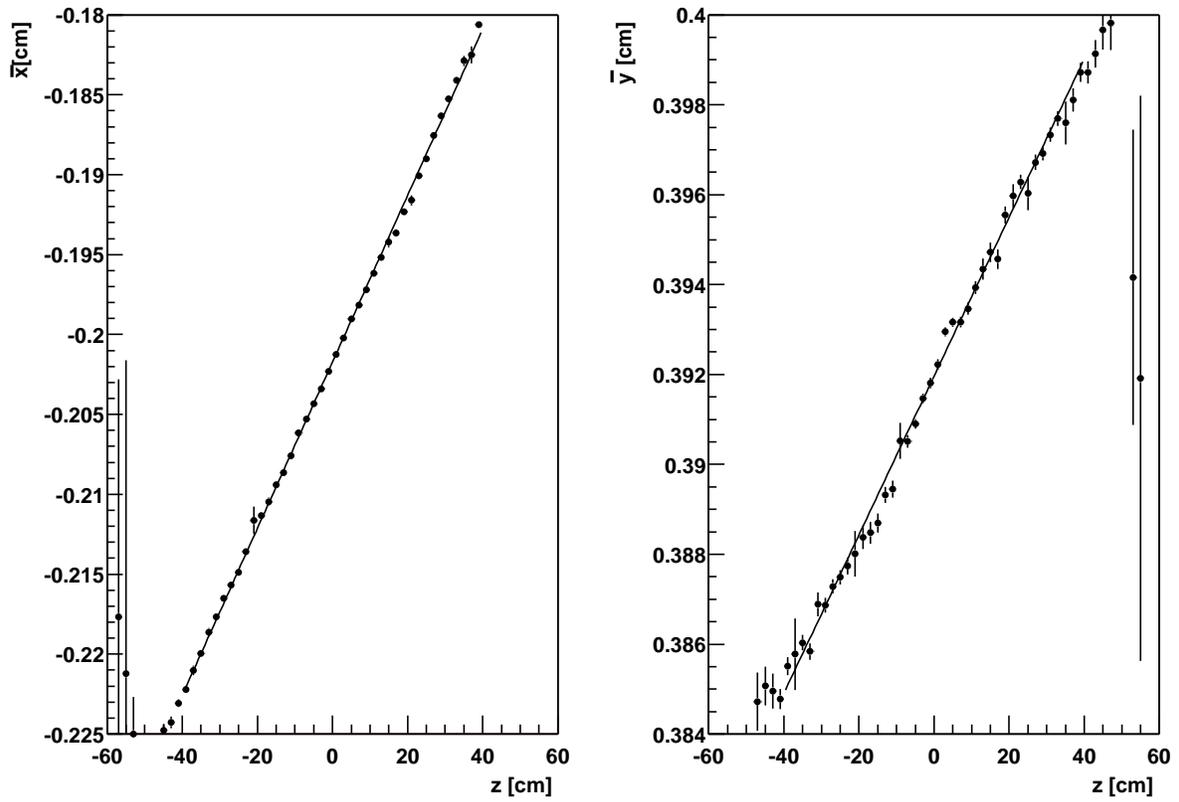
Figure 28: *Profile plots of the vertices reconstructed with silicon tracks for run 155130.*
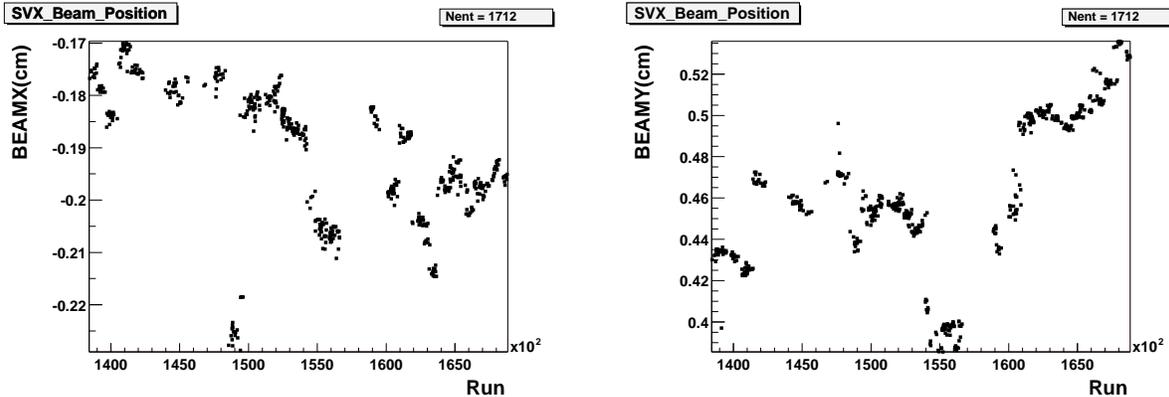
Figure 29: *x and y coordinates of the beam line for all the runs used in this analysis.*

along the running period[3]. Using the proper value (point) of the beam position in x and y for each particular run we obtain the d_0 (with respect to the beam position) versus ϕ_0 distribution shown in figure 30. As expected we observe a distribution with no dependence over ϕ_0, i.e., flat on ϕ_0. The spread of the distribution is ± 3 μm. This tells us that we know with such a precision the location of the beam along the different runs. Such a precision allows us to pursue our studies on the measurement of exclusive lifetimes of B hadrons.

3.3.3   Scaling of the COT error matrix

It has been shown that for COT tracks, the errors returned on the track parameters are underestimated, since the COT tracking code does not account for multiple scattering in the COT volume [51]. To correct for this, the covariance matrix of the COT parent track is scaled using p_T, θ-dependent scale factors from [52] shown in table 7. This scaling implements the parametrized effect of multiple scattering in the COT as measured in a GEANT Monte Carlo study.

---

[3]The movements experienced by the beam have reached values of more than 100 microns in several occasions. This happens most of the times during a shutdown, time when the detector is opened so that different repairs can take place.
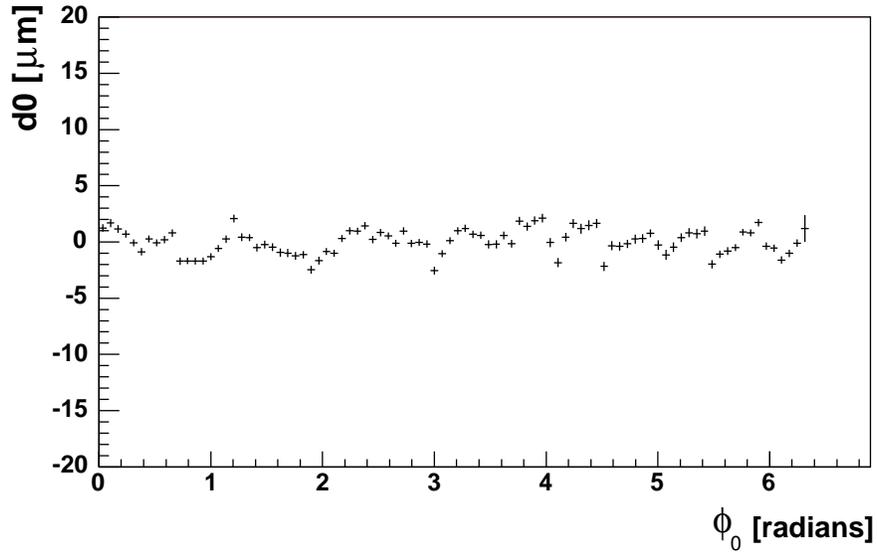
Figure 30: *Expected d_0 (with respect to the beam position) versus ϕ_0 distribution for the daughter tracks of J/ψ 's.*

Table 7: Track parameter error scaling factors.

Track parameter	Error scale factor
$cot\theta$	$\sqrt{1 - 0.580/(p_T^2 \cdot sin^3\theta)}$
z_0	$\sqrt{1 - 0.653/(p_T^2 \cdot sin^3\theta)}$
$curvature$	$\sqrt{1 - 5.33/p_T^2}$
d_0	$\sqrt{1 - 3.01/p_T^2}$
ϕ_0	$\sqrt{1 - 3.70/p_T^2}$

3.3.4 Track refitting

In order to properly account for the energy loss that a particle undergoes when going through the silicon active material (the silicon wafers) and the multiple scattering suffered when interacting with the silicon passive material (hybrids, port cards, cooling system, support frames, HDI cables, etc), all tracks used in this analysis are refitted.

The refitting process starts from the COT parent (with errors already scaled as described in the previous section) of the track[4] and puts back the silicon hits while correcting the track parameters for energy loss and multiple scattering effects in the silicon material.

Before refitting, we drop any silicon hit attached to tracks which are from Layer 00 or ISL, as well as all the SVX II stereo hits (both small angle and 90-degree stereo). The reason for this is that the L00 was still under commissioning while this analysis was developed, ISL alignment was not completely finished, and keeping the stereo hits does not directly contribute to the quality of the lifetime analysis. Besides, we do not lose any lifetime information by dropping the stereo hits, since we need only L_{xy} (the transverse decay length, see 3.4 for definition) and p_T (the transverse momentum) to compute the proper decay time of any particle.

Tracks forming a K_s^0 candidate need special treatment. Since the K_s^0's have a mean decay distance of 2.7 cm, their daughter particles often originate outside some silicon layers. Hence, hits in those inner layers could not have been produced by the particles, and should not be added to the track. On the other hand, during basic reconstruction, all tracks are treated in a generic way; basic reconstruction proceeds as if all tracks came from the beam or near the beam. As a result, wrong hits may be attached to tracks from K_s^0 daughters. Similar considerations apply with respect to the effect of the material traversed by the K_s^0 daughters. Multiple scattering and energy loss corrections should only be applied down to the radius in which the K_s^0 decayed (instead of all the way down to $r = 0$).

---

[4]If the track has no silicon hits, it is the track itself.

In view of the above, when refitting K_s^0 daughter tracks we do not use hits or material with a radius smaller than that of the estimated K_s^0 decay point. We obtain this point using the COT-only versions of the tracks, since tracks with incorrect silicon hits attached may bias the vertex constrained fit for the K_s^0.

The Kaon in the $B^+ \to J/\Psi K^+$ decay, and the Pion in the decay $K^{*\pm} \to \pi^\pm K^+$ are treated in the same way as muon tracks from J/ψ's regarding refitting, except that a Pion or Kaon mass hypothesis is assigned to the track, which makes a small difference in the energy loss correction.

3.3.5 $J/\psi \to \mu^+\mu^-$ reconstruction

We begin the reconstruction of J/ψ's by selecting muons with $p_T > 1.5\ GeV$. The muons can be of any type (CMU, CMP, CMUP, CMX). We select only muons whose corresponding tracks have 3 or more $r - \phi$ hits in the SVX II. We also impose track quality requirements by requiring a minimum number of COT axial (20) and stereo (16) hits on the muon tracks.

The dimuon invariant mass is determined using a general kinematic fitter ([53]), which determines the position of one or more vertices in a decay chain, plus the momenta of particles from those vertices, from two or more measured tracks. The two muons are constrained to come from a common vertex. The dimuon invariant mass distribution is shown in figure 31 together with p_T, η and the distribution of $Prob(\chi^2)$ for signal (white histogram) and background (gray). The signal is assumed to have a mass within 0.08 GeV of the PDG value (when fitted to a single Gaussian, the observed peak has a width of about 16 MeV; which comes entirely from detector resolution and not from the natural width of the particle. The background for these plots is taken from the sidebands of the J/ψ mass peak. The region on the left side of the peak, right below the signal region, is affected by the radiative tails; i.e., it contains J/ψ's with muons that have emitted a photon. For that reason, such region is not used as a sideband.

A perfect distribution of $Prob(\chi^2)$ for signal is flat from zero to one. A large number of events in bins near zero indicates the presence of background. We do not apply a cut on the J/ψ vertex probability; instead, we use a cut on the B vertex probability, which is highly correlated with the former. The p_T distribution for background peaks at lower values and the η distribution is more concentrated at 0.

3.3.6 $K_s^0 \rightarrow \pi^+\pi^-$ reconstruction

To search for K_s^0, we examine pairs of oppositely charged tracks in J/ψ events to find a secondary vertex. We refer to these tracks as daughter tracks. The daughter tracks should neither be associated with the primary vertex nor be the muon tracks. Track pairs which do not intersect when projected into the transverse plane are rejected. For each remaining track pair, we obtain the invariant mass by calculating the momenta of the individual tracks at their intersection point and adding them. In order to reduce background we require that the transverse momenta of the K_s^0 daughter tracks be greater than 0.4 GeV.

The L_{xy} (the transverse decay length) of the K_s^0 candidate, represents its distance to the J/ψ vertex (i.e., the intersection point formed by the two muons, since the J/ψ decays immediately) in the transverse plane, and is also obtained. K_s^0 candidates with an invariant mass outside a 20 MeV window around the PDG K_s^0 mass value or with an L_{xy} below 0.4 cm are removed.

Since the K_s^0 have a mean decay distance of 2.7 cm, they can decay in between or even outside the SVX II. For this reason no requirement is set on the silicon content of the K_s^0 daughter tracks but those hits at radii smaller than that of the intersection point are dropped (see section 3.3.4). Figure 32 shows the K_s^0 invariant mass distribution as it comes from the J/ψ sample with no specific K_s^0 reconstruction cuts applied other than a valid K_s^0 vertex.

3.4   RECONSTRUCTION OF EXCLUSIVE MODES

Since we are interested in the ratio of lifetimes, and in order to take advantage of the similarities among the decay modes considered, an effort was made to use common reconstruction
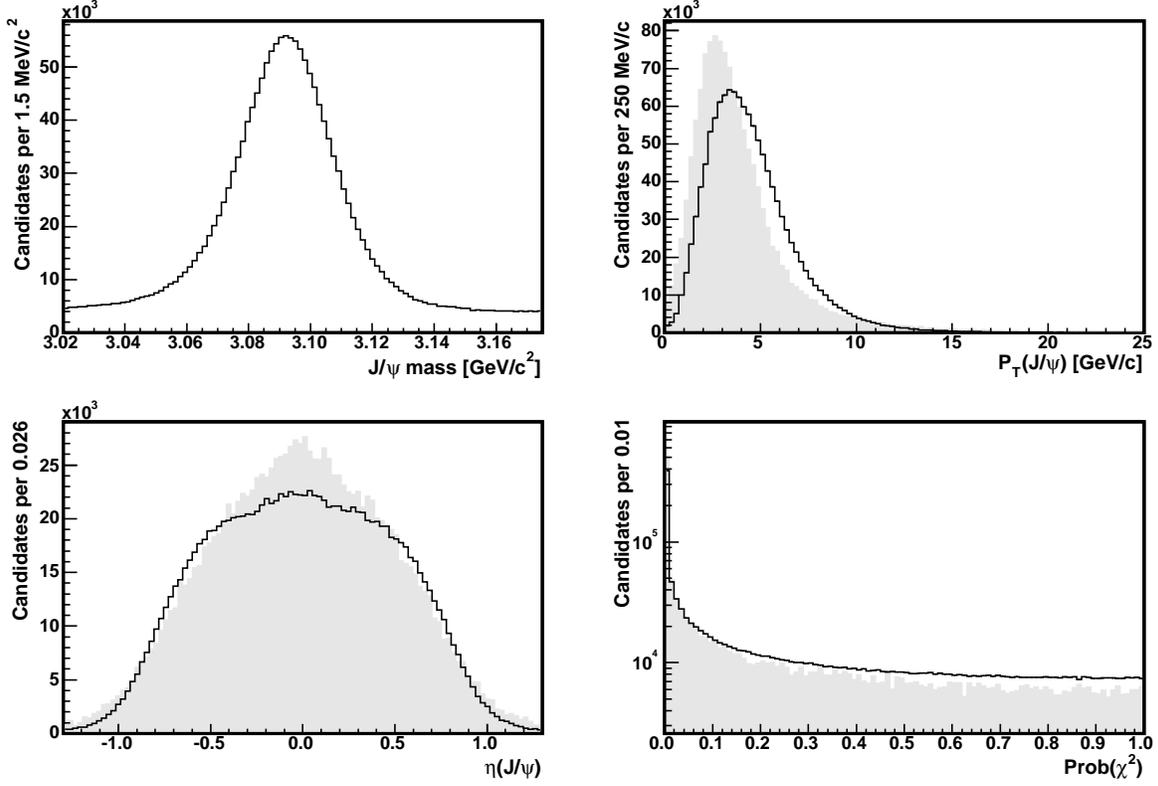
Figure 31: J/ψ *mass distribution from the* J/ψ *sample (top left)* P_T (J/ψ) *(top right),* $\eta(J/\psi)$ *(bottom left), and distribution of* $Prob(\chi^2)$ *(bottom right) for the vertex constrained dimuon fit, for signal (white histograms) and background (gray). For the last three plots, the background histogram is normalized to the signal histogram.*
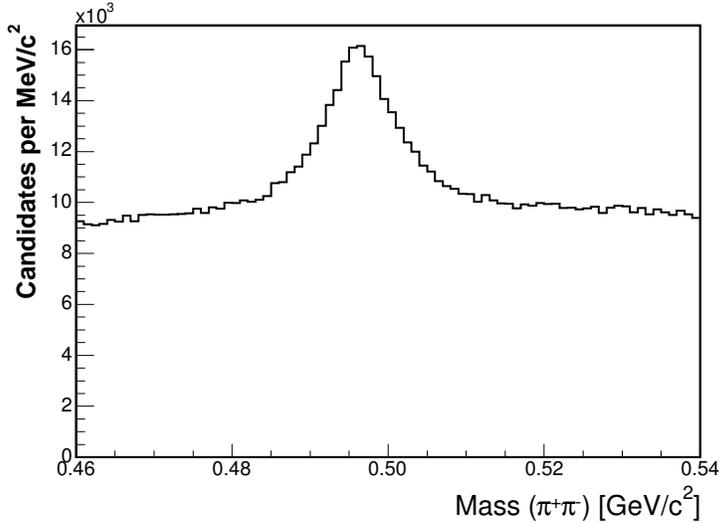
Figure 32: K^0_s *invariant mass distribution as it comes from the* J/ψ *sample.*

procedures whenever possible (beyond those imposed by the use of CDF's reconstruction platform, already described). Common procedures include the treatment of tracks, the selection of muons, the reconstruction of J/ψ candidates (used in all modes) and K^0_s candidates (for the two modes that include them), mass windows for kaons, *etc.* In this section, these common procedures are described, followed by the channel-specific cuts used for the analysis.

The B meson decay length L_{xy} and the relative proper decay time $c\tau$ are obtained in two ways; first, using all tracks to estimate the secondary vertex position; second, using only the two muon tracks from the J/ψ to estimate the secondary vertex position. In the following, we will refer to the first method as decay length or lifetime "*from the B vertex*" and the second method as "*from the J/\psi vertex*". The decay length L_{xy} is defined as

$$L_{xy} = \frac{\mathbf{V} \cdot P_T}{|P_T|} \ ,$$

where $\mathbf{V}$ is the vector going from the primary vertex to the secondary vertex. In both cases, P_T is computed using all the tracks that form the B vertex. The proper decay time is computed as:

$$c\tau = \frac{M L_{xy}}{P_T} \ .$$

72

3.4.1 Common reconstruction cuts

The following common cuts and procedures were applied:

Every track used was required to have a minimum of 20 axial hits and 16 stereo hits. In order to ensure the quality of lifetime information, the two muon tracks used for the J/ψ vertex are required to have at least 3 $r\phi$ hits. We also require 3 $r\phi$ hits for the Kaon in $B^+ \to J/\psi\, K^+$ and the Pion in the decay $K^{*+} \to \pi^+ K_s^0$. Since we are using the J/ψ vertex for the determination of the lifetime, these extra requirements do not affect significantly the quality of the vertices; however, they do reduce the amount of background in these channels.

As mentioned before, muons that left hits in the CMU detector are required to have a maximum χ^2 (as described in section 3.2.1) of 9. Since we want to use all suitable candidates, muons from all available muons are used, regardless of the detector that identified them (CMU, CMX, CMP).

The transverse momentum of each muon track considered for the J/ψ candidate is required to be over $1.5\,\mathrm{GeV}$. Since the cutoff due to the rangeout of the calorimeter is already 1.4 GeV, this cut does not significantly affect the sample (hardly any candidate is dropped that would have passed the rest of the cuts) but we keep it for definiteness. Only J/ψ candidates with masses within $80\,\mathrm{MeV}$ of the PDG value are considered.

Two of the modes considered include a K_s^0 ($J/\psi\, K_s^0$ and $J/\psi\, K^{*+}$). For both of them, a minimum P_T of $0.4\,\mathrm{MeV}$ is imposed on K_s^0 daughter tracks, assumed to be charged pions. The mass of K_s^0 candidates is required to be within $20\,\mathrm{MeV}$ of the PDG value. Finally, a considerable reduction on the combinatorial background can be achieved by a requirement on the minimum decay lenght of the K_s^0 (defined, as usual, as the distance from the B vertex to the K_s^0 vertex, projected into the direction of the K_s^0 momentum); in both channels, we set this minimum value to 0.4 cm.

Two modes include the reconstruction of a K^* as part of the decay. Also in this case, we consider only tracks with a P_T over $0.4\,\mathrm{GeV}$ as viable kaon daughters. K^* candidates are not considered if their mass is farther than $80\,\mathrm{MeV}$ from the corresponding PDG value ($0.892\,\mathrm{GeV}$). An extra cut on a mass-like quantity is needed in the case of the decay $B^0 \to J/\psi\, K^*$ to get rid of background coming from misreconstructed $B_s \to J/\psi\, \phi$: the cor-

responding K^* mass is required to be greater than 840 MeV. This will be addressed in more detail in section 3.4.2.

In all four channels, the overall vertex fits include only a J/ψ mass constraint (to the PDG value). The natural width of the K^* is large, so no mass constraint should be used in that case. For the K_s^0, we decided not to use a mass constraint based on both uniformity among the modes and the fact that such constraint reduces the yield without a significant improvement elsewhere (via the signal to noise ratio, or the width of the distribution). In all cases, a cut is used on the probability for the vertex fit chi^2, which needed to be greater than 0.1%.

For all modes, only candidates with a reconstructed B mass between 5.155 and 5.405 MeV/c^2 are used for the lifetime measurements. This window is chosen so that decays that are similar to the ones we are interested in, but for which a particle was not found, fall outside our allowed mass window.

3.4.2 Channel-specific reconstruction cuts

Although much of the reconstruction procedure allows a uniform treatment of all the decay modes considered, there are also a few channel-specific cuts that are needed. The need for channel-specific reconstruction cuts comes from the differences in the topology of the decay channels. The particular cuts used in the reconstruction of each decay channel are described below.

The z_0 of the K^+ track is required to be within 5 cm of the J/ψ vertex. The minimum transverse momenta considered are 2.0 GeV for the K^+ track and 6.0 GeV for the B^+ candidate.

A pointing angle cut (described in section 3.4.3) is required between the K_s^0 momentum and the vector going from the B vertex to the K_s^0 vertex. The transverse momentum cuts used for this specific channel are: $P_T(K^{*+}) > 2.5$ GeV, $P_T(B^+) > 7.5$ GeV.

We require the z_0 of the two K^* pions to be less than 5cm from each other; also, candidates are only considered if the z distance between the J/ψ vertex and the K^* vertex is less than 5 cm. When multiple candidates are found in one event, and they share tracks,
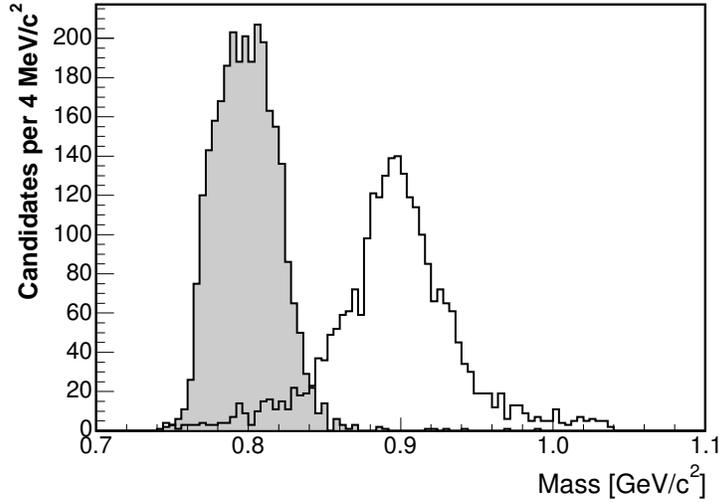
Figure 33: *Mass distribution for K^* candidates from $B_s \to J/\psi \, \phi$ (gray) and $B^0 \to J/\psi \, K^*$ (white), both reconstructed as $B^0 \to J/\psi \, K^*$ (Monte Carlo).*

only the one with a K^* mass closest to the PDG value is kept. The transverse momentum cuts are, in this case: $P_T(K^*) > 2.5$ GeV, $P_T(B^0) > 7.5$ GeV.

In this channel, we need to apply an extra cut on the mass of the K^* candidate in order to remove the possible contamination coming from misreconstructed B_s decays. When a B_s decays into $J/\psi \, \phi$, with $\phi \to K^+ K^-$, one of the Kaons can be misidentified as a pion; since they do form a vertex, and since there is also a J/ψ present in the event to be paired with this vertex, a candidate can be formed that, in principle, would be able to pass the reconstruction cuts. In order to determine if this was indeed the case, we simulated the aforementioned B_s decay and reconstructed the resulting events as $B^0 \to J/\psi \, K^*$. The reconstructed candidates do fall into our mass window. In order to avoid their inclusion in the sample, a cut can be applied on the mass of the K^* candidates. Figure 33 shows the comparison between the mass distributions of the K^* candidates from both samples. The white histogram corresponds to the K^* candidates coming from $B^0 \to J/\psi \, K^*$; the gray histogram corresponds to the misreconstructed $\phi \to K^+ K^-$. We set a cut value of 0.84 GeV on the mass of the K^*; only candidates above this value are considered; estimating from Montecarlo, this eliminates over 98% of the misreconstructed B_s^0 candidates.

A pointing angle cut (described in section 3.4.3) is required. The transverse momentum cuts are: $P_T(K_s^0) > 1.5$ GeV, $P_T(B^0) > 4.5$ GeV.

There was some concern in this case about the possible contamination coming from the misreconstruction of $\Lambda_b \to J/\psi\,\Lambda$, with $\Lambda \to p\pi$, if the proton track is used as a pion. Using realistic Monte Carlo, we found a relative efficiency of about 3% for the reconstruction of $J/\psi\,\Lambda$ as $J/\psi\,K_s^0$ after all our reconstruction cuts. This, together with the fact that (based on previous CDF Run II results) we expect the number of $J/\psi\,\Lambda$ to be roughly one third of the number of $J/\psi\,K_s^0$, bounds the contamination from this source to less than 1%. We take no further measures to eliminate these candidates.

The cuts on the transverse momenta of B^+, B^0, K^*, K^+, K_s^0, and $L_{xy}(K_s^0)$ were obtained via an optimization procedure (similar to the one described in [54]). The procedure maximizes the quantity $S/\sqrt{S+B}$, where S is the number of signal events and B the number of background events. The signal events were produced with a full simulation (see section 4.2.1), which included trigger emulation and complete GEANT representation of CDF. The background events were taken from the sidebands of the data mass distribution. For each mode, the maximum of $S/\sqrt{S+B}$ was found simultaneously with respect to all varying quantities.

3.4.3 Angle cut for reconstruction of K_s^0 channels

The reconstruction of exclusive states with a long-lived particle (K_s^0, Λ) usually includes a pointing constraint in the vertex fit, which forces the momentum of the long-lived particle to "point back" toward the vertex in which the particle was produced. Figure 34 shows a diagram of the situation, based on one of the decays we consider ($B^0 \to J/\psi\,K_s^0$). In general, measurement errors can cause the K_s^0 momentum (shown as a solid arrow in the figure) to not point back exactly toward the J/ψ vertex. However, the pointing will be better in the case of real $J/\psi\,K_s^0$ decays than for combinatorial background. We can take advantage of this by requiring the fit to force a perfect pointing, and then discarding the low probability candidates. This can be regarded as a replacement of the two pion tracks by a K_s^0 "pseudotrack" that is required to form a common vertex with the J/ψ muons.

The procedure has been used in the past [8, 55], and it allows a significant reduction of the background level. However, in order to apply the constraint, we first need to evaluate the quality of the K_s^0 pseudotrack, to verify that it has no potential to bias the measurement. We did this by studying inclusive K_s^0 candidates from the J/ψ sample, and found that pointing information from the K_s^0 pseudotracks is not well behaved, and therefore we do not apply the pointing constraint when fitting the B vertices. Appendix describes this study.

Once we drop the pointing constraint on the K_s^0 modes, the background level almost doubles, with a corresponding increase in the $c\tau$ statistical error. In order to reduce the amount of background, we applied an angle cut, being careful with its possible effect on the measurement. The pointing angle cut makes use of the angle α, in the transverse plane, between the K_s^0 momentum vector and the vector joining the J/ψ and the K_s^0 vertices.

The distribution of α for the selected $J/\psi\, K_s^0$ candidates is shown in figure 35. As expected, it peaks strongly at zero since, even when no constraint is applied, the badly pointing candidates still have a lower probability, and are hence removed from the sample; for the study of a pointing-based cut, a more convenient parameter is the negative logarithm of the angle, also shown in figure 35.

Figure 36 shows the behavior of the number of $J/\psi\, K_s^0$ signal candidates as a function of $-\log\alpha$. The lower plot shows the number of background candidates for each value of the cut. The signal is not affected by cut values below 4, while half of the background is removed.

To make sure that the use of this cut does not bias the fitted lifetime, we looked at the behavior of the fitted $c\tau$ as a function of cut value. Figure 37 shows the resulting lifetimes for the simultaneous and the non-simultaneous fit procedures (described in section 4.1). Given the errors, there is no evidence for a bias on the lifetime due to this cut. Moreover, cutting at $-\log\alpha > 4$ does not remove any signal events, so that there can be no lifetime bias induced by this cut. For our final sample, we select candidates for which $-\log\alpha > 4$ for channels $J/\psi\, K_s^0$ and $J/\psi\, K^{*+}$.
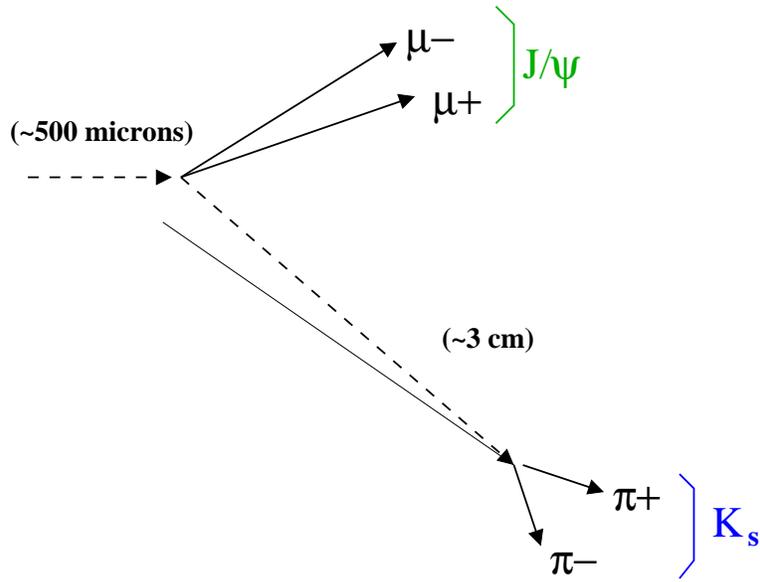
Figure 34: *In the decay $B^0 \rightarrow J/\psi\, K_s^0$, the K_s^0 momentum (shown as a solid arrow) does not point back exactly towards the K_s^0 production vertex. In principle, when fitting the B^0 vertex, a pointing constraint can be used to force this pointing.*
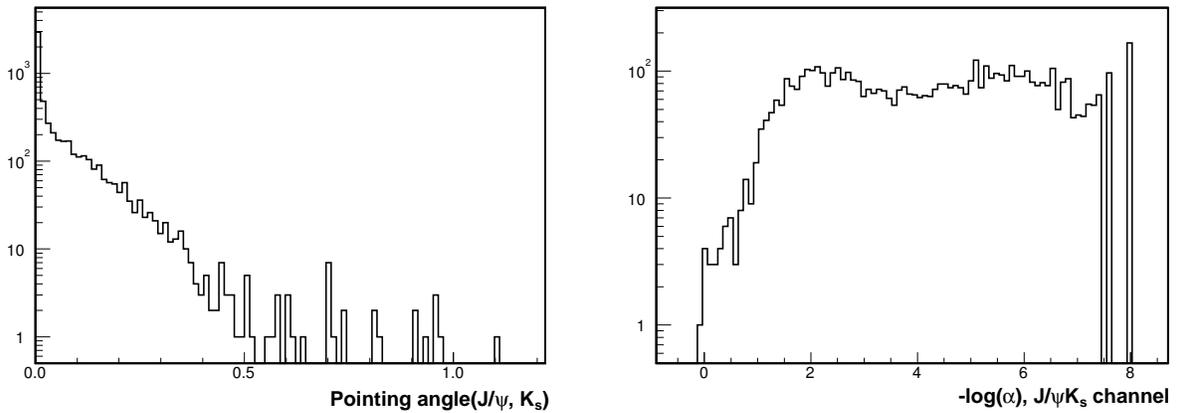
Figure 35: *Left: Pointing angle for the $J/\psi\, K_s^0$ candidates. Right: Negative logarithm of the pointing angle.*
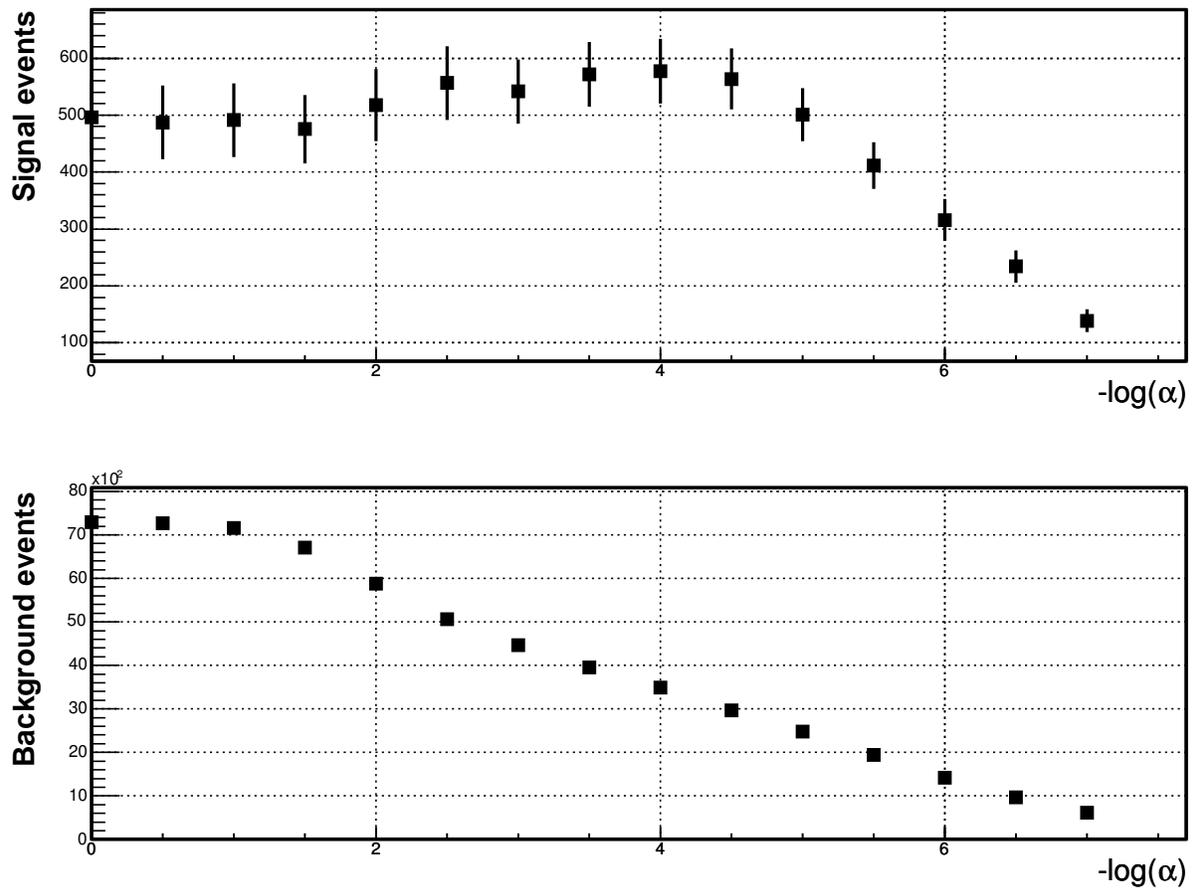
Figure 36: *Number of signal (up) and background (down) candidates as a function of the cut value for* $-\log \alpha$
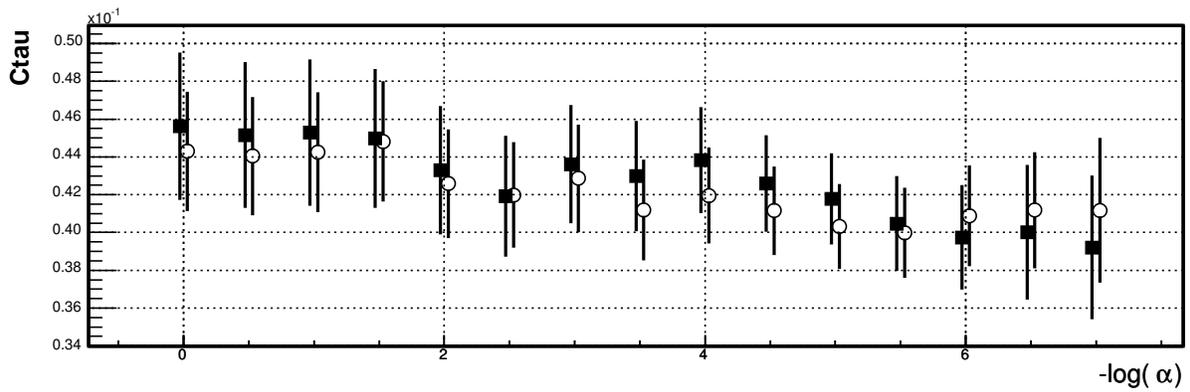
Figure 37: *Behavior of the fitted lifetime as a function of cut value for* $-log\alpha$. *Black squares: non-simultaneous fit; white circles: simultaneous mass-lifetime fit (nominal fit for the measurement). Only half-integer values of* $-\log\alpha$ *were used; the small separation among corresponding points is only to allow error bars to be read. Both fit methods are described in section 4.1.*

4.0   LIFETIME MEASUREMENT

From each B candidate, there are four quantities that we use for the extraction of the lifetime; they are the proper decay length, its estimated error, the mass, and the estimated error for the mass. As mentioned in section 3.4, the proper decay length is obtained from the 2d vector difference between the beamspot and the decay vertex. Roughly speaking, the distribution of this variable is a smeared exponential for the signal, and a Gaussian with exponential tails for the background. The mass distribution can be modeled by a Gaussian over a linear background. The errors in mass and proper decay distance are extracted from the vertex fitter. In this chapter, we model these distributions in order to obtain a measurement of the lifetime. The measurement of the lifetime ratio is derived from these measurements. The technique we use to extract the parameters is an unbinned maximum likelihood fit. This chapter describes the details of that fit, the studies we performed to validate the procedure, and the method used to combine the results of four decay channel fits into two meson lifetimes.

4.1   FITTING MODELS AND RESULTS

In this section we describe the fitting model used in extracting lifetimes. We use two techniques for lifetime extraction:

- **The "non-simultaneous" fit:** we analyze the B mass distribution in order to identify signal and sideband regions. We then divide the events up according to whether they lie in either of these two regions, and perform a simultaneous fit of the background shape of the $c\tau$ distribution in both regions, and of the B lifetime shape in the signal region. We call

this method "non-simultaneous" because the mass information is obtained independently of the lifetime fit.

- **The "simultaneous" fit:** We perform a simultaneous fit for mass and lifetime for all events in the B mass window.

In either case the lifetime is extracted using an unbinned maximum log-likelihood technique, incorporating event-by-event estimated errors. Floating error scale factors are added to the fit to model the underestimation of the errors. The first technique may be regarded as deriving from the second by replacing the continuous variable (mass) by a discrete variable taking three values: signal, sideband, or neither. It is of interest because it does not use event-per-event mass errors and is therefore independent of certain systematics. In the sections below we describe the fitting functions for each method in more detail.

4.1.1   Method 1: Non-simultaneous fit

In this technique, we first fit the mass spectrum of all candidates to a linear background plus a Gaussian distribution, using a binned likelihood fit. These fits are then used to define the signal and sideband regions based on the fitted mass, m_0 and its width σ_0. We use a mass window of $\pm 3\sigma_0$ around the fitted mass to define the signal region. The lower sideband region extends from 5155 MeV to $m_0 - 4\sigma_0$, while the upper sideband extends from $m_0 + 4\sigma_0$ to 5405 MeV.

In addition to defining the signal and sideband regions, the fit also provides an estimate of the signal fraction, f_1^0, and its error, σ_f, which are used in a constraint during the lifetime fit. The results of the mass fit for all the modes are displayed in Table 8.

The measured proper decay length $c\tau$ and its estimated error $\sigma_{c\tau}$ for events in the sideband region are next used in an unbinned maximum log likelihood fit to the B lifetime. A fraction of the events in this region, f_g, is assumed to come from prompt J/ψ production, and are modeled as a Gaussian distribution with a width described by the estimated error in $c\tau$. Systematic underestimation of $\sigma_{c\tau}$ is accommodated through an error scale factor s_l (*i.e.*, the "proper decay length error scale factor"; the subscript l stands for "length", we use it to distinguish this parameter from the scale factor for the mass errors, s_m, used in method

Table 8: *Mass fits for the four exclusive decay channels. Parameters m_0, σ_m, and N are, respectively, the mean, sigma and number of events of the signal peak (modeled as a single Gaussian). N_{bg} and S_{bg} are the number of events and the slope in the linear background model. The signal fraction f_1^0 for each mode is estimated from these fits.*

Param	$B^+ \to J/\psi\, K^+$	$B^+ \to J/\psi\, K^{*+}$	$B^0 \to J/\psi\, K^*$	$B^0 \to J/\psi\, K_s^0$
m_0 (MeV)	5278.4 ± 0.7	5282.2 ± 2.2	5278.6 ± 1.0	5278.0 ± 1.6
σ_m (MeV)	15.1 ± 0.8	14.3 ± 2.8	13.4 ± 1.0	17.3 ± 2.0
N	2091.3 ± 98.2	163.6 ± 26.1	894.2 ± 60.5	576.8 ± 56.8
N_{bg}	14237.7 ± 147.6	876.4 ± 37.5	6724.8 ± 97.5	3485.2 ± 78.3
S_{bg}	-0.8 ± 0.5	-9.6 ± 1.8	-2.8 ± 0.7	-0.8 ± 1.0
f_1^0	0.1281 ± 0.0059	0.1573 ± 0.0241	0.1174 ± 0.0078	0.1420 ± 0.0136

2). Of those events that remain, a fraction f_p is assumed to come from heavy flavor decays and modeled as a smeared exponential distribution for $c\tau > 0$ with a lifetime λ_{++}. Those that remain are assumed to come from pattern recognition errors; a fraction f_n is described by a negative smeared exponential tail with a lifetime of λ_-, and the remaining candidates in the sideband region are described by a second positive smeared exponential with an effective lifetime λ_+.

All exponential tails are smeared via their convolution with a Gaussian function whose width is determined by the event-per-event error $\sigma_{c\tau}$ as well as the global error scale factor s_l.

The likelihood function for the sideband region models only the background elements and is described by:

$$\mathcal{L}_{BG}(c\tau; \sigma_{c\tau}) = \frac{f_g}{\sqrt{2\pi} \cdot s_l \cdot \sigma_{c\tau}} e^{-(c\tau)^2/2(s_l\sigma_{c\tau})^2} + \int L(ct) \cdot \frac{1}{\sqrt{2\pi} \cdot s_l \cdot \sigma_{c\tau}} e^{-(c\tau - ct)^2/2(s_l\sigma_{c\tau})^2} d(ct)$$

(4.1)

where

$$L(c\tau) = (1 - f_g) \cdot \left[\frac{f_p}{\lambda_{++}} e^{-c\tau/\lambda_{++}} + (1 - f_p) \left(\frac{f_n}{\lambda_-} e^{c\tau/\lambda_-} + (1 - f_n) \frac{1}{\lambda_+} e^{-c\tau/\lambda_+} \right) \right]$$

(4.2)

The likelihood function for the signal region contains one component which is exactly that of the sideband region, together with another component which describes the signal events. The latter component is a smeared positive exponential:

$$\mathcal{L}_{SIG}(c\tau; \sigma_{c\tau}) = \int \frac{1}{c\tau_B} e^{-ct/c\tau_B} \cdot \frac{1}{\sqrt{2\pi} \cdot s_l \cdot \sigma_{c\tau}} e^{-(c\tau - ct)^2/2(s_l \sigma_{c\tau})^2} d(ct) \tag{4.3}$$

For events within the signal region, the likelihood is a linear superposition of $\mathcal{L}_{BG}$ and $\mathcal{L}_{SIG}$:

$$\mathcal{L} = f_1 \cdot \mathcal{L}_{SIG} + (1 - f_1) \cdot \mathcal{L}_{BG} \tag{4.4}$$

A Gaussian constraint term of the form $\exp\left[(f_1 - f_1^0)^2/2\sigma_f^2\right]$ is added to the likelihood function, using for f_1^0 and σ_f the values obtained from the fit to the mass distribution (which are shown in table 8). The $-\log\mathcal{L}$ is computed for all events in the signal and sideband regions and minimized with respect to the parameters in tables 9 and 10; the fitted parameter values and their statistical errors are shown in those tables for the two types of vertices (B vertex and J/ψ vertex). Figures 38 and 39 show the mass and lifetime distributions for all modes (histograms). The result of the binned likelihood fits (to a Gaussian plus a linear term) is overlaid on the mass peaks (solid black line); the $c\tau$ distributions show the unbinned likelihood fit results. The plots in the central column correspond to the signal region; on each, the solid black line depicts the full likelihood function (including both signal and background models), the signal model is shown filled in light gray, and the dotted line shows the expected distribution of the background in the signal region. The left column shows the $c\tau$ distribution for candidates in the sidebands (points), together with the fitted background model (dotted line).

4.1.2   Method 2: Simultaneous fit for B mass and lifetime

The second way in which we extract the B lifetime is using a simultaneous fit for B mass and lifetime. In this case, the sample is not separated into "signal" and "sideband" regions; instead, the overall probablity density depends on both $c\tau$ and m, and all candidates in the mass window are used. As a result, this method uses more data than method 1 (which does not use candidates in the regions going from $3\sigma_0$ to $4\sigma_0$ away from the fitted value of

Table 9: *Lifetime fit results for the four exclusive decay channels when the B vertex is used for the estimation of transverse decay length. The first four parameters (m_0, σ_m, N, P_1) are obtained from a binned likelihood fit to the mass distribution. The rest are the result of the unbinned maximum likelihood fit.*

Parameter	$B^+ \to J/\psi\, K^+$ value	error	$B^+ \to J/\psi\, K^{*+}$ value	error	$B^0 \to J/\psi\, K^*$ value	error	$B^0 \to J/\psi\, K_s^0$ value	error
m_0 (MeV)	5278.4	0.7	5282.3	2.2	5278.7	1.0	5277.9	1.6
σ_m (MeV)	15.0	0.8	14.2	2.8	13.6	1.0	17.4	2.0
N	2106.3	98.0	163.7	25.9	910.0	61.2	578.0	57.2
P_1	-0.8	0.5	-9.7	1.8	-2.8	0.7	-0.8	1.0
f_1	0.296	0.006	0.373	0.028	0.313	0.010	0.294	0.015
λ_{++} (μm)	517.3	66.2	1121.1	756.3	479.0	50.1	1001.0	340.5
λ_+ (μm)	114.3	16.7	234.3	35.1	81.5	13.0	268.0	51.7
λ_- (μm)	297.4	25.7	703.7	340.3	183.2	23.2	540.0	59.7
f_g	0.854	0.009	0.748	0.029	0.800	0.015	0.757	0.016
f_p	0.230	0.047	0.050	0.060	0.288	0.042	0.121	0.093
f_n	0.184	0.019	0.045	0.025	0.166	0.024	0.191	0.028
s_l	1.265	0.012	1.472	0.068	1.248	0.020	1.211	0.026
$c\tau_B$ (μm)	494.4	12.9	466.9	43.0	454.4	17.5	436.9	28.0

Table 10: *Lifetime fit results when the J/ψ vertex is used. The first four parameters (m_0, σ_m, N, P_1) are obtained from a binned likelihood fit to the mass distribution. The rest are the result of the unbinned maximum likelihood fit.*

Parameter	$B^+ \to J/\psi\, K^+$ value	error	$B^+ \to J/\psi\, K^{*+}$ value	error	$B^0 \to J/\psi\, K^*$ value	error	$B^0 \to J/\psi\, K^0_s$ value	error
m_0 (MeV)	5278.4	0.7	5282.2	2.2	5278.6	1.0	5278.0	1.6
σ_m (MeV)	15.1	0.80	14.3	2.8	13.4	1.0	17.2	2.0
N	2091.3	98.2	163.6	26.1	894.2	60.5	576.8	56.8
P_1	-0.8	0.5	-9.6	1.8	-2.8	0.7	-0.8	1.0
f_1	0.291	0.007	0.372	0.029	0.311	0.01	0.293	0.015
λ_{++} (μm)	421.16	45.42	1147.4	763.3	498.4	55.92	991.1	374.2
λ_+ (μm)	98.2	12.9	230.9	31.3	90.83	14.1	276.6	53.5
λ_- (μm)	261.5	22.3	439.9	131.66	127.6	14.9	530.4	58.71
f_g	0.825	0.01	0.712	0.029	0.77	0.017	0.761	0.016
f_p	0.255	0.044	0.042	0.049	0.25	0.041	0.118	0.1
f_n	0.185	0.02	0.084	0.027	0.218	0.03	0.199	0.03
s_l	1.204	0.012	1.209	0.054	1.179	0.019	1.188	0.025
$c\tau_B$ (μm)	501.66	13.3	468.8	43.5	459.7	18.4	438.3	28.0
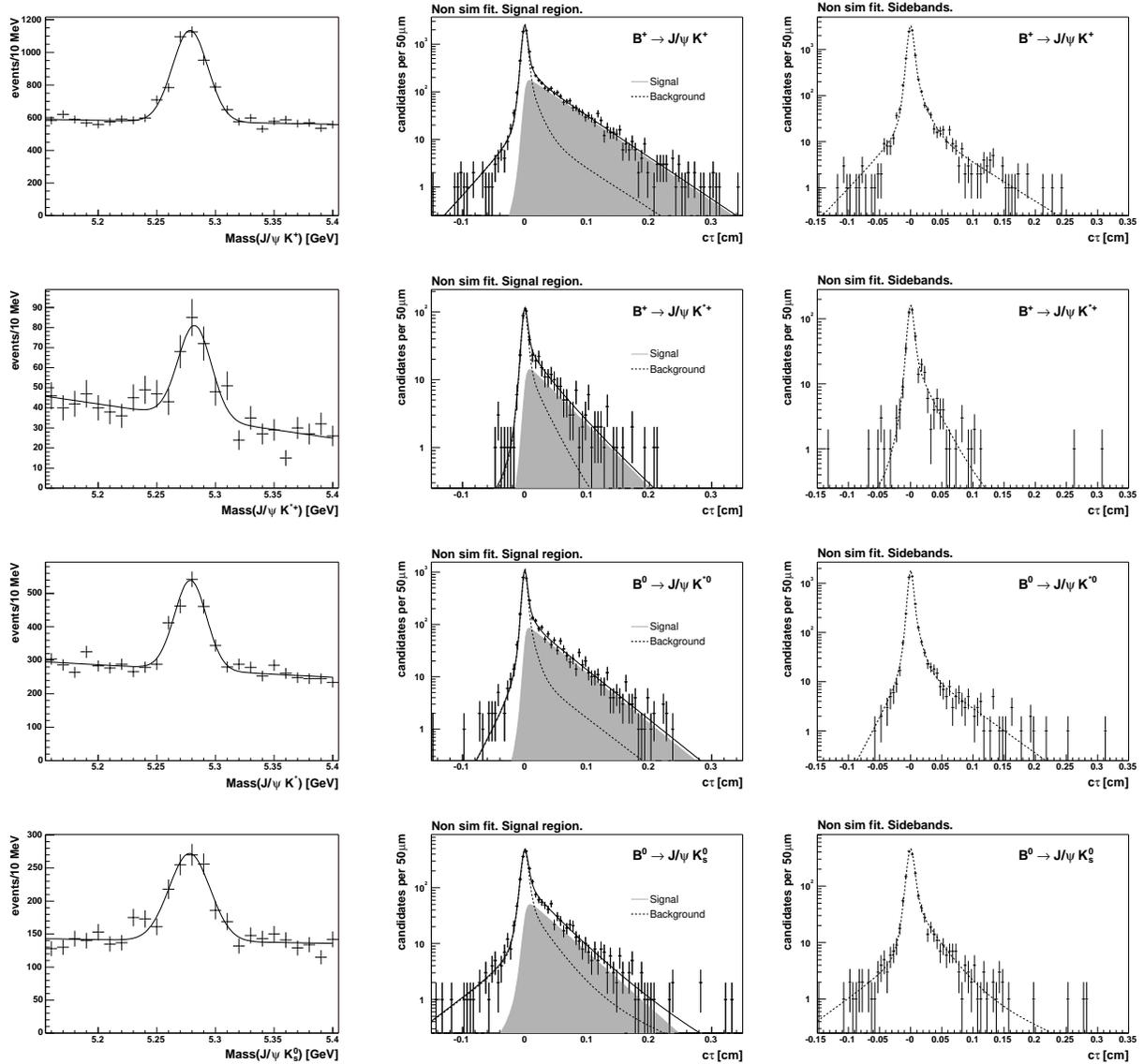
Figure 38:   *Results of the Method 1 fits for the four channels, using the B vertex. Rows (top to bottom): $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, and $B^0 \to J/\psi\, K_s^0$. Columns (left to right): fitted mass distributions, $c\tau$ distribution for the signal region, $c\tau$ distributions for the sidebands.*
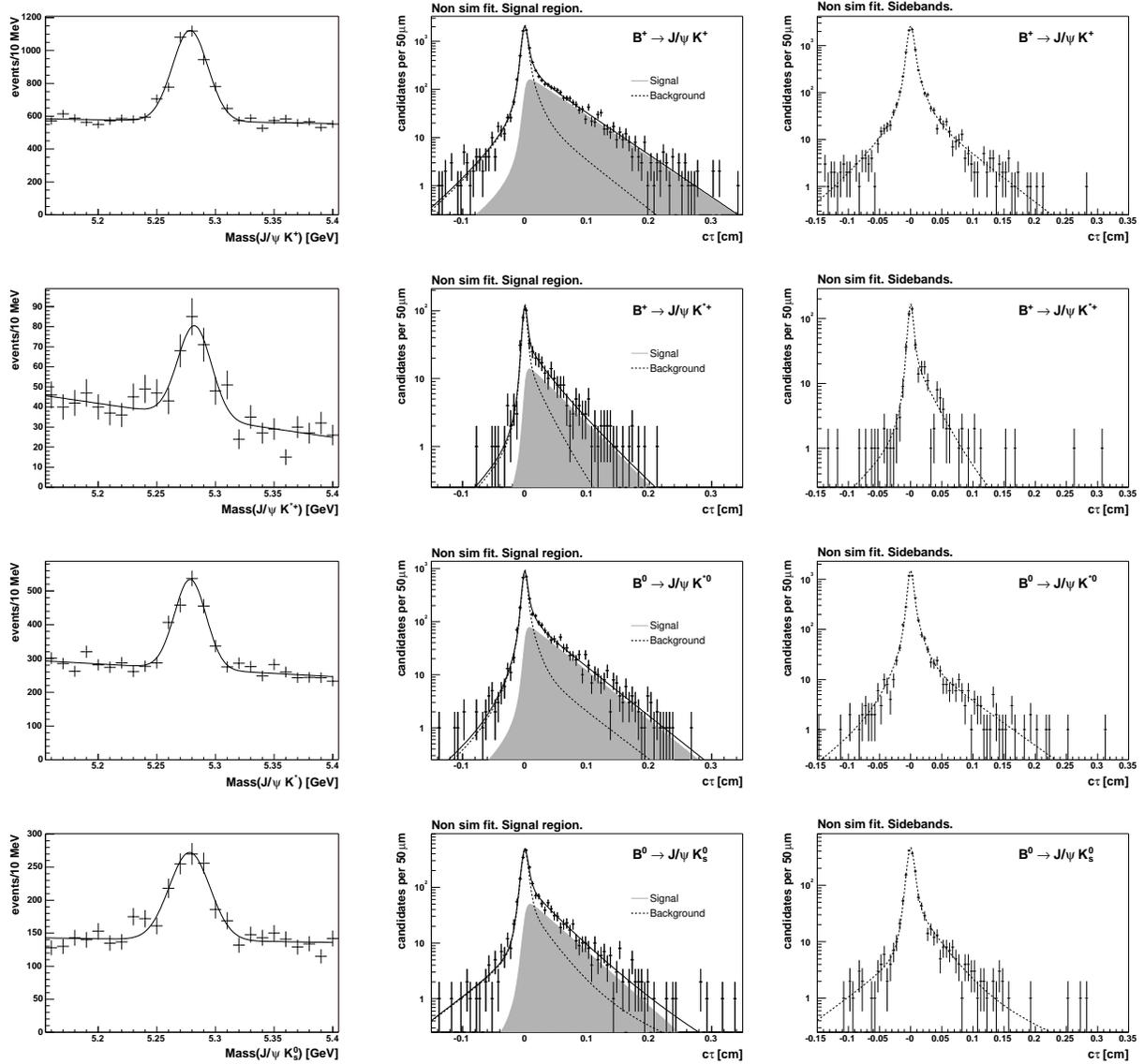
Figure 39: *Results of the Method 1 fits for the four channels, using the J/ψ vertex. Rows (top to bottom): $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, and $B^0 \to J/\psi\, K_s^0$. Columns (left to right): fitted mass distributions, cτ distribution for the signal region, cτ for the sidebands.*

the mass peak's mean), resulting in a smaller overall error on the measurement. Another difference with method 1 is that, in this case, we use event-per-event estimated errors on both $c\tau$ and m, and error scale factors for both lifetime and mass errors are included in the fit (method 1 uses only the $c\tau$ estimated errors, and hence only a single error scale factor can be determined in that case). We use the results of method 2 (applied to the J/ψ vertex information) as our nominal values; the results of method 1 are used for comparison, and to validate our estimation of systematic errors.

Instead of using two stages (one to determine how to divide the sample, and another for the lifetime fit), mass-dependent terms are incorporated into the overall fitting function; all parameters (including mass mean and width) are then obtained by minimization of the negative log-likelihood. We use a first order polynomial to describe the mass distribution of background events:

$$P(m) \equiv P_0 + P_1 \cdot (m - m_c) \tag{4.5}$$

(where m_c is the center of the mass window), and the function $\mathcal{L}_{BG}(c\tau; \sigma_{c\tau})$ discussed above to model the lifetime distribution. The combined likelihood term for background events $\mathcal{L}_{BG}^{(2)}(m, c\tau; \sigma_m, \sigma_{c\tau})$ is a direct product of these two functions:

$$\mathcal{L}_{BG}^{(2)}(m, c\tau; \sigma_m, \sigma_{c\tau}) = P(m) \cdot \mathcal{L}_{BG}(c\tau; \sigma_{c\tau}) \tag{4.6}$$

To model the bivariate mass and lifetime distribution of the signal events, the fitting function employs a direct product of two simple 1-D functions (a function of $c\tau$ and a function of m). The mass distribution of the signal is modelled as a Gaussian:

$$G(m; \sigma_m) \equiv \frac{1}{\sqrt{2\pi} s_m \sigma_m} e^{-(m-m_0)^2/2(s_m \sigma_m)^2}. \tag{4.7}$$

where m is the candidate's mass, σ_m is the corresponding estimated error (both quantities measured on a per-candidate basis), m_0 is the meson's mass, and s_m is the mass error scale factor (these last quantities are parameters to be determined by the minimization of $-\log\mathcal{L}$).

The combined likelihood term for signal events $\mathcal{L}_{SIG}^{(2)}(m, c\tau; \sigma_m, \sigma_{c\tau})$ is a direct product of these two functions:

$$\mathcal{L}_{SIG}^{(2)}(m, c\tau; \sigma_m, \sigma_{c\tau}) = G(m; \sigma_m) \cdot \mathcal{L}_{SIG}(c\tau; \sigma_{c\tau}). \tag{4.8}$$

Finally, the complete fitting function is a linear superposition of signal and background terms, with a coefficient f_1 (the *signal fraction*) that is determined in the unbinned likelihood fit:

$$\mathcal{L}^{(2)}(m, c\tau; \sigma_m, \sigma_{c\tau}) = f_1 \cdot \mathcal{L}^{(2)}_{BG}(m, c\tau; \sigma_m, \sigma_{c\tau}) + (1 - f_1) \cdot \mathcal{L}^{(2)}_{SIG}(m, c\tau; \sigma_m, \sigma_{c\tau}) \qquad (4.9)$$

The likelihood function is computed for all events in the mass window from 5115 MeV to 5405 MeV, and minimized with respect to the parameters in the Tables 11 and 12. Here, the parameter f_1 is not directly comparable to the same parameter in the previous section because it now refers to the signal fraction in the entire window, not just within the signal region.

Figures 40 and 41 show the result of method 2 when using the B vertex and the J/ψ vertex. The left column shows the mass histograms for all modes; the projection of the likelihood function is overlaid. The number of signal candidates shown in the plots is obtained from the signal fraction and the number of candidates. The column on the right shows the $c\tau$ distributions for all candidates on the mass window (histograms); on each plot, the solid black line shows the overall likelihood function, the signal component is shown filled in light gray, and the dotted line depicts the background. The procedure used to obtain the probabilities quoted on figure 41 is described in section 4.3.2.

4.2   CROSS CHECKS

4.2.1   Results on fully simulated Monte Carlo

As a crosscheck of the procedures used in the measurement of exclusive B lifetimes, we used CDF's full simulation, based upon GEANT, to simulate the production of B mesons and their decay to exclusive final states.

The Monte Carlo sample used for the expected signal was produced with an event generator ([56, 57]) with a cut of 3 GeV/c on the b quark and of 4 GeV/c on the P_T of the B meson. The quark was also cut in rapidity at 6. The angular and momentum distributions used were the NDE spectrum with effective b quark mass of 4.75 GeV. The B mesons were

Table 11: *Combined mass/lifetime fit results for the four exclusive B channels. Results use the B vertex. N denotes the number of signal candidates; it is obtained from the signal fraction and the number of candidates (i.e., it is not a parameter in the fit).*

	$B^+ \to J/\psi\, K^+$		$B^+ \to J/\psi\, K^{*+}$		$B^0 \to J/\psi\, K^*$		$B^0 \to J/\psi\, K^0_s$	
Parameter	value	error	value	error	value	error	value	error
m_0 (MeV)	5278.7	0.4	5278.7	1.4	5279.2	0.6	5277.2	0.9
s_m (MeV)	1.25	0.03	1.30	0.12	1.40	0.05	1.04	0.06
P_1	-0.9	0.5	-9.2	1.8	-2.9	0.7	-0.9	1.0
f_1	0.135	0.003	0.189	0.017	0.129	0.005	0.150	0.009
λ_{++} (μm)	477.1	51.6	263.5	99.4	461.1	41.8	2825.7	1073.6
λ_+ (μm)	107.0	14.8	143.4	132.1	77.7	13.1	373.6	30.6
λ_- (μm)	318.7	26.4	562.0	217.1	178.2	20.6	596.7	64.8
f_g	0.852	0.008	0.753	0.033	0.801	0.014	0.773	0.012
f_p	0.249	0.044	0.696	0.830	0.330	0.040	0.028	0.017
f_n	0.176	0.017	0.187	0.583	0.189	0.025	0.163	0.018
s_l	1.270	0.012	1.415	0.061	1.243	0.019	1.255	0.024
N	2216.0	57.4	197.1	17.9	991.9	38.5	607.5	35.5
$c\tau_B$ (μm)	493.7	12.2	498.0	41.2	453.3	16.9	419.1	25.3

Table 12: *Combined mass/lifetime fit results for the four exclusive B channels. Results use the J/ψ vertex. N denotes the number of signal candidates, as obtained from the signal fraction and the number of candidates.*

Parameter	$B^+ \to J/\psi\, K^+$ value	error	$B^+ \to J/\psi\, K^{*+}$ value	error	$B^0 \to J/\psi\, K^*$ value	error	$B^0 \to J/\psi\, K^0_s$ value	error
m_0 (MeV)	5278.5	0.4	5278.7	1.4	5279.1	0.6	5277.3	0.9
s_m (MeV)	1.24	0.03	1.29	0.12	1.41	0.06	1.04	0.06
P_1	-0.9	0.5	-9.1	1.8	-2.9	0.7	-0.8	1.0
f_1	0.133	0.004	0.190	0.017	0.127	0.005	0.149	0.009
λ_{++} (μm)	405.4	41.7	261.2	144.9	480.1	47.8	2896.3	1110.8
λ_+ (μm)	102.4	13.4	162.7	132.8	90.5	14.0	379.9	30.5
λ_- (μm)	284.4	23.3	415.4	118.0	129.8	13.9	580.4	61.9
f_g	0.830	0.009	0.724	0.032	0.772	0.016	0.774	0.012
f_p	0.264	0.047	0.603	0.761	0.279	0.040	0.027	0.016
f_n	0.180	0.019	0.214	0.568	0.233	0.029	0.171	0.018
s_l	1.206	0.011	1.213	0.052	1.176	0.018	1.222	0.023
N	2165.7	58.6	197.5	18.0	966.3	39.5	603.4	35.4
$c\tau_B$ (μm)	499.3	12.6	500.4	41.6	459.9	17.8	419.6	25.3
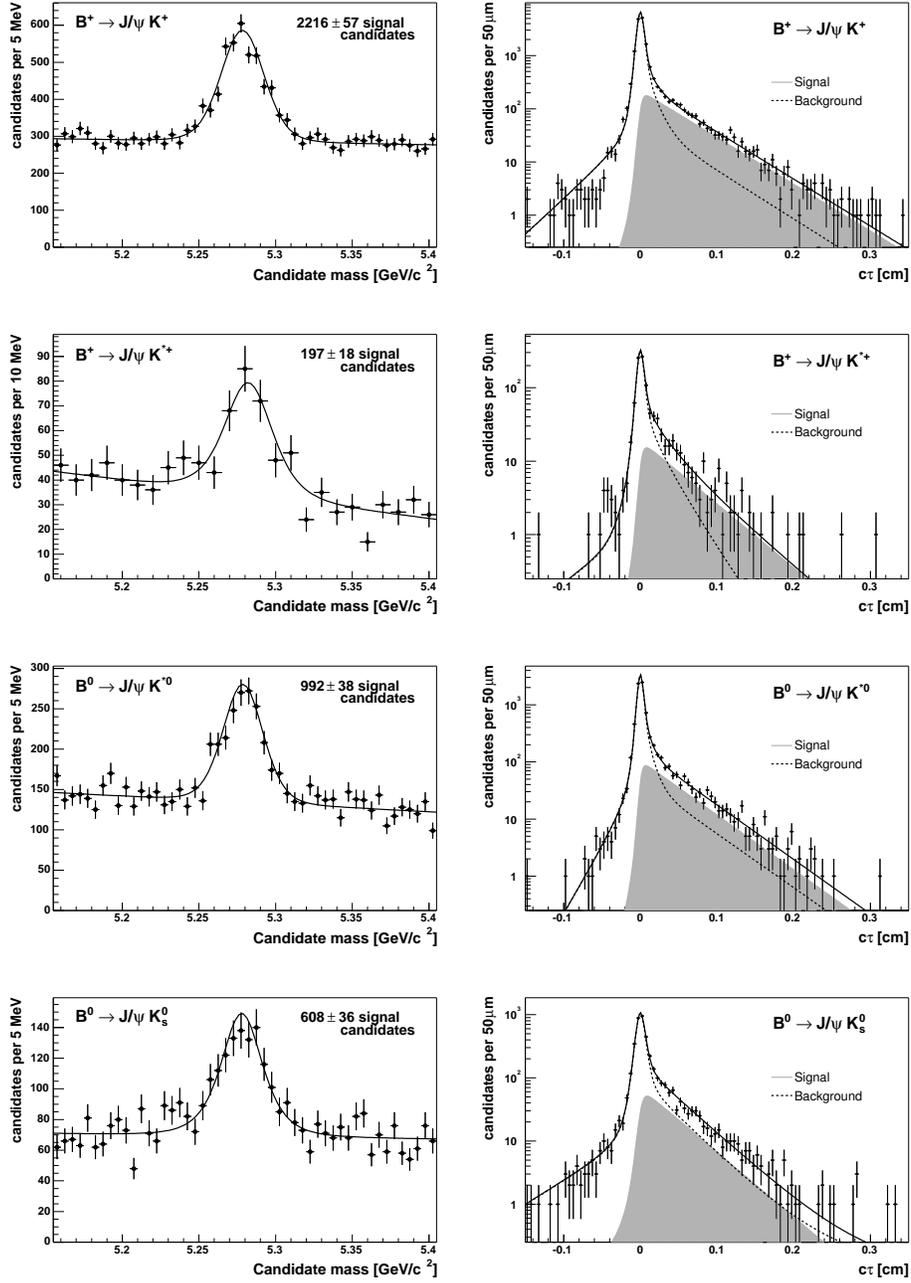
Figure 40: *Results of the Method 2 (simultaneous) fits for the four channels, using the*
B vertex. Rows (top to bottom): $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$ and
$B^0 \to J/\psi\, K^0_s$. Columns: Mass and $c\tau$ distributions. The results of the maximum likeli-
hood fits are overlaid.

Figure 41: *Results of the Method 2 (simultaneous) fits for the four channels, using the J/ψ vertex. Rows (top to bottom): $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$ and $B^0 \to J/\psi\, K_s^0$. Columns: Mass and $c\tau$ distributions. The results of the maximum likelihood fits are overlaid. The fit probabilities shown are obtained through a chi squared test, described in section 4.3.2.*

decayed [58], filtered using a simulation of the dimuon trigger [59], and reconstructed using the full CDF Run II detector simulation and basic reconstruction. For each of the four modes about 4.6 million events were generated, out of which about 3% passed the trigger filter and were thoroughly simulated and reconstructed. The beam spot was smeared using an ideal beam (no tilt or shift) with $\sigma_z = 30$ cm and $\sigma_t = 25\mu$m.

The resulting Monte Carlo events were then fed to basic reconstruction process (see section 3.3) and passed through the same analysis code we applied to data. A major difference between the data and the Monte Carlo, however, was that no background was present in the Monte Carlo samples. The study should be considered as a check of the procedures in place for estimating the proper decay time, and for selecting events in a way free of lifetime biases.

We checked the pull distributions, defined as $(c\tau - c\tau_{true})/\sigma_{c\tau}$, for each of the four channels. If there is no systematic bias on the measured values $c\tau$, and if the errors $\sigma_{c\tau}$ are properly estimated, the pull distributions should be Gaussian, with mean zero and unit width. The resulting distributions (in semi-log plots) are shown for the two ways of estimating the proper decay time (from the B vertex and from the J/ψ vertex; see explanation in section 3.4) in figures 42 and 43. These plots establish that the estimated error describes the deviations adequately. This is true for both ways of estimating the vertex. The resolutions are higher, in most channels[1], when the B vertex is used, however, the pulls justify the use of either means of estimating the proper decay time.

We also fit the sample using both methods described in Section 4.1. In each case we used the proper time as estimated both using the B vertex and the J/ψ vertex. The results are compared in Table 13, and the fits are shown in Figures 44, 45, 46, and 47. One can draw the following conclusions:

- All of the extracted values are consistent with the input values within their errors; this indicates that no problems with the procedure used to estimate proper decay time can be attributed to the procedure itself.
- There is no gross difference between the lifetime values that come from the two fitting procedures.

---

[1]The exception is $B^0 \to J/\psi\, K_s^0$; in that case, the J/ψ vertex is not significantly altered by the addition of the K_s^0 daughter tracks because they are displaced.
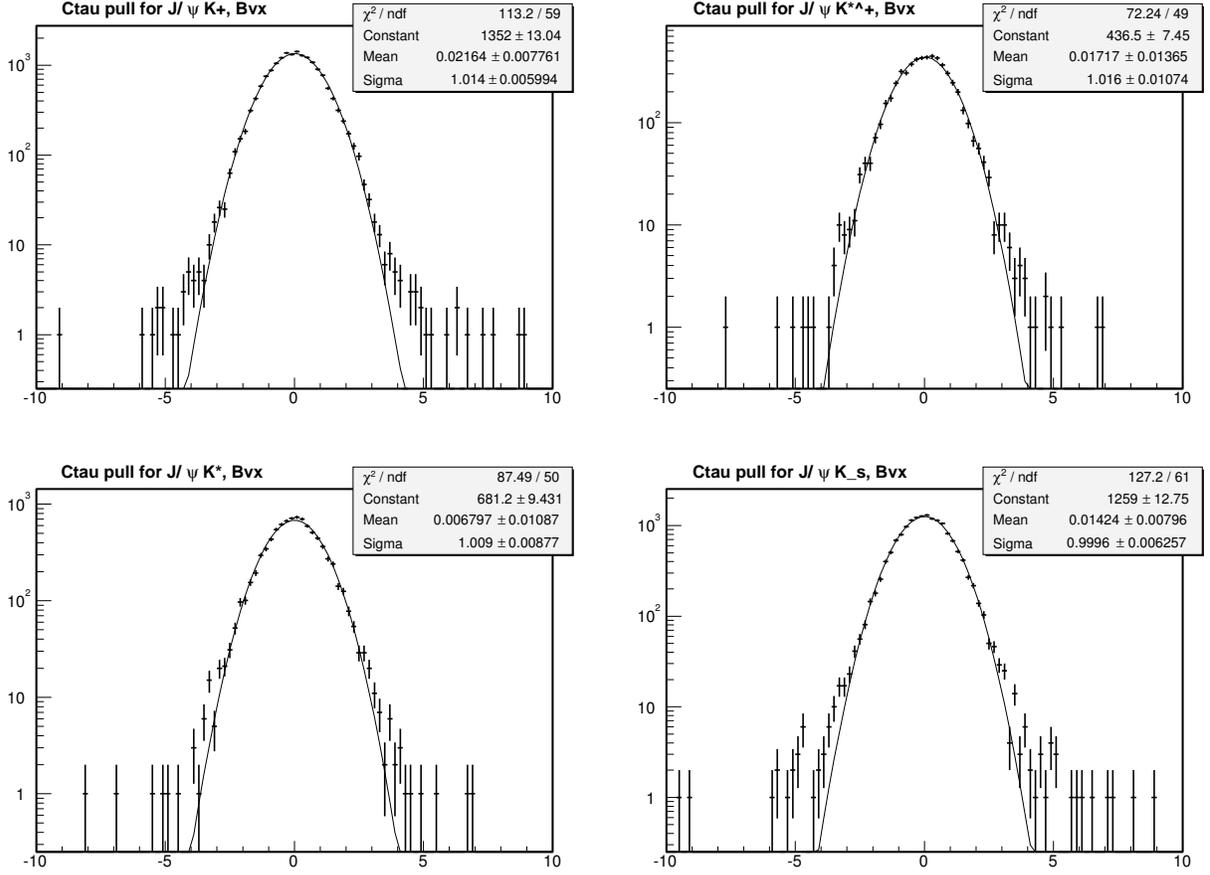
Figure 42: $c\tau$ pull distributions for the channels $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, and $B^0 \to J/\psi\, K^0_s$. The histograms show the event per event pull $(c\tau - c\tau_{true})/\sigma_{c\tau}$. The solid lines are single Gaussian fits. For these plots, the B vertex is used.
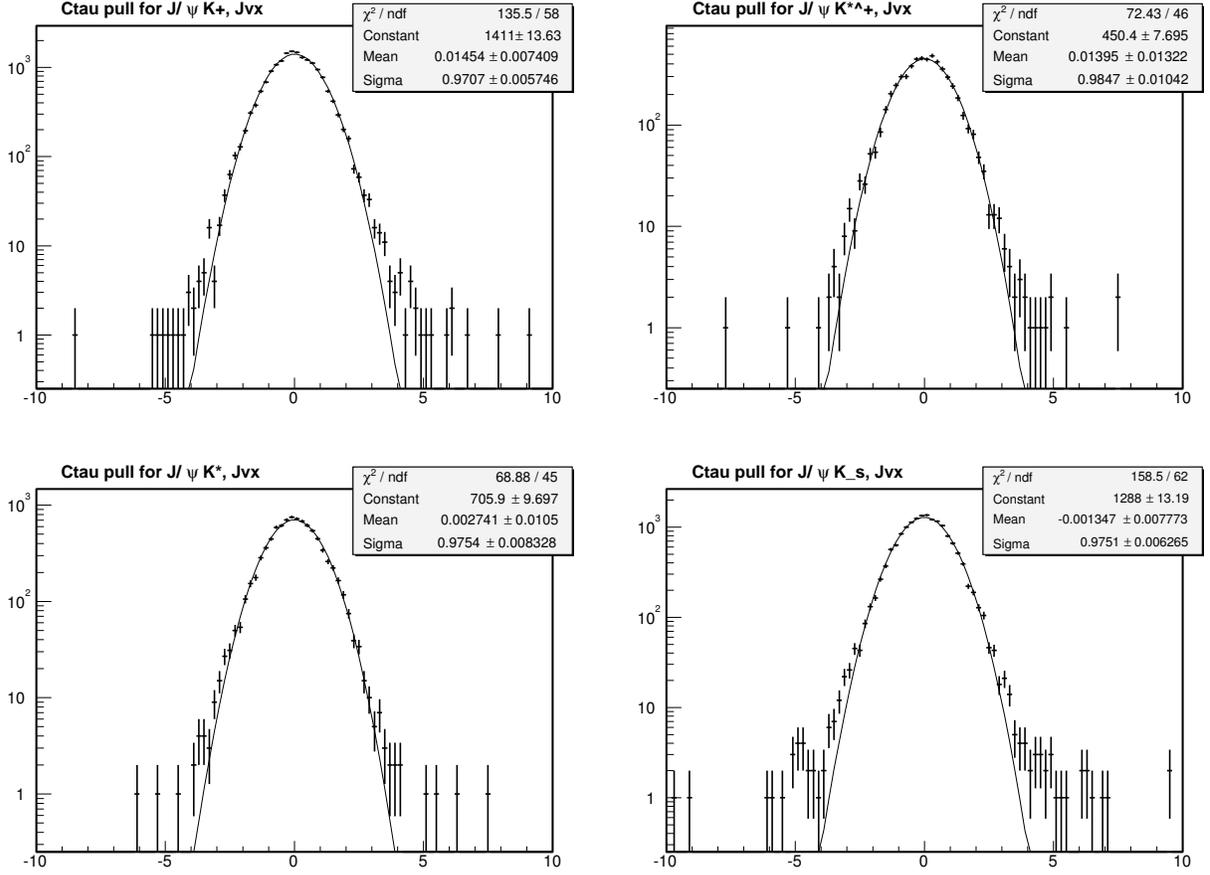
Figure 43: $c\tau$ pulls for the channels $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, and $B^0 \to J/\psi\, K_s^0$. The histograms show the event per event pull $(c\tau - c\tau_{true})/\sigma_{c\tau}$. The solid lines are single Gaussian fits. For these plots, the J/ψ vertex is used.

- The precision of a high-statistics measurement does not suffer when the J/ψ vertex is used instead of the B vertex; this is despite the appearance of visible negative lifetime tails. This is due, presumably, to the accuracy of the error estimation and the sufficiency of the proper time measurement from the J/ψ alone.

These studies justify the use of either Method I or Method II, and also the use of the B vertex or the J/ψ vertex.

4.2.2   Comparison of fitting methods

Both the simultaneous bivariate fit to mass and lifetime (method II, on which we base our results) and the two-step fit (in which the mass fit is performed to determine signal and sideband regions for a subsequent lifetime fit), appear from full Monte Carlo simulation to be equally justified. The methods as we have observed are correlated. The degree of correlation can be derived from Toy Monte Carlo simulation. The differences between the results of the two methods are shown in figure 48. For each channel, a number of Toy Monte Carlo experiments was generated using our nominal model and the parameters measured in data; each experiment had as many candidates as were actually observed in the corresponding channel. Each experiment is fitted using both methods and the distribution of differences is fit to a single Gaussian. We use the width of the Gaussian fits as an estimation of the expected difference between the two methods.

The expected differences are: $4\,\mu$mfor $B^+ \to J/\psi\,K^+$, $15\,\mu$mfor $B^+ \to J/\psi\,K^{*+}$, $7\,\mu$mfor $B^0 \to J/\psi\,K^*$, and $12\,\mu$mfor $B^0 \to J/\psi\,K_s^0$, while the actual deviations in the data are: $2.2\,\mu$m, $-31.6\,\mu$m, $-0.2\,\mu$m, and $18.7\,\mu$m, respectively. The fact that these deviations are consistent with the expected statistical fluctuations gives us further confidence in the quoted values.
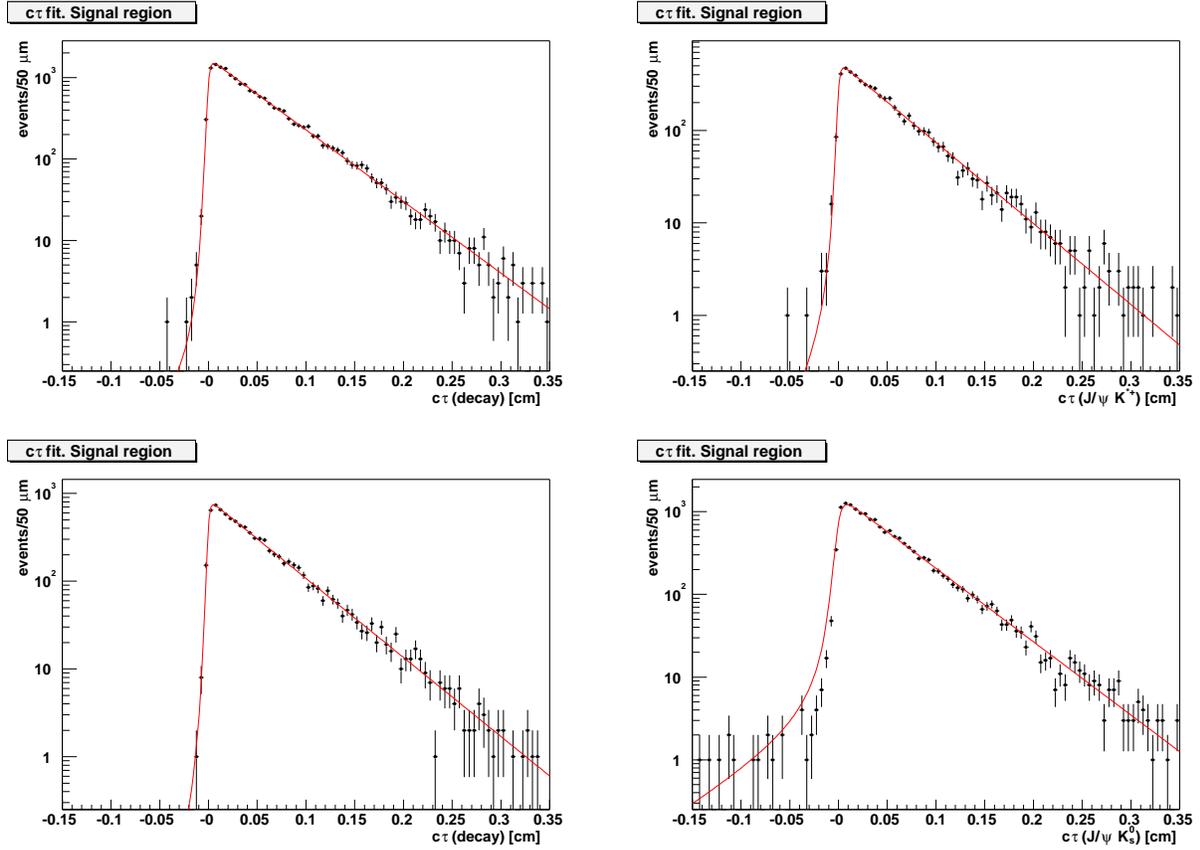
Figure 44: *Method 1 Fits to the fully simulated Monte Carlo for the channels $B^+ \rightarrow J/\psi\, K^+$, $B^+ \rightarrow J/\psi\, K^{*+}$, $B^0 \rightarrow J/\psi\, K^*$, and $B^0 \rightarrow J/\psi\, K^0_s$. The B vertex was used to estimate the proper decay time.*
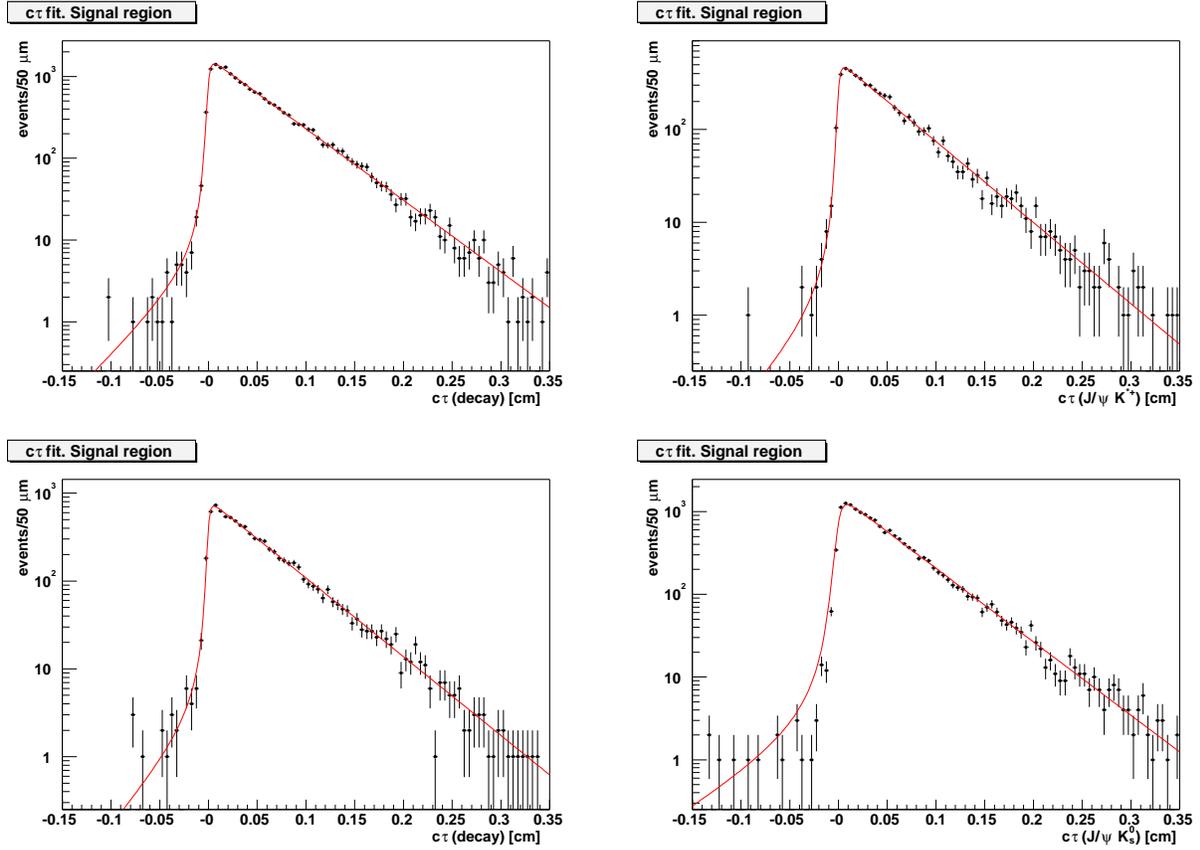
Figure 45: *Method 1 Fits to the fully simulated Monte Carlo for the channels $B^+ \rightarrow J/\psi\,K^+$, $B^+ \rightarrow J/\psi\,K^{*+}$, $B^0 \rightarrow J/\psi\,K^*$, and $B^0 \rightarrow J/\psi\,K_s^0$. These fits use the J/ψ vertex to estimate proper decay time.*
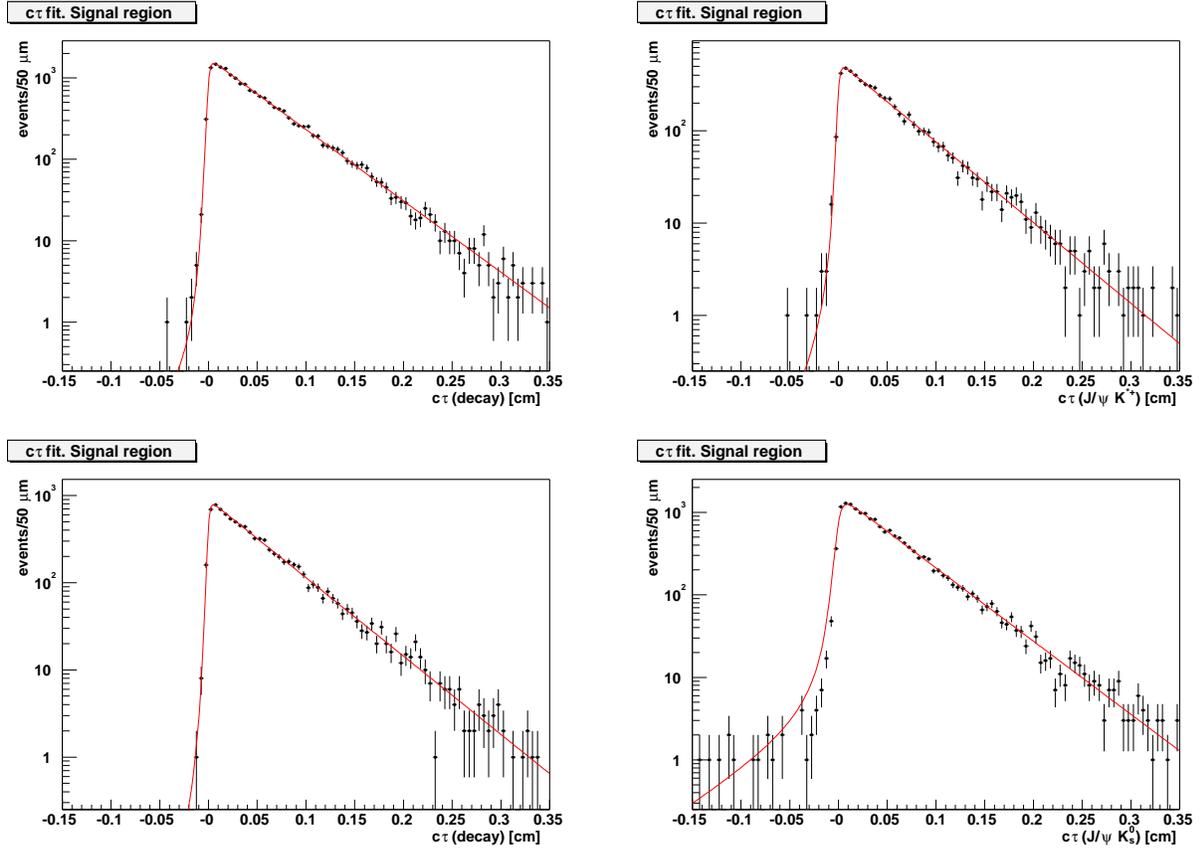
Figure 46: *Method 2 Fits to the fully simulated Monte Carlo for the channels $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, and $B^0 \to J/\psi\, K_s^0$. These fits use the B vertex to estimate proper decay time.*
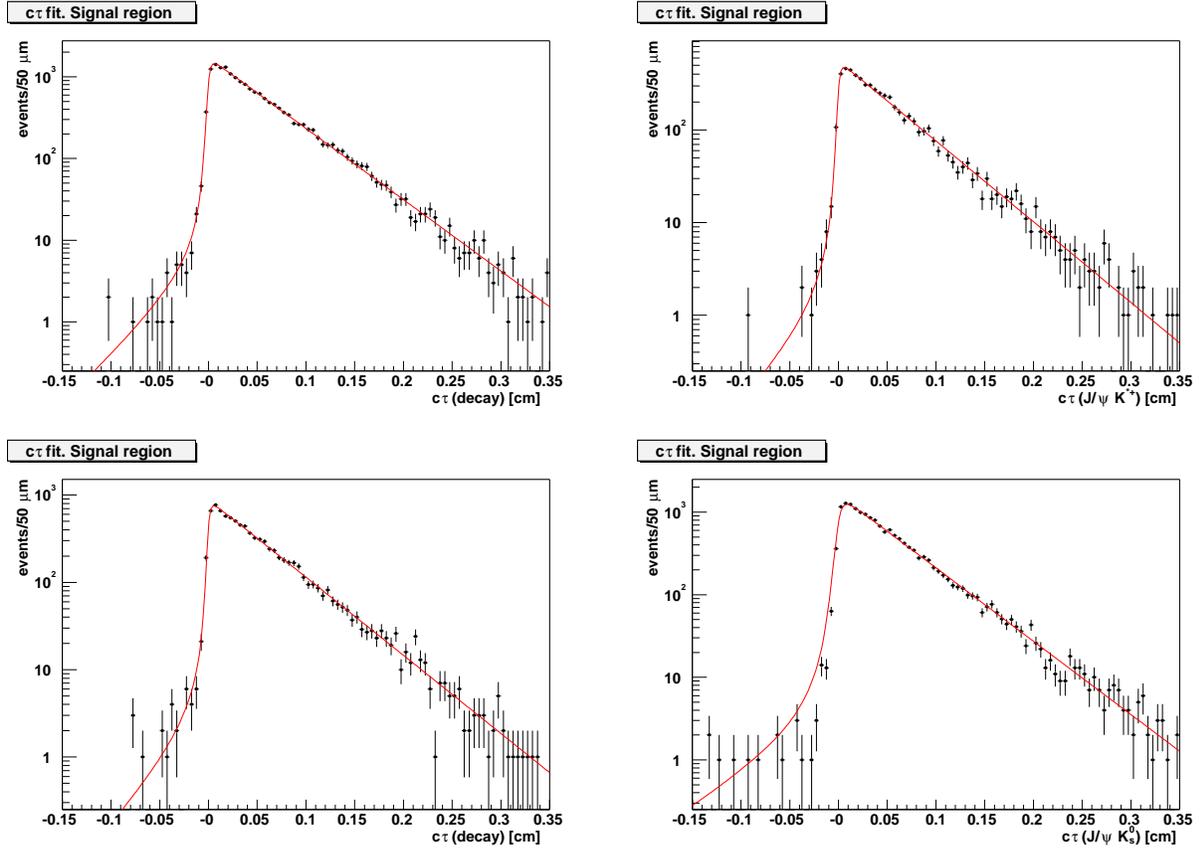
Figure 47: *Method 2 Fits to the fully simulated Monte Carlo for the channels $B^+ \to J/\psi \, K^+$, $B^+ \to J/\psi \, K^{*+}$, $B^0 \to J/\psi \, K^*$, and $B^0 \to J/\psi \, K_s^0$. These fits use the J/ψ vertex to estimate proper decay time.*

Table 13: *Comparison of full Monte Carlo lifetime measurements for the four channels using both methods. The results of using both the J/ψ vertex and the B vertex are shown. All values are in microns.*

Channel	True Value	M1 BVTX	M1 JVTX	M2 BVTX	M2 JVTX
$B^+ \to J/\psi \, K^+$	493	494.2 ±3.8	495.6 ±3.8	495.2 ±3.8	496.6 ±3.8
$B^+ \to J/\psi \, K^{*+}$	493	495.8 ±6.7	496.2 ±6.7	496.6 ±6.6	497.2 ±6.7
$B^0 \to J/\psi \, K^*$	486	482.4 ±5.3	483.8 ±5.4	483.4 ±5.2	484.6 ±5.2
$B^0 \to J/\psi \, K_s^0$	486	485.7 ±3.9	485.2 ±3.9	485.7 ±3.9	485.2 ±3.9

4.3 GOODNESS-OF-FIT ESTIMATION AND PULL DISTRIBUTIONS

4.3.1 Toy Monte Carlo

We checked both kinds of fit (simultaneous and non-simultaneous) for biases in the fitting procedure and for goodness-of-fit. Both types of check were performed using a Toy Monte Carlo which generated mass and lifetime points according to the model used in the fit. In each case, over 4000 Toy Monte Carlo experiments were performed. The parameters governing the mass and lifetime distributions were taken to be the values extracted from the data. Each experiment has the same statistics as the corresponding channel in data. Each toy Monte Carlo candidate uses, as its $\sigma_{c\tau}$, an estimated error taken from the data. In method II, which also makes use of estimated mass errors, the σ_m is also taken from a real data candidate.

A goodness-of-fit estimator (a "confidence level") was obtained by comparing the likelihood of the actual experiment (data) against the likelihood distribution of the Toy Monte Carlo experiments. We determined the number of Toy Monte Carlo experiments with a higher value of $-\ln(\mathcal{L})$ than that obtained for data. For Method 1, The central column of figure 49 shows the spectrum of all experiments and of those with a larger $-\ln \mathcal{L}$. The
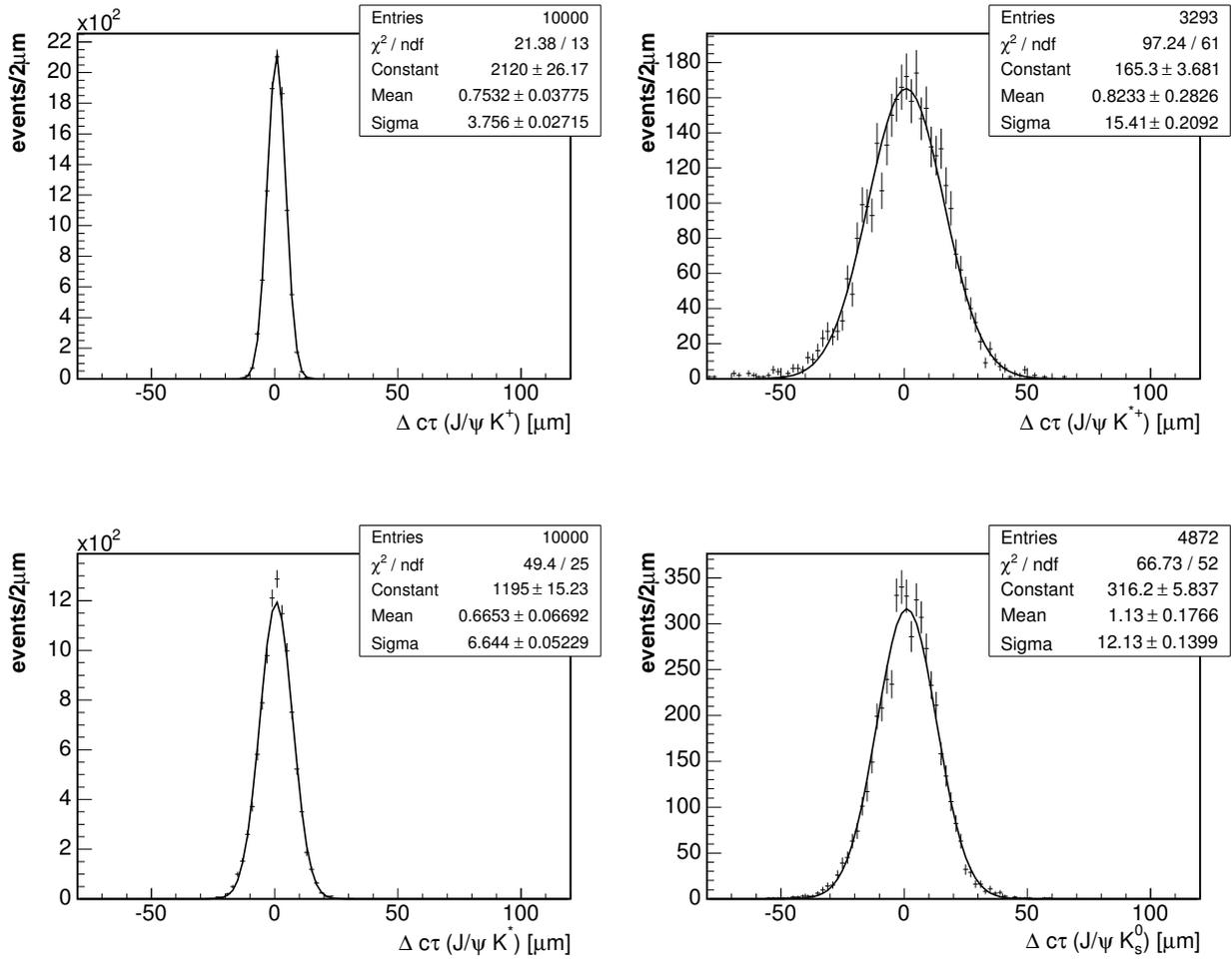
Figure 48: *Difference between cτ from methods 1 and 2 using the toy Monte Carlo.*

confidence levels that one obtains are: 18.6% for $B^+ \to J/\psi \, K^+$, 11.0% for $B^+ \to J/\psi \, K^{*+}$, 24.6% for $B^0 \to J/\psi \, K^*$, and 20.2% for $B^0 \to J/\psi \, K^0_s$. For Method II, the plots are shown in figure 50, and the corresponding confidence levels are: 3.4%, 35.2%, 25.6%, and 26.3%.

We also obtained the lifetime pull distributions for the ensemble of experiments for each of the four channels. In this case, the quantities used are the proper decay length $c\tau_{data}$ used for generation of all Toy Monte Carlo experiments (for each channel), the value $c\tau_{fit}$ obtained by the fit, and the error on this quantity, σ_{fit}. As before, the distribution of $(c\tau_{fit} - c\tau_{data})/\sigma_{fit}$ should be close to a unit Gaussian. The pulls are shown in the plots on the left column of figures 49 and 50. Single Gaussian fits are overlaid on the semilog plots. Fitted means and sigmas of the pull distributions are shown in Table 14. Note that small downward shifts in pull distributions may occur in low-statistics samples, as discussed in [60]. The origin of these shifts is the correlation in the estimates of lifetime and lifetime error.

Also of interest is the behavior of $-\log \mathcal{L}$ around the minimum. The plots in the right column of figures 49 and 50 show that, for all channels, the shape around the minimum, as expected, has no secondary minima or other anomalies.

4.3.2 Chi squared test

As another estimator of the goodness of fit, we used a χ^2 test. Both the data and the fit distribution are binned using unequal bin sizes chosen so that no bin has less than 20 expected events (with a minimum bin size of 10 microns). A χ^2 is then calculated as

$$\chi^2 = \sum_{i=1}^{n_{bins}} \frac{(d_i - p_i)^2}{\sigma_{d_i}^2} \; , \tag{4.10}$$

where n_{bins} is the number of bins, d_i is the number of data points in bin i and p_i is the expected number of events for that bin.

An estimation of the goodness of fit can then be obtained as the probability for χ^2, with $n_{bins} - 9$ degrees of freedom (since nine parameters are used to describe the $c\tau$ distribution). Figure 51 shows the comparison between expected number of events (solid line) and data (points) for each bin. It depicts the results for the procedure we will use for our nominal
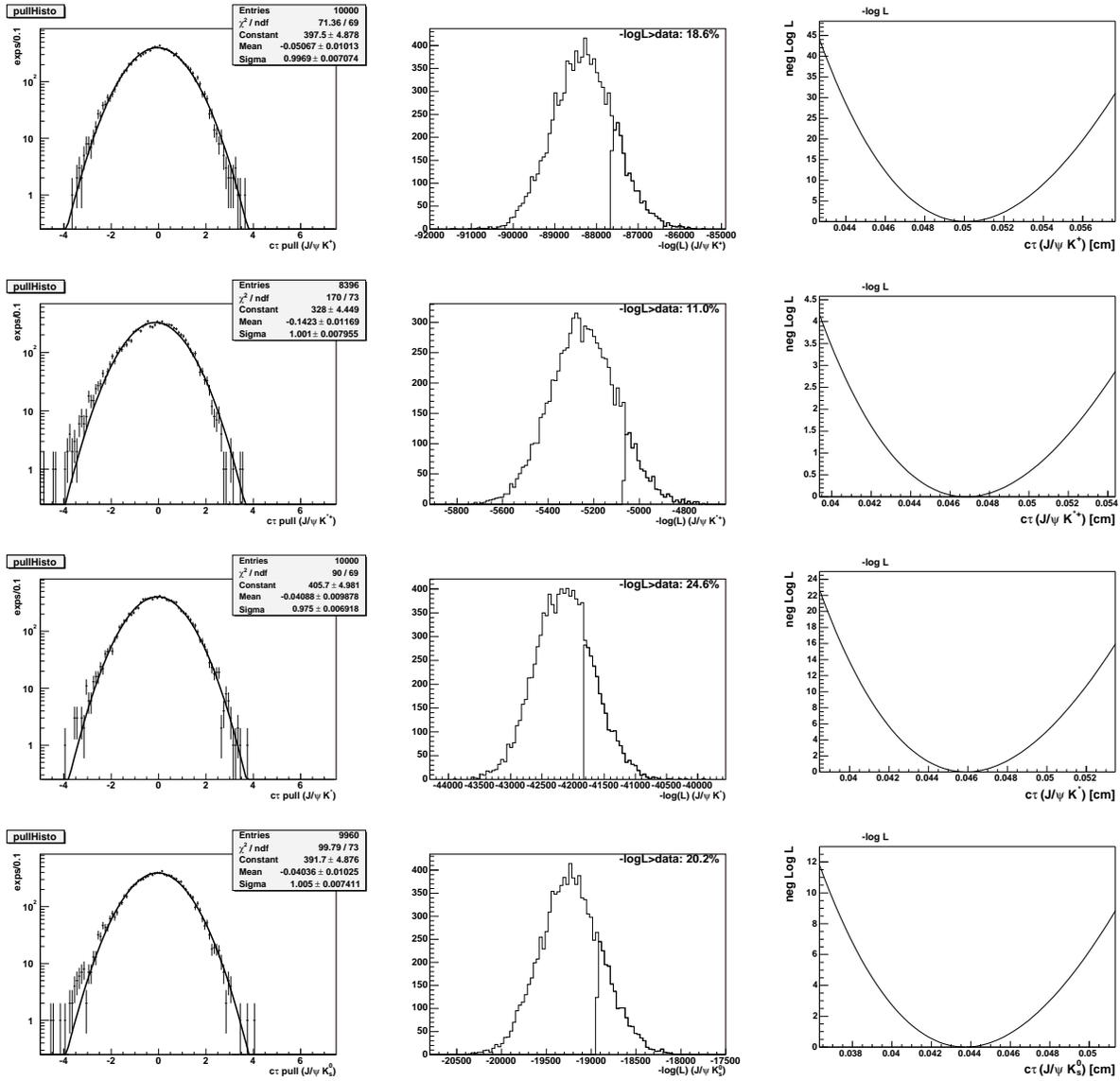
Figure 49: *Left: Lifetime pull distributions for 500 Toy Monte Carlo Experiments using Method I. Center: likelihood distributions (confidence levels: 18.6%, 11.0%, 24.6%, 20.2%). Right: Likelihood profiles. Rows (top to bottom): $B^+ \to J/\psi\,K^+$, $B^+ \to J/\psi\,K^{*+}$, $B^0 \to J/\psi\,K^*$, $B^0 \to J/\psi\,K_s^0$.*
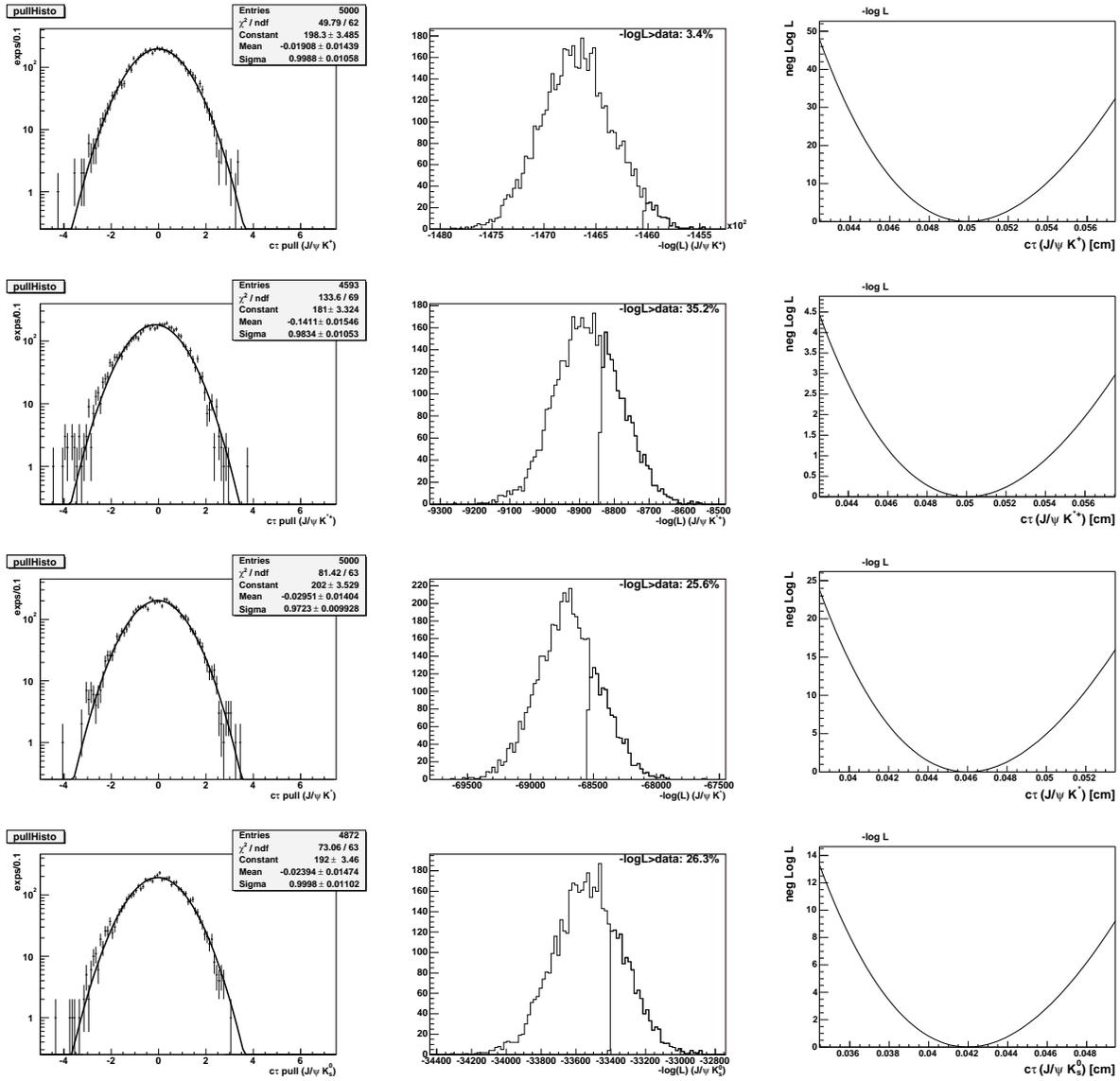
Figure 50: *Left: Lifetime pull distributions for 500 Toy Monte Carlo Experiments using Method II. Center: likelihood distributions (confidence levels: 3.4%, 35.2%, 25.6%, 26.3%); right: Likelihood profiles. Rows (top to bottom):* $B^+ \to J/\psi\, K^+$, $B^+ \to J/\psi\, K^{*+}$, $B^0 \to J/\psi\, K^*$, $B^0 \to J/\psi\, K_s^0$ *.*

Table 14: *Means and sigmas of the Toy Monte Carlo Pulls for the four channels.*

Toy Monte Carlo lifetime pulls		
$B^+ \to J/\psi\, K^+$		
Parameter	Meth1	Meth2
Mean	-0.051 ± 0.014	-0.019 ± 0.014
Sigma	0.997 ± 0.007	0.999 ± 0.011
$B^+ \to J/\psi\, K^{*+}$		
Parameter	Meth1	Meth2
Mean	-0.142 ± 0.012	-0.141 ± 0.015
Sigma	1.001 ± 0.008	0.983 ± 0.011
$B^0 \to J/\psi\, K^*$		
Parameter	Meth1	Meth2
Mean	-0.041 ± 0.010	-0.030 ± 0.014
Sigma	0.975 ± 0.007	0.972 ± 0.010
$B^0 \to J/\psi\, K_s^0$		
Parameter	Meth1	Meth2
Mean	-0.040 ± 0.010	-0.024 ± 0.015
Sigma	1.005 ± 0.007	1.000 ± 0.011

results (simultaneous fit using the J/ψ vertex). The peculiar shape of these plots is a result of the procedure used. In the region close to $c\tau=0$, no change of binning is needed, due to the prompt background. Once the expected number per bin falls below 20, the bin width is increased (in increments of $10\,\mu$m) until 20 or more events are expected. As a result of changing the bin size, both the histogram and the expected number of candidates increase their value per bin. As the value of $c\tau$ gets farther away from 0, both the number of data candidates and the likelihood model expectation decrease, until the bin size has to be changed again.

As shown in figure 51, the probabilities for the fits are 38.2%, 30.1%, 4.2% and 48.1% for $J/\psi\,K^+$, $J/\psi\,K^{*+}$, $J/\psi\,K^*$ and $J/\psi\,K^0_s$, respectively.

4.4 COMBINATION OF SINGLE-CHANNEL RESULTS

We have described four measurements of fully reconstructed B decays; two of them for B^0 and two for B^+. In each measurement the mass is obtained, plus a scale factor and the lifetime of a neutral or charged B meson. In either case a number of other parameters are determined, such as prompt fraction and the effective lifetime of a positive or negative tail.

To combine the results from the two decay channels available for each meson species, we treat each channel's results as a joint measurement of mass, lifetime and proper decay length error scale factor; the two channels are regarded as separate, independent determinations of the *same* triplet of numbers. We then apply the following procedure, which is essentially a vector weighted average of the measurements.

Let us denote the Likelihood functions as:

$$\mathcal{L}_1 = \mathcal{L}_1(\alpha_{11}, \alpha_{12}, ...\alpha_{1M}) = \mathcal{L}(\vec{\alpha}_1) \tag{4.11}$$

$$\mathcal{L}_2 = \mathcal{L}_2(\alpha_{21}, \alpha_{22}, ...\alpha_{2N}) = \mathcal{L}(\vec{\alpha}_2) \tag{4.12}$$

where the $\vec{\alpha}_i$ is the vector of parameters of the ith maximum likelihood fit. We shall suppose these parameters to have been adjusted so that $-\ln(\mathcal{L}_i)$ is at a minimum. In addition to the array of values, $\vec{\alpha}_i$, the fitting procedures return real symmetric covariance matrices $\mathbf{C_i}$. Let $\vec{\alpha}_1$ have dimension M and $\vec{\alpha}_2$ have dimension N.
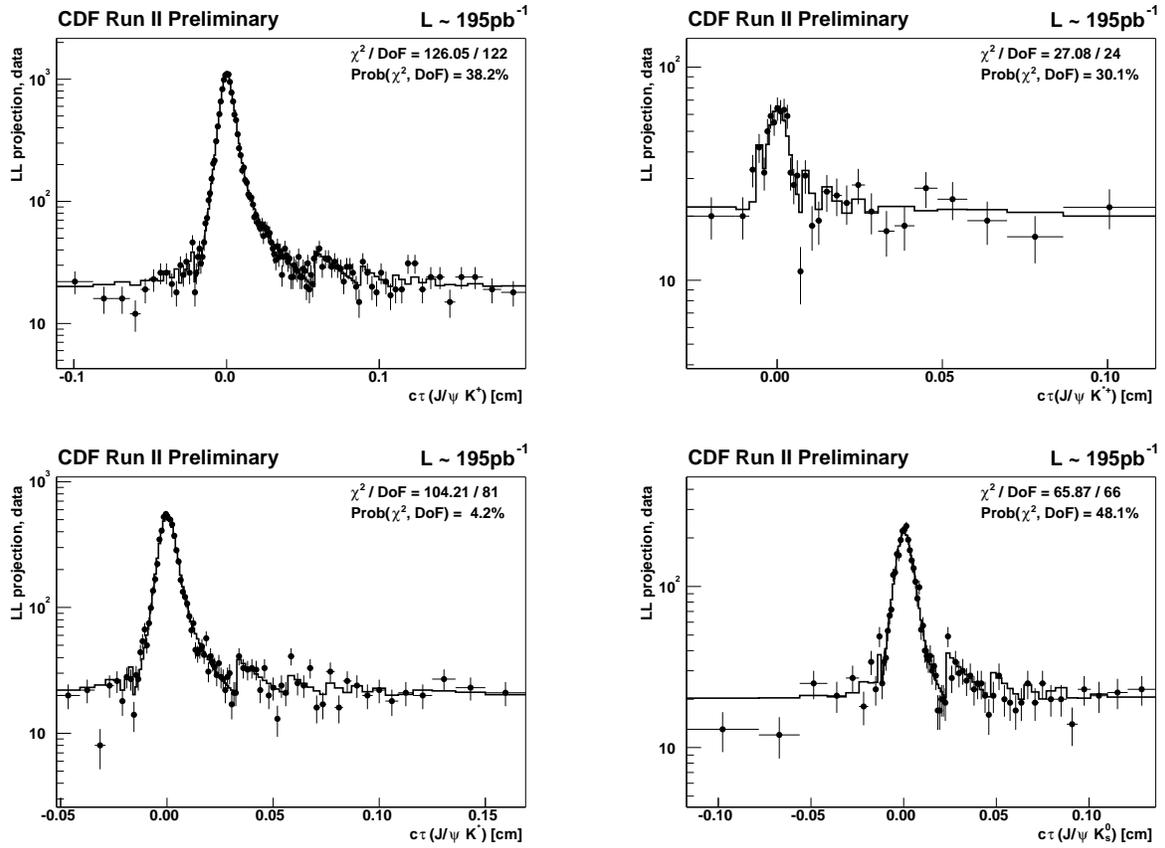
Figure 51: χ^2 test for the nominal fits (method 2, J/ψ vertex).

We make a vector out of all the distinct quantities that the two likelihood fits constrain and call it $\vec{\beta}$; it has dimension P, which is greater than $\min(M, N)$ and less than or equal to $M + N$. We can order this vector however we like. The order in which the elements of $\vec{\alpha}_1$ are mapped into $\vec{\beta}$ is encapsulated in a $P \times M$ matrix called $\mathbf{H_1}$, whose elements are:

$$\{\mathbf{H_1}\}_{ij} = \frac{\partial \alpha_{1i}}{\partial \beta_j} \tag{4.13}$$

Similarly, the matrix

$$\{\mathbf{H_2}\}_{ij} = \frac{\partial \alpha_{2i}}{\partial \beta_j} \tag{4.14}$$

describes the way that element of $\vec{\alpha}_2$ maps into $\vec{\beta}$. The matrix elements of $\mathbf{H_1}$ and $\mathbf{H_2}$ are either one or zero. For this analysis we define $\vec{\beta} = \{\tau, m, s_l; f_{g1}, f_{g2}, f_{p1}, f_{p2}...\}$. The matrices $\mathbf{H_1}$ and $\mathbf{H_2}$ in this case would be:

$$\mathbf{H_1} = \left(\begin{array}{ccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & ... \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & ... \\ ... & & & & & & & \end{array} \right) \tag{4.15}$$

and

$$\mathbf{H_2} = \left(\begin{array}{ccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & ... \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & ... \\ ... & & & & & & & \end{array} \right) \tag{4.16}$$

The first three rows/columns from both matrices form an identity matrix; when acting on $\vec{\alpha}_1$ and $\vec{\alpha}_2$, this puts the first three parameters from both measurements in the same positions on $\vec{\beta}$. From there on, the two matrices separate the remaining entries from $\vec{\alpha}_1$ and $\vec{\alpha}_2$ into different entries in $\vec{\beta}$.

The combined measurement is computed as:

$$\vec{\beta} = \left(\mathbf{H_1^T} \cdot \mathbf{C_1^{-1}} \cdot \mathbf{H_1} + \mathbf{H_2^T} \cdot \mathbf{C_2^{-1}} \cdot \mathbf{H_2}\right)^{-1} \cdot \left(\mathbf{H_1^T} \cdot \mathbf{C_1^{-1}} \cdot \vec{\alpha_1} + \mathbf{H_2^T} \cdot \mathbf{C_2^{-1}} \cdot \vec{\alpha_2}\right) \quad (4.17)$$

The error on the combined measurement is:

$$\mathbf{C} = \left(\mathbf{H_1^T} \cdot \mathbf{C_1^{-1}} \cdot \mathbf{H_1} + \mathbf{H_2^T} \cdot \mathbf{C_2^{-1}} \cdot \mathbf{H_2}\right)^{-1} \quad (4.18)$$

The consistency of the two measurements can also be checked *via*:

$$\chi^2 = \left(\mathbf{H_1^T} \cdot \vec{\alpha_1} - \mathbf{H_2^T} \cdot \vec{\alpha_2}\right)^T \cdot \left(\mathbf{H_1^T} \cdot \mathbf{C_1^{-1}} \cdot \mathbf{H_1} + \mathbf{H_2^T} \cdot \mathbf{C_2^{-1}} \cdot \mathbf{H_2}\right) \cdot \left(\mathbf{H_1^T} \cdot \vec{\alpha_1} - \mathbf{H_2^T} \cdot \vec{\alpha_2}\right) \quad (4.19)$$

This is equivalent to performing a combined likelihood fit whenever the likelihood functions are parabolic in the parameters.

Applying this procedure to the values obtained from the method 1 fit using the J/ψ vertex for all channels, and the corresponding error matrices, we obtain:

$$c\tau_{B^+} = 499 \pm 12 \mu m$$

$$c\tau_{B^0} = 446 \pm 15 \mu m$$

Figure 52 shows these results, together with the lifetimes for the individual channels. The solid horizontal lines show the combined value and the dotted horizontal lines show the error for each measurement. A detailed discussion of systematic errors coming from silicon alignment, the choice of resolution function, and background model is presented in the next section.
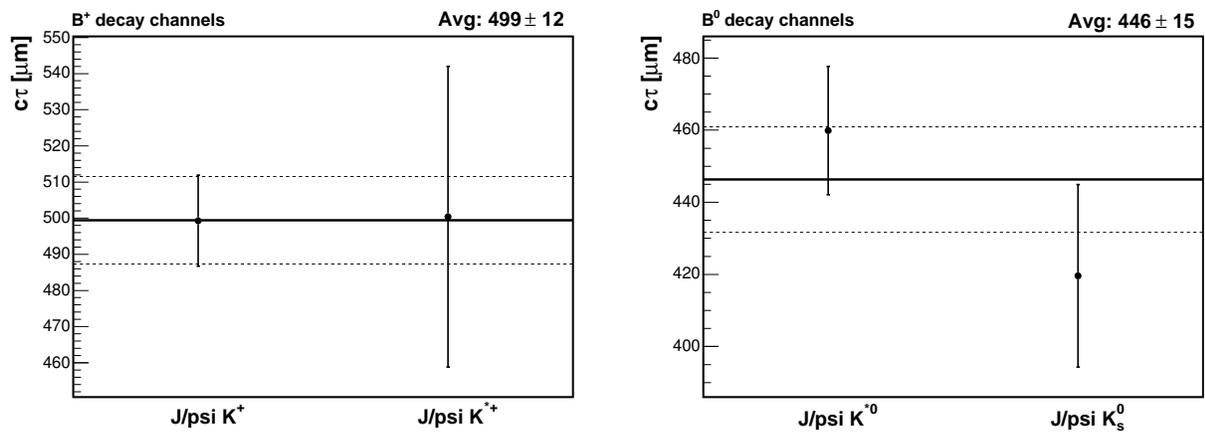
Figure 52: *Combination of $c\tau$ measurements for B^+ and B^0 decay modes. Errors are statistical only.*

5.0 SYSTEMATIC ERRORS

In this chapter, we study the systematic uncertainties that arise from the procedure used to extract and model our data. Since our measurements are based on information obtained from silicon tracks, we need to quantify the effect, on our measured quantities, of the uncertainty in the alignment of the silicon detector. We also evaluate the effect of choosing a different, but equally reasonable, model for the $c\tau$ resolution. Finally, we assess the uncertainty associated with the shape of the background distribution.

5.1 ALIGNMENT

Misalignments of the silicon systems can be either random displacements of individual detector components or collective motions of the detector as a whole. The type of displacement that affects lifetime measurements most strongly is a radial dilation or contraction of the system. Such an effect can be due to a bowing of silicon ladders, which is known to occur in CDF. The size of the effect is conservatively estimated at 50 microns by the tracking group [61].

One can derive a simple yet robust estimation of the effect of a 50 micron bowing on lifetime measurement by noting that the fractional change to the position of the first silicon layer, at 2.44 cm, is 0.2%; if such a distortion is present, it will shift the value of a measured lifetime by the same fraction, which, for a particle with a $c\tau$ of $500\,\mu$m, is $1\,\mu$m. We expect that the true alignment systematic error is about $1\,\mu$m.

In CDF, the information about the alignment of the silicon detector is stored in "alignment tables". They are basically a set of constants that specify the position and orientation

Table 15: *Fitted values of cτ for various alignments when reconstructing $B^- \to J/\psi K^-$ decays in MonteCarlo using method 2 and the B vertex.*

Deviation from nominal	Fitted Lifetime (μm)
None	476 ± 14
Shift out	477 ± 14
Shift in	478 ± 14
Bow out	477 ± 14
Bow in	478 ± 14

of each component of the system. In order to measure the alignment systematic uncertainty, we used five different alignment tables, specially prepared by the alignment group to represent the following distortions (with respect to the table used for data analysis):

1. None (version used for reconstruction).

2. All ladders shifted radially +50 μm.

3. All ladders shifted radially -50 μm.

4. All ladders bowed outward 50 μm.

5. All ladders bowed inward 50 μm.

we computed the effect, on the fitted lifetimes in Monte Carlo and data, of using these alignment tables.

The Monte Carlo sample was obtained from a CDF-wide available dataset produced with the same beamline as that used for data and a dimuon trigger requirement of two muons with $p_T > 1.4$ GeV/c. The results are listed in table 15. The fitted value of 476 ± 14 is consistent within statistics with the generated value of 493 μm. Changing the alignment table to account for bowing produces a maximum deviation of 2 μm.

For data, we reconstructed each of the four channels using each of the five alignment tables. In order to factor out the effect of selection differences, we considered, for the

Table 16: *Shifts on the fitted $c\tau$ values of all modes for different alignments of the silicon detector. The first row shows the values obtained with the alignment used for the nominal fit. All numbers are in microns.*

Alignment	$J/\psi\,K^+$	$J/\psi\,K^{*+}$	$J/\psi\,K^*$	$J/\psi\,K_s^0$
Nominal	494.5	490.5	454.2	420.2
Shift out	0.7	-1.1	-0.3	1.5
Shift in	-1.4	-1.1	0.2	-1.8
Bow out	0.7	-0.6	-0.7	0.3
Bow in	-1.1	-0.2	-0.7	-0.7

subsequent lifetime fits, only events that were selected by all alignment versions of the reconstruction. Table 5.1 shows the effect, on each channel's fitted lifetime, of the use of each alignment table. All fits use the nominal procedure (method II, J/ψ vertex). The first row shows the fitted $c\tau$ (in μm) obtained with the best alignment (the one used for our nominal results). Rows 2 to 5 show the shifts (with respect to row one) of the fitted lifetime when *the same B candidates* are used, while the relevant quantities were obtained with the corresponding alignment table. All shifts are also expressed in μm.

The channels $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K_s^0$ have the behavior expected from the type of distortion introduced on each alignment table (alignments bowed or shifted out yield higher lifetimes). However, the two remaining modes ($J/\psi\,K^{*+}$ and $J/\psi\,K^*$) do not. This is due to the fact that the lifetime fit can assimilate the changes in reconstructed quantities into different components of the fitting model. In particular, the lifetime of the long lived component of the background, λ_{++}, (and the corresponding fraction) can vary, effectively modifying the number of candidates assigned to signal and background. In the case of $J/\psi\,K^+$, λ_{++} varies by $8\,\mu$m (1.9%) among all alignment versions, while for $J/\psi\,K^*$ its value varies by up to $28\,\mu$m (6.1%).

This may be interpreted as saying that part of the variation we observe in this test overlaps with the one coming from the background model. The observed variation would then be an upper bound for this error. This upper bound is already small compared to other sources, so we assign $2\,\mu$m as the systematic uncertainty coming from alignment.

5.2   RESOLUTION FUNCTION

As described in section 4.1, our nominal fit uses a single Gaussian resolution function for decay lengths, and allows for an overall misestimation of errors via an error scale factor for the width. As shown in section 4.3, this choice provides an adequate description of our data, and comes out with a lifetime error scale factor of about 1.2 in all modes.

On the other hand, it may well be that the resolution function for our vertices is inherited from that of the inclusive J/ψ's. In this section, we study the resolution function of inclusive J/ψ's and use it to evaluate the systematic uncertainty induced by our ignorance of the details of the true resolution function in J/ψ vertices from our B decays.

5.2.1   Resolution function for J/ψ vertices

As described in section 4.1.2, our measurement is based on information obtained using the J/ψ vertices. This has the advantage of having shared systematics across modes, hence reducing their impact on the ratio. A second advantage of using the J/ψ vertices for the measurement is that the mere size of the inclusive J/ψ sample allows us to carry out a detailed study of its L_{xy} resolution function.

In order to characterize the inclusive J/ψ resolution function, we used a J/ψ sample of approximately 1.2 million signal events, reconstructed with the same procedure used for the J/ψ candidates formed during the reconstruction of B mesons. Figure 53 shows the mass distribution of J/ψ candidates used (histogram), together with a fit (double Gaussian over a linear background, solid line) and the norm (number of candidates), mean and sigma for both Gaussian components.
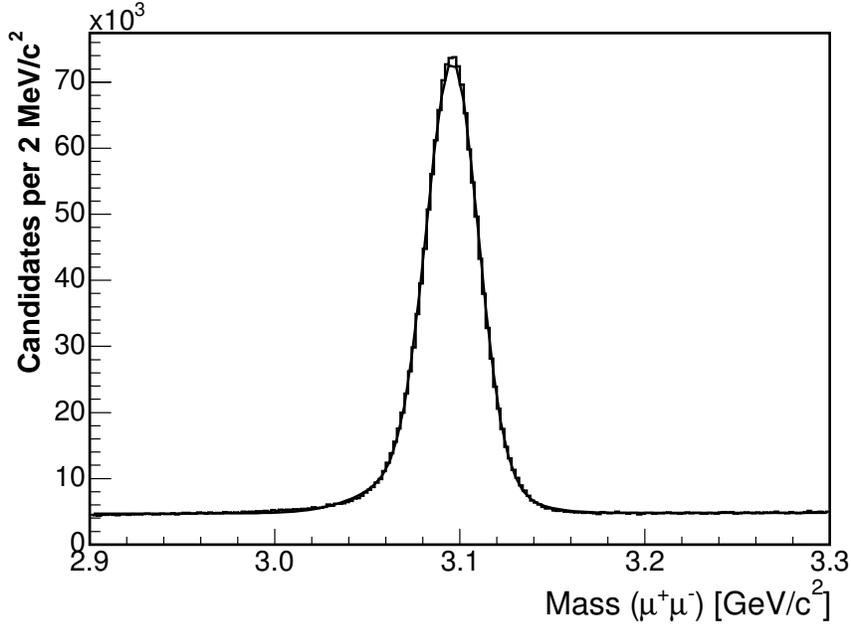
Figure 53: *Invariant mass distribution of the sample used for studying the resolution function of the J/ψ vertex. The fit shown is a double Gaussian over a linear background.*

A first evaluation of the resolution function for the J/ψ vertices was performed using the estimated L_{xy} errors returned by the (two-track) vertex fit. Given the size of the sample, a full-fledged maximum likelihood fit cannot be applied. Instead, we use the following procedure, depicted in figure 54:

1. The sideband-subtracted distribution of the estimated L_{xy} errors (as returned by the two-track vertex fit) is used to weight corresponding Gaussian components (for now, *single* Gaussian components). As suggested in figure 54, the width of the Gaussian is determined by the value of the L_{xy} error, while its weight in the overall sum is given by the bin content for that error value in the L_{xy} error histogram.

2. The resulting sum is convolved with a model for the signal (a delta function plus an exponential positive tail). Only the signal is modeled because we expect the sideband subtraction to do a good job eliminating the background, especially given the S/N ratio of our J/ψ sample.

3. The resulting function is then fit to the sideband-subtracted L_{xy} distribution of the inclusive J/ψ sample. The following parameters are determined by the fit:

118

- The characteristic length of the exponential tail

- The fraction of events on the tail, and

- An overall scale factor

The distribution of residuals (determined as the difference between the model so obtained and the sideband-subtracted L_{xy} distribution) is shown in figure 55. The distribution shows a clear structure. Two features can be emphasized at this point:

1. The size of the effect is of $\mathcal{O}(1\%)$ (compare figure 55 with figure 54, with peak values close to 200,000), and

2. The structure is statistically significant,

Of particular importance for lifetime measurements are the long lived tails of the distribution of residuals (tails that reach up to $500\,\mu$m). The residuals close to L_{xy}=0, on the other hand, may have an effect on the lifetime scale factors, but are not expected to significantly bias the $c\tau$ measurements.

The distribution of residuals was used to guide the selection of extra components to be added to the resolution function in order to better model the data. In particular, the structure shown in figure 55 can be interpreted as indicating the presence of a bimodal distribution of errors, as will be shown next.

The distribution of L_{xy} errors is indeed bimodal, as can be seen in figure 56. The left plot in this figure shows the distribution of J/ψ vertices along the z axis of the detector. The two valleys on this plot correspond to regions in which the presence of more material reduces the efficiency of J/ψ reconstruction and increases the L_{xy} estimated error. The right plot in figure 56 is a color-coded two dimensional histogram; the horizontal axis is the z coordinate of the J/ψ vertex (in cm) and the vertical axis is the error on L_{xy}. Each bin's content is shown using the color of the bin. Notice the higher errors in the regions about $|z| = 15$.

In order to observe the effect of this on the distribution of residuals, we performed a simple experiment. The left plot on figure 57 shows the mean value of the L_{xy} error distribution as a function of the z position of the J/ψ. It shows that there are two modes for the error distribution: one approximately at $80\,\mu$m, and one at about $54\,\mu$m. The plot on the right shows a partition used to make a rough estimation of the fraction that corresponds to each
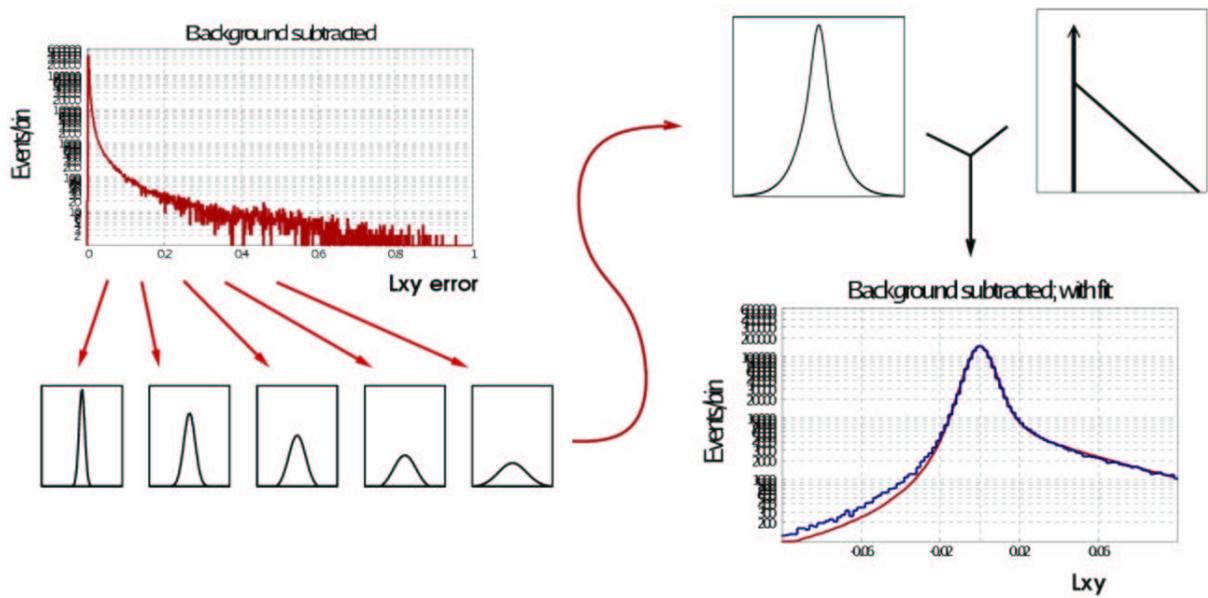
Figure 54: *Evaluation of the resolution function for $L_{xy}(J/\psi)$. The sideband-subtracted distribution of L_{xy} errors weights Gaussian components of a function that is convolved with a model for the J/ψ signal (a delta function plus a single exponential). The resulting function, with floating parameters, is used to fit the sideband-subtracted L_{xy} distribution.*
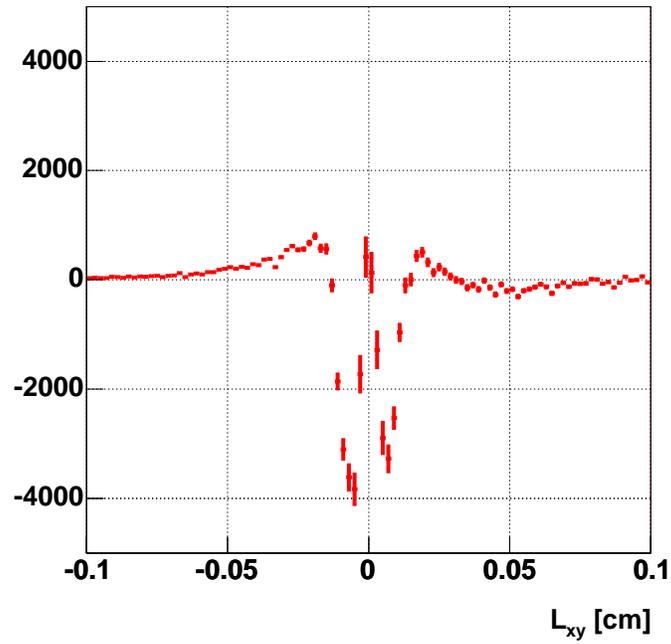
Figure 55: *Residuals of the model.*
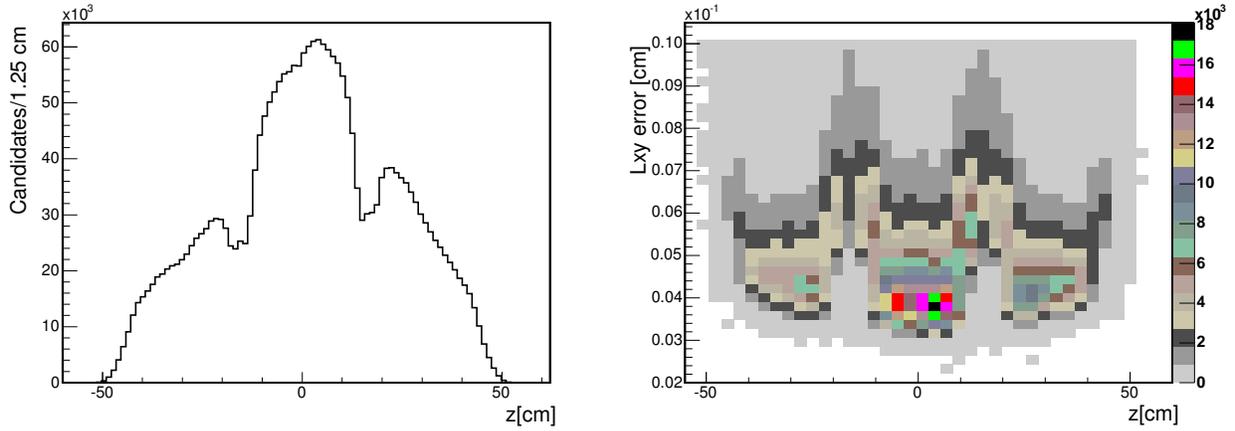
Figure 56: *z distribution from inclusive J/ψ 's (left) and σ(L_xy) vs z of the J/ψ (right)*

error (85% to 54 μm and 15% to $\sim$ 80 μm). Using these estimations, the main features of the distribution of residuals can be reproduced with a simple model: fitting *a single Gaussian* to the sum of two distributions with the widths and fractions corresponding to the two error components. Figure 58 shows the residuals for this toy model. Notice the similarity with figure 55.

In order to account for this, we add a second Gaussian component to the resolution function, and allow the fit to determine separately the mean value and the scale factor of each component. Figure 60 shows the effect of this on the distribution of residuals. Note especially the effect on the long-lived tails. Figure 59 shows the comparison between data and model for the two resolution function models (right: one-component resolution function; left: two-components).

The resulting resolution function has two components: one enters with a relative weight of 92%, a scale factor of 1.08 and an offset of 6 μm; the second component has a weight of 8%, a scale factor of 3.56 and an offset of 362 μm. This can be interpreted as saying that some fraction of the time (8%) the errors are strongly underestimated.
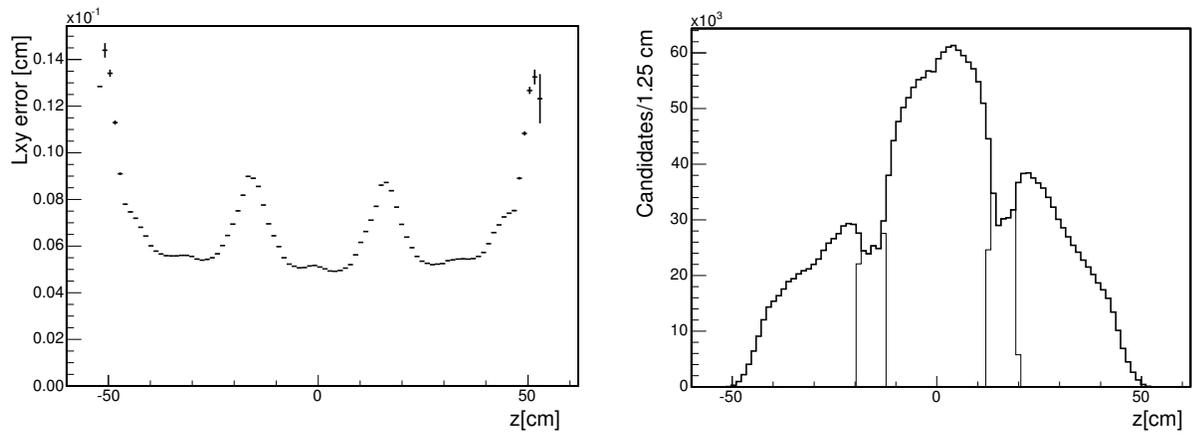
Figure 57: *L_{xy} error distribution as a function of the z position of the J/ψ (left), and partition in z regions used for the estimation of the relative weight of each error component for the toy model.*
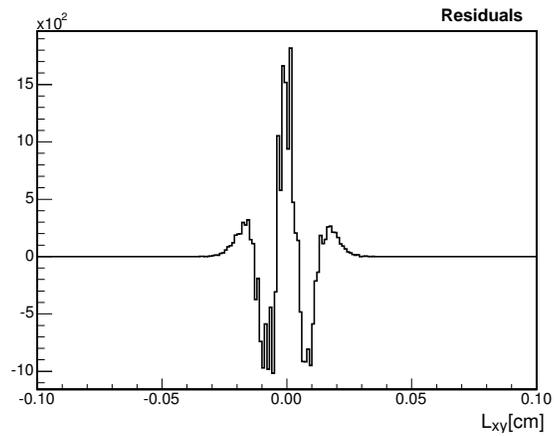

Figure 58: *Distribution of residuals obtained using a toy model.*
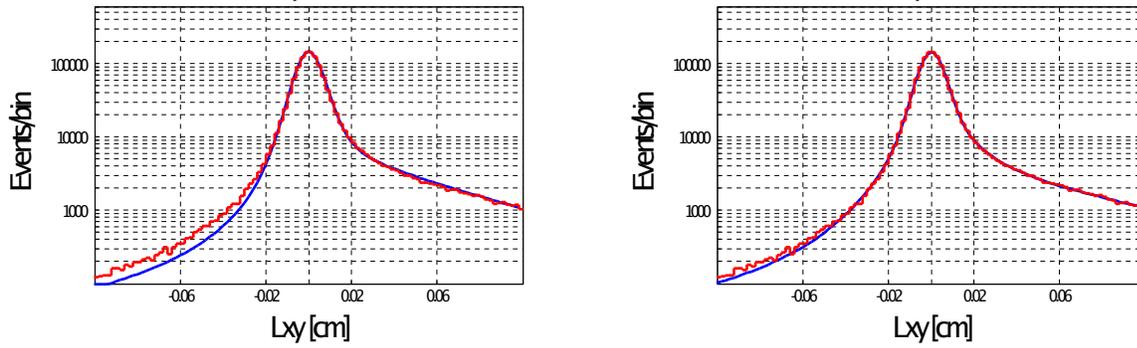
Figure 59: *Right: Data (histogram) and fitted model (line) for a single-Gaussian resolution function; Left: double-Gaussian resolution function.*
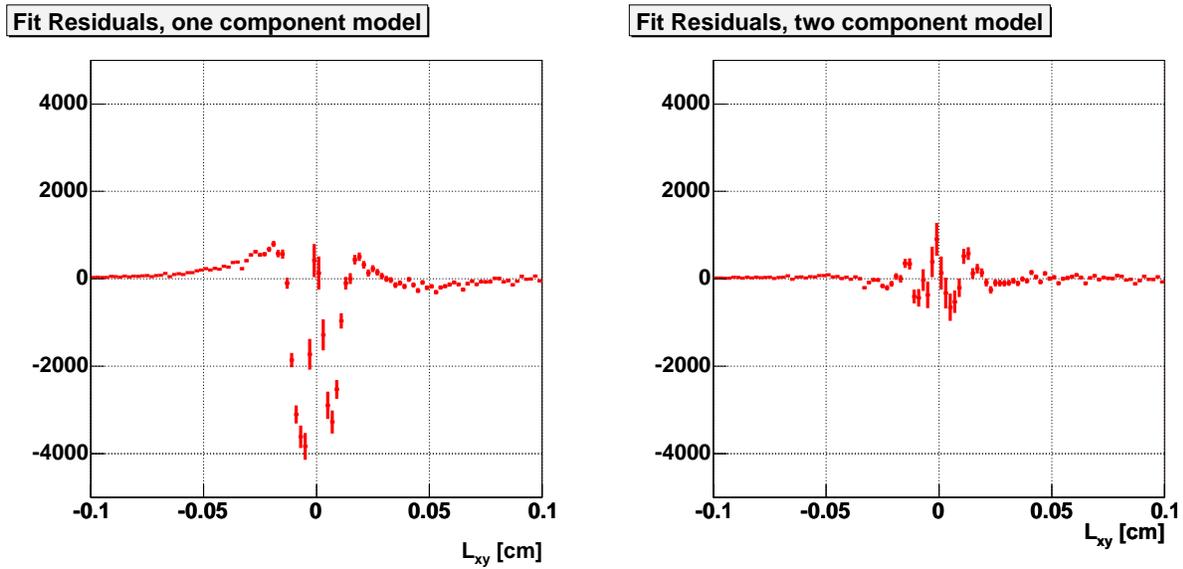

Figure 60: *Distribution of residuals before and after the introduction of a second Gaussian component in the resolution function.*

Table 17: *Effect of using the resolution function obtained from the inclusive J/ψ sample on: the four channels, each of the two species, and the lifetime ratio.*

Quantity	Shift
$\tau(J/\psi K^+)$	-2.5 μm
$\tau(J/\psi K^{*+})$	-3.4 μm
$\tau(J/\psi K^{*0})$	-2.2 μm
$\tau(J/\psi K_s^0)$	-4.8 μm
$\tau(B^+)$	-2.6 μm
$\tau(B^0)$	-2.9 μm
τ_+/τ_0	0.0014

5.2.2 Evaluation of the resolution function systematic uncertainty

We use the obtained model to evaluate the systematic uncertainty due to the use of different resolution function models on the lifetime measurements. Table 17 shows, for each channel, the shifts induced in the fitted $c\tau$ by the use of the J/ψ resolution function in our nominal fit.

We shall take, as a systematic error on the lifetime on each mode and on the the combination of modes for each channel, the maximum observed variation, 5 μm; and as a systematic error on the lifetime ratio, 0.0014. The latter is practically a negligible error on the ratio, and shows the benefit of using the J/ψ vertex in all modes.

5.3 BACKGROUND MODEL

We estimate this systematic uncertainty by adding an extra background component to our nominal fitting model. As this extra component, we used: (a) a flat distribution, and (b) Gaussian distributions with fixed widths varying from 100 μm to 2mm.

Table 18: *Uncertainties assigned to the background model. Obtained as the maximum variation among fitting models including different background components.*

Mode	Max $\Delta c\tau$ (μm)
$J/\psi K^+$	1
$J/\psi K^{*+}$	17
$J/\psi K^{*0}$	7
$J/\psi K_s^0$	9

Since the behavior of the background is different for each mode, we assign a different systematic error for each channel, obtained as the maximum difference among the values obtained for all background models. Table 18 shows the error assigned to each mode.

The effect of the background model uncertainty on each meson's lifetime is obtained by varying each channel's lifetime by the corresponding error shown in table 18, and combining the measurement (as described in section 4.4) with the other mode available for the same meson (for example, altering the lifetime of $J/\psi K_s^0$ by $9\,\mu$m and combining it with an unaltered $J/\psi K^*$ to obtain the B^0 lifetime). Error matrices are not altered for this evaluation.

The effect of this variation is shown in table 19. The systematic uncertainty from this source on the ratio is obtained from the total uncertainties shown for each mode, with a result of 0.014.

5.4   SUMMARY

Table 20 summarizes the final systematic uncertainties. On the individual channels, our largest systematic uncertainty is the background model, which mainly affects the channels with lower statistics ($J/\psi K^{*+}$ and $J/\psi K_s^0$). However, when combining the measurements, the lower statistical weight of these modes limits the effect of their higher errors, making the

Table 19: *Effect of the background model on each meson species' combined lifetime*

Mode	Shift induced (μm)	
	$\Delta c\tau(B^+)$	$\Delta c\tau(B^0)$
$J/\psi K^+$	0.9	-
$J/\psi K^{*+}$	1.4	-
$J/\psi K^{*0}$		4.7
$J/\psi K^0_s$	-	3.0
Total	1.7	5.6

background model systematic comparable to the one coming from the resolution function. As for the ratio, the effect of resolution function is strongly reduced by the fact that it affects all channels in a similar way, nearly canceling in the ratio, and leaving the background model as the main uncertainty in this case.

Table 20: *Summary of systematic uncertainties on measured $c\tau$.*

	$J/\psi\,K^+$	$J/\psi\,K^{*+}$	$J/\psi\,K^*$	$J/\psi\,K_s^0$	B^+	B^0	τ_+/τ_0
Alignment	$2\,\mu$m				$2\,\mu$m		0
Resolution function	$5\,\mu$m				$5\,\mu$m		0.0014
Background model	$1\,\mu$m	$17\,\mu$m	$7\,\mu$m	$9\,\mu$m	$2\,\mu$m	$6\,\mu$m	0.014
Total	$5\,\mu$m	$18\,\mu$m	$9\,\mu$m	$10\,\mu$m	$6\,\mu$m	$8\,\mu$m	0.014

6.0   CONCLUSION

We have measured the lifetimes of four exclusive decays of B^+ and B^0 mesons (two each) to a J/ψ and a Kaon. The results are shown below and in figure 61. The central values and statistical uncertainties are taken from the values for simultaneous mass-lifetime fits using the J/ψ vertex listed in table 12, and the systematic uncertainties are taken from table 20.

$$c\tau(B^+ \to J/\psi\,K^+) = 499 \pm 13(stat) \pm 5(syst)\mu m$$

$$c\tau(B^+ \to J/\psi\,K^{*+}) = 500 \pm 42(stat) \pm 18(syst)\mu m$$

$$c\tau(B^0 \to J/\psi\,K^*) = 460 \pm 18(stat) \pm 9(syst)\mu m$$

$$c\tau(B^0 \to J/\psi\,K^0_s) = 420 \pm 25(stat) \pm 10(syst)\mu m$$

We combine these four measurements into two lifetime measurements, for B^+ and B^0, according the the procedure described in Section 4.4. The systematic errors are combined in the following way. The error coming from the resolution function is propagated by noting the change in combined lifetime produced by a simultaneous variation of the resolution function for both channels. This error is then added in quadrature with the systematic errors from alignment and from the background models in both channels, assumed to be independent. The results are:

$$c\tau_{B^+} = 499 \pm 12(stat) \pm 6(syst)\mu m$$

$$\tau_{B^+} = 1.66 \pm 0.04(stat) \pm 0.02(syst)ps$$

$$c\tau_{B^0} = 446 \pm 15(stat) \pm 8(syst)\mu m$$

$$\tau_{B^0} = 1.49 \pm 0.05(stat) \pm 0.03(syst)ps \ .$$
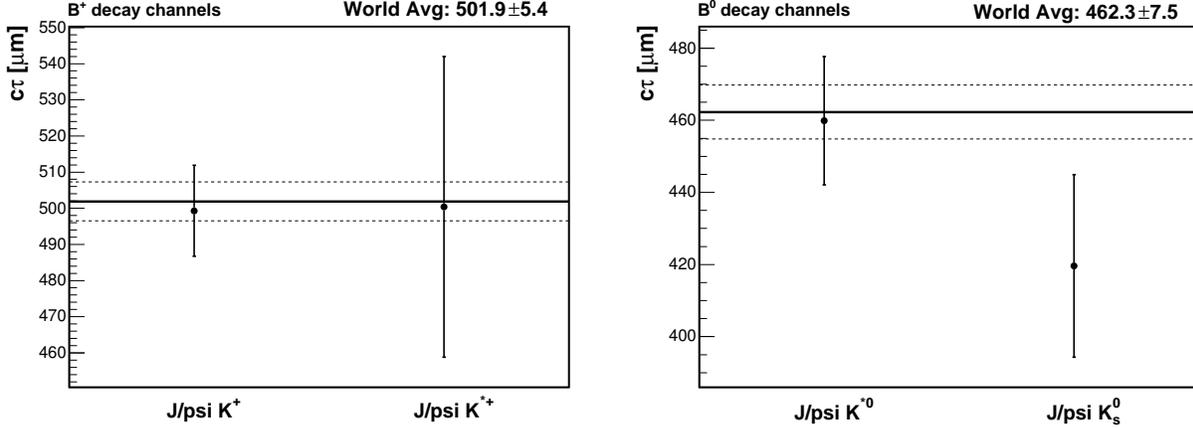
Figure 61: *Measured $c\tau$ for B^+ and B^0 decay modes. Errors are statistical only. The dashed lines indicate the uncertainty on the world averages.*

Finally, we calculate the ratio of lifetimes. The systematic error due to the resolution function is obtained as the change in the lifetime ratio produced by a simultaneous variation of the resolution function for all four channels. The systematic error from alignment is considered to cancel in the ratio. The systematic errors from the background model in each channel are added in quadrature with each other. The final systematic error on the lifetime ratio comes from adding in quadrature the systematic errors from the resolution function and from the background model. This procedure yields

$$\tau_{B^+}/\tau_{B^0} = 1.119 \pm 0.046(stat) \pm 0.014(syst),$$

which differs from the world average of 1.083 ± 0.017 by less than 1σ.

Figure 62 shows a comparison between this measurement of the lifetime ratio and the PDG value; also shown are the theoretical predictions obtained in the framework of the HQE at leading order (LO) and next to leading order (NLO), which have the following values [15]:

LO: 1.041 ±0.040 ±0.013 ,

NLO: 1.053 ±0.016 ±0.017 .

All of them are compatible with each other. Both experimental values are higher than the theoretical predictions, but the difference is small compared to the errors involved. Overall, our measurement is in agreement with the HQE.
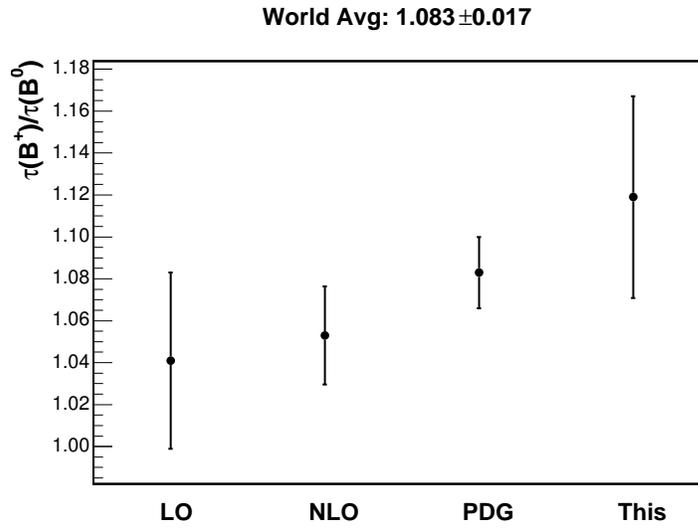
Figure 62: *Comparison among this measurement of $c\tau(B^+)/c\tau(B^0)$, its current PDG value and predictions from the HQE*

By the end of Run II, with approximately ten times as much data available than what is used in this measurement, the statistical error can be expected to be, roughly, a third of the error obtained in this measurement. On the other hand, since the error on the theoretical prediction is inherited, in part, from experimental measurements, it is reasonable to expect that it will also decrease, hence allowing a stronger test of the validity of the HQE.

APPENDIX

K_S^0 AS A TRACK

The reconstruction of exclusive states including a long lived particle (K_s^0 in our case, for both $B^+ \to J/\psi\, K^{*+}$ and $B^0 \to J/\psi\, K_s^0$) needs special attention.

As mentioned in section 3.4, silicon hits are dropped from each track of the K_s^0 candidate when they are below the radius of intersection of the COT parents of both tracks. This reduces the average number of silicon hits on the K_s^0 daughters, potentially amplifying the effect of any remaining misalignment between COT and SVX.

This is especially relevant for these states because their reconstruction can in principle make use of a pointing constraint in the B vertex fit. This constraint can significantly reduce the combinatorial background, loosing only a small fraction of the signal candidates; however, it can also bias the measurement if the COT K_s^0 do not point correctly back to the corresponding J/ψ vertex.

In order to validate the possible use of a pointing constraint in the reconstruction of these states, we reconstructed inclusive $K_s^0 \to \pi^+\pi^-$ and looked at their behavior "as tracks", meaning that we looked at the quantities that may affect their use in constraining a multitrack vertex fit. When adding a displaced vertex with a mean radius on the order of 10 cm to a J/ψ vertex located inside the beampipe (and within 3mm of the beamline), the tracks forming the displaced vertex act on the fit in a way that can be regarded as the addition of a single extra track (or "pseudotrack"). To evaluate the quality of this "pseudotrack", we studied its missed distance d_0 with respect to the primary vertex; for K_s^0 candidates, d_0 is

defined just as it was for tracks in equation (2.4), as

$$d_0 \equiv \frac{\hat{z} \cdot (\vec{r} \times \vec{P_T})}{\left|\vec{P_T}\right|},$$

where $\vec{P_T}$ and r now correspond to the K_s^0 and not to a single track. d_0 and its error are calculated from the fitted vertex position and momentum (and their uncertainties); from the former quantities, we obtain the d_0 significance (i.e., d_0 divided by its error).

We used a sample of events reconstructed with the same version of basic reconstruction than the one used for our exclusive B decays. The sample was divided into bins of $L_{xy}(K_s^0)$ and $\phi(K_s^0)$. In each bin, background subtracted distributions were obtained for the d_0 significance of the K_s^0 pseudotrack, and they were fit to a Gaussian plus a symmetric exponential (with the same mean as the Gaussian, which is allowed to float in the fit). To clean the K_s^0 signal, cuts were applied on the two track vertex probability (0.1%), and on the d_0 significance of both pion tracks (greater than 3).

For the sideband subtraction, the mass distribution (for each bin of $L_{xy}(K_s^0)$ or $\phi(K_s^0)$) was fit using a single Gaussian over a (second order) polynomial background. Signal and sideband regions were defined in each bin using the corresponding fitted mean μ and width σ; the signal region is defined as $|m - \mu| < 3\sigma$, and the sidebands are defined as all events with $|m - \mu| > 4\sigma$, but inside the mass window (0.45 GeV to 0.55 GeV).

Figure 63 shows the mass distributions and the corresponding fits for the bins of $L_{xy}(K_s^0)$. The first twelve correspond to 1 cm bins, covering from 0 to 12 cm; the last eight cover 5cm each, going from 12 to 52 cm. This figure (and the next) are included to allow the reader get a feeling of the signal and sideband regions used, as well as the S/N ratio of the K_s^0 candidates used.

Figure 64 shows the mass distributions in bins defined by the azimuthal angle of the K_s^0 momentum. The upper left bin starts at $\phi = -\pi$, and the lower right corner corresponds to $\phi = \pi$.

Figure 65 shows the raw d_0 error as a function of $L_{xy}(K_s^0)$. As expected, the estimated error is larger in the region between COT and SVX (L00 and ISL were not used for this study, since they are not used in the B lifetime analysis).
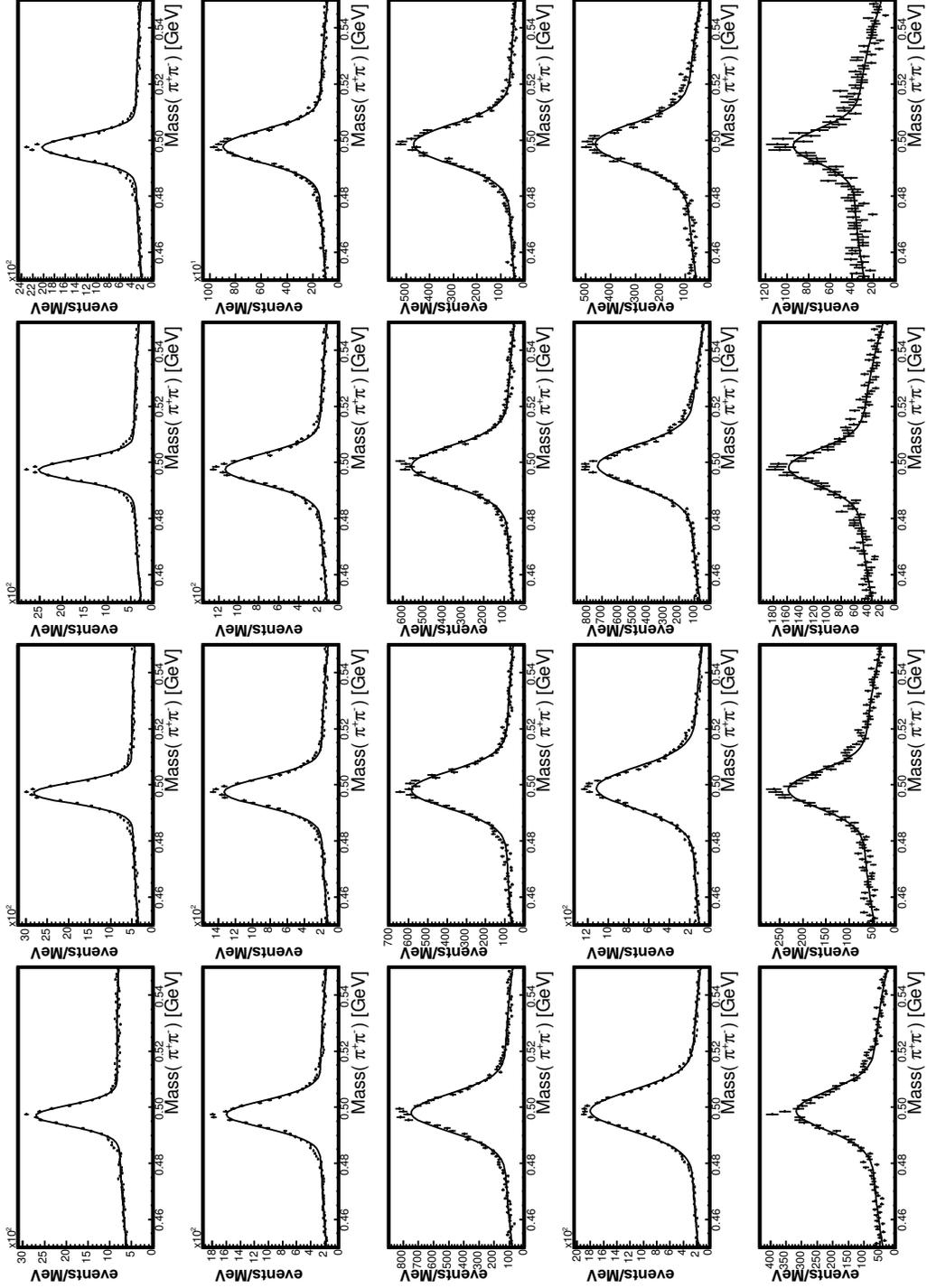
Figure 63: K_s^0 mass distributions in bins of $L_{xy}(K_s^0)$. The first twelve bins correspond to L_{xy} between 0 and 12 cm (1 cm each). The remaining eight correspond to 5 cm bins, covering from 12 to 52 cm.
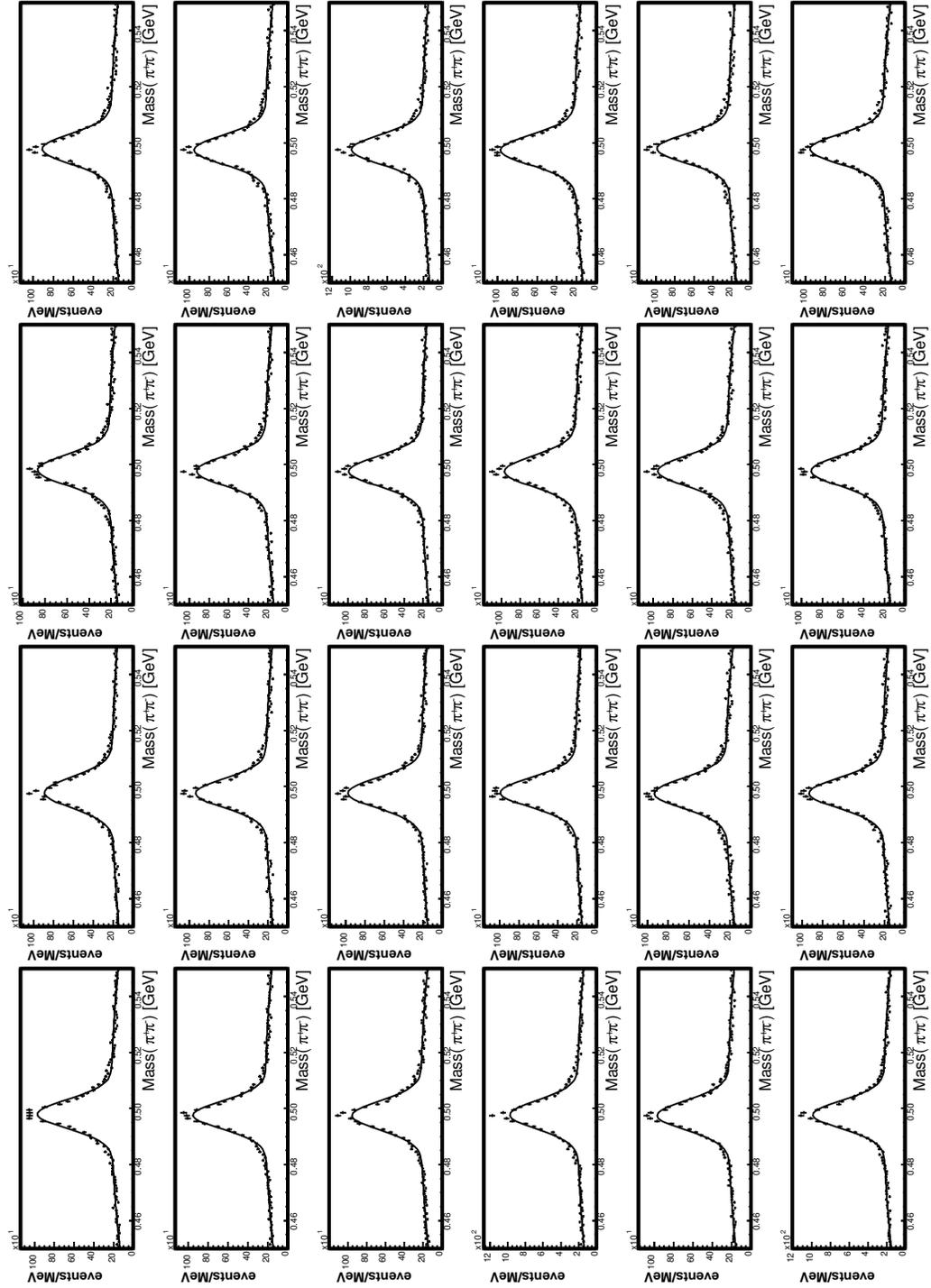
Figure 64: K_s^0 mass distributions in bins of $\phi(K_s^0)$. Upper left: $-\pi$; lower right: π

135
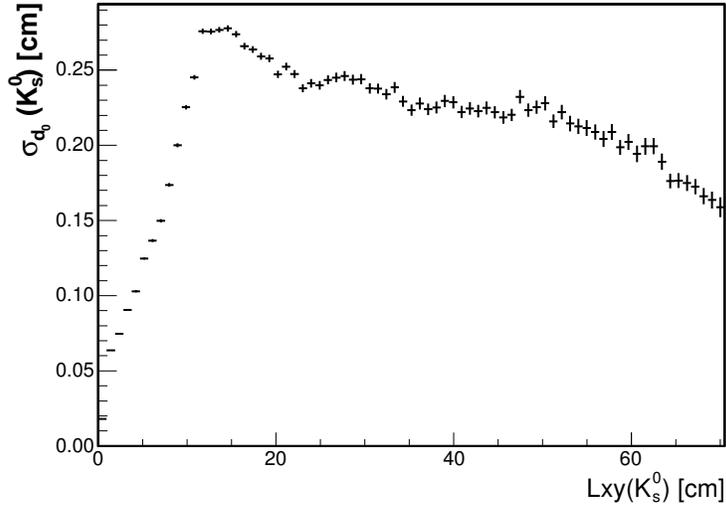
Figure 65: d_0 *estimated error for inclusive* K_s^0 .

In figure 69, the sideband subtracted distributions of the K_s^0 d_0 significance are shown in blue. Sideband distributions are shown in light brown. The horizontal axis goes from -10 to 10 (dimensionless, since it is a pull) in all plots. They all have the same vertical scale. The increase in the number of events between bins 12 (last in third row) to 13 (first in fourth row) is due to the change in bin size (1 cm vs 5 cm) mentioned above.

Figure 66 summarizes the behavior of the width and mean of the d_0 significance as a function of $L_{xy}(K_s^0)$. Each point corresponds to the fit value of the width(mean) of one bin in figure 69. The width varies from 0.8, in the region between SVX and COT, to 1.0, when the K_s^0 has decayed inside SVX. These values are not too far from 1.0 so, although they can in principle affect the measurement via a cut on χ^2, their effect on the pointing need not be large. The mean values, however, vary from -0.05 to 0.20, which is a significant variation and can affect the reconstructed quantities if a pointing constraint is imposed. Most of it occurs within SVX. Few B candidates are reconstructed with $L_{xy}(K_s^0)$>30cm, so the variation may be limited to up to 15% of the d_0 error (from -0.05 to 0.10).

Also important is the behavior of the K_s^0 pseudotrack's d_0 significance as a function of ϕ (the direction of flight of the K_s^0). Figure 70 shows the sideband subtracted distributions (in blue), sidebands (light brown) and fits (solid lines) in bins of ϕ. The fit values for widths
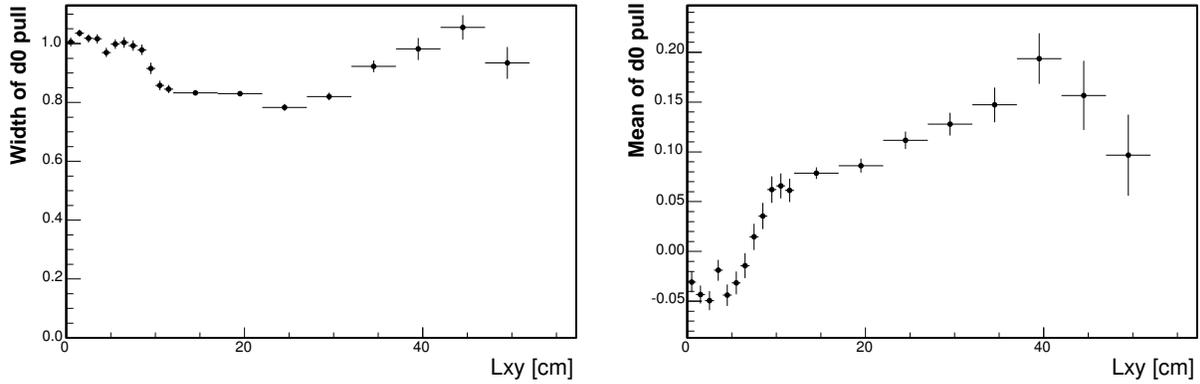
Figure 66: *Mean and width of the pulls of $d_0(K_s^0)$ (dimensionless) vs. $L_{xy}(K_s^0)$ (in cm).*

and means are plotted in figure 67 for all bins in ϕ. The width stays constant (at about 0.9) for all values of ϕ, but the mean shows a clear departure from 0. There are big shifts at about -2 and 1.5 radians, which correspond to opposite sides of the detector, and they are all shifted up (by about 0.03).

The effect is more pronounced for K_s^0 decaying out of SVX. Figure 68 shows the behavior of the mean values for K_s^0 with L_{xy}>12cm. They have a larger shifts at -2 and 1.5 radians, and the overall shift is also larger, resulting in a total range of close to 40% of the statistical error for the shift in $d_0(K_s^0)$.

From this, we concluded that the K_s^0 pseudotrack should not be used to constrain the multitrack vertex fit in the reconstruction of B mesons. This is the reason we do not impose pointing constraints on the B vertex fits. The procedure described in section 3.4.3 is used to work around this problem.
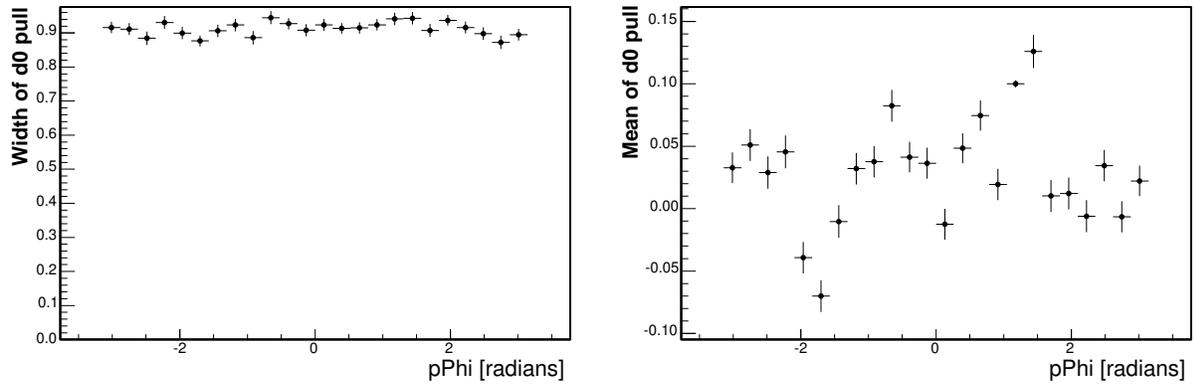
Figure 67: *Mean and width of the pulls of $d_0(K_s^0)$ (dimensionless) vs $\phi(K_s^0)$ (in radians).*
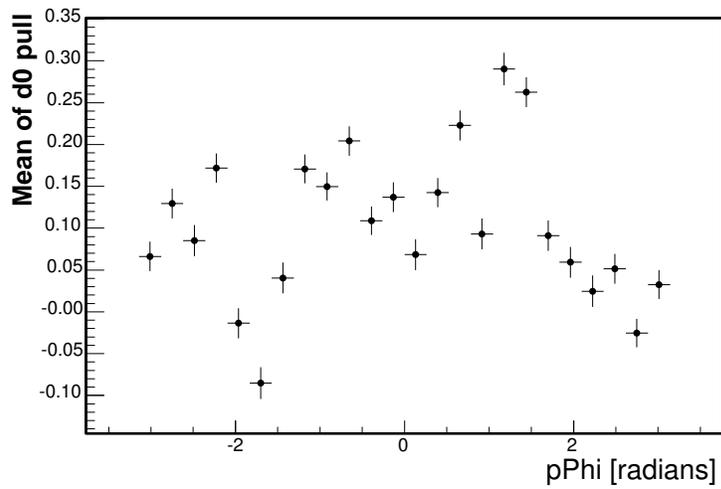
Figure 68: *Mean of the K_s^0 d_0 significance vs. ϕ for $L_{xy}{>}12$ cm*

138
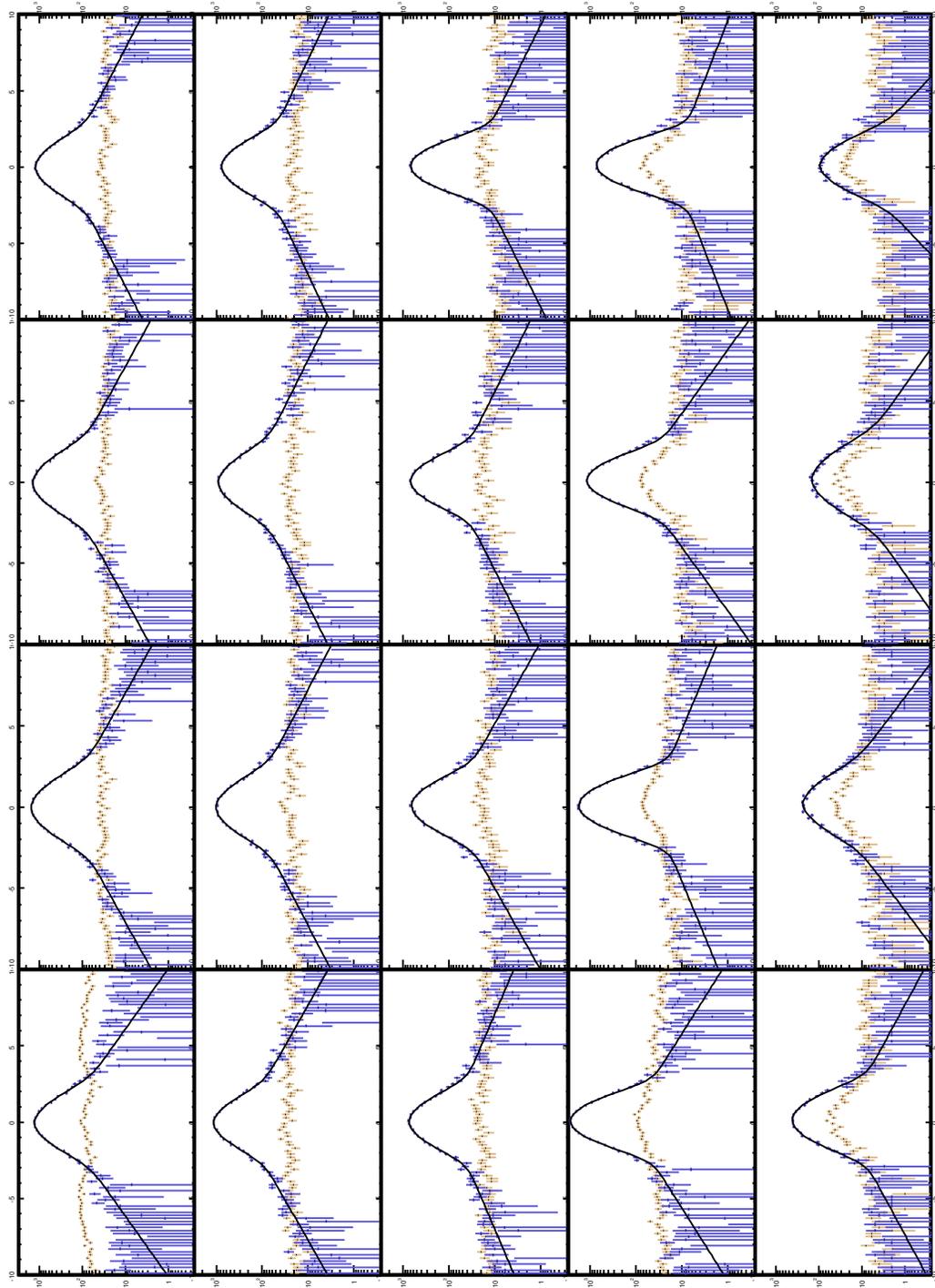
Figure 69: K_s^0 d_0 significance (linear scale) in bins of $L_{xy}(K_s^0)$ (log scale). d_0 significance is shown from -10 to 10 (dimensionless). All plots have the same vertical scale (logarithmic, from 0.25 to 3000).
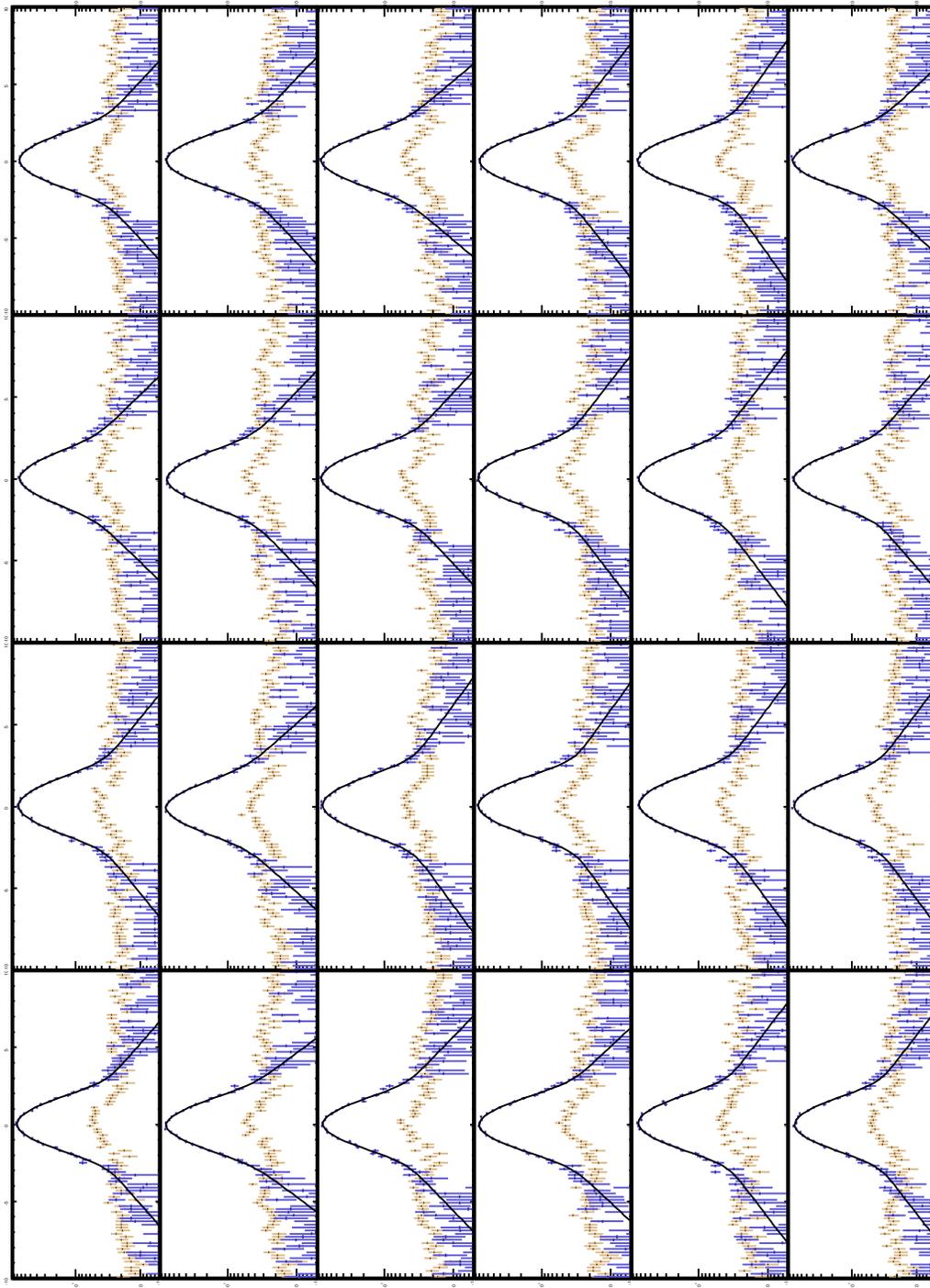
Figure 70: K_s^0 d_0 significance (linear scale) in bins of $\phi(K_s^0)$ (log scale). d_0 significance is shown from -10 to 10 (dimensionless). All plots have the same vertical scale (logarithmic, from 0.25 to 960).

140

BIBLIOGRAPHY

[1] S. L. Glashow. Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1962.

[2] S. Weimberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[3] E. D. Commins and P. H. Bucksbaum. *Weak Interactions of Leptons and Quarks.* Cambridge University Press, 1983.

[4] T. Nakano et al. Evidence for a narrow S= +1 baryon resonance in photoproduction from the neutron. *Phys. Rev. Lett.*, 92:012002, 2003.

[5] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413, 1957.

[6] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the K_0^2 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[7] G. Buchalla. Heavy Quark Theory. *Lectures from the 55th Scottish Universities Summer School in Physics*, 2001. St. Andrews, Scotland: August 7–23, 2001. arXiv:hep-ph/0202092.

[8] A. Heiss *et. al.* A new measurement of exclusive B lifetimes using run 1 data. *CDF internal note 5331*, 2001.

[9] D. Boutigny et al. (The BABAR Collaboration). BABAR Technical Design Report. 1995. SLAC-R-0457, SLAC-R-457, SLAC-0457, SLAC-457, SLAC-R-95-457, Mar . 1995. 618pp.

[10] D. R. Marlow. Status of KEKB and Belle. *AIP Conf. Proc.*, 424:161, 2004.

[11] K. Hagiwara et al. (Particle Data Group). Review of Particle Physics. Phys. Rev. D, 66:010001, 2002. URL: http://pdg.lbl.gov.

[12] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[13] K. M. Ecklund. *Proceedings of the International Europhysics Conference in High Energy Physics.* Aachen, Germany, 2003. arXiv:hep-ex/0312053v1, July 2003.

[14] N. Uraltsev. Heavy Quark Expansion in beauty and its decays. 1997. Lectures from the "International School of Physics Enrico Fermi", Varenna: July 7–18, arXiv:hep-ph/9804275v2.

[15] M. Beneke, G. Buchalla, C. Greub, A. Lenz, and U. Nierste. The $B^+ - B^0_d$ lifetime difference beyond leading logarithms. *Nuclear Physics B*, 639:389–407, 2002.

[16] K. Anikeev et al. B physics at the Tevatron. 2001. FERMILAB-Pub-01/197.

[17] I. I. Bigi. The lifetimes of heavy flavour hadrons - a case study in quark-hadron duality. 1999. UND_HEP-99-BIG 07.

[18] J. P. Marriner et al. Tevatron luminosity upgrade project. *6th European Particle Accelerator Conference (EPAC 98)*, 1998. Stockholm, Sweden, 22–26 June, 1998.

[19] R. Blair et al. The CDF II Detector Technical Design Report. 1996. FERMILAB-Pub-96/390-E, November, 1996.

[20] C. W. Schmidth. The Fermilab 400-MeV Linac upgrade. 1993. FERMILAB-CONF-93/111.

[21] J. Marriner. Stochastic Cooling Overview. 2003. FERMILAB-CONF-03/158.

[22] C. E. Clayton and P. Muggli (editors). Advanced accelerator concepts. *Proceedings of the 10th Workshop AAC 2002*, 2002. Mandalay Beach, Oxnard, USA. June 22-28, 2002.

[23] F. Abe et al. (the CDF Collaboration). The CDF detector: an overview. *Nucl. Instr. Meth.*, A271:387–403, 1988.

[24] F. Abe et al. (the CDF Collaboration). Measurement of the W-boson mass. *Phys. Rev.*, D52:4784–4827, 1995.

[25] R. Wagner et al. CDF Central Outer Tracker. *CDF internal notes 6227, 6267*, (2002,2003).

[26] W. Riegler et al. COT detector physics simulations. *CDF internal Note 5050*, 2000.

[27] A. Sill. *Nucl. Instrum. Meth.*, A447:1, 2000.

[28] S. Leone and others (the ISL group). The Intermediate Silicon Layers (ISL) detector for the Collider Detector At Fermilab. *CDF internal Note 5129*, 1999.

[29] The CDF II Collaboration. Proposal for enhancement of the CDF II detector: "an inner silicon layer and a time of flight detector". 1998. Fermilab-Proposal-909, presented to the Fermilab Director and PAC, October, 1998.

[30] M. Bishai et al. An SVX3D Chip User's Companion. *CDF internal Note 5062*, 1999.

[31] A. Bardi et al. *Nucl. Instrum. Meth.*, A409:658, 1998.

[32] W. Ashmanskas et al. The CDF Silicon Vertex Tracker. FERMILAB-CONF-00/238-E.

[33] W. Ashmanskas et al. The CDF Silicon Vertex Trigger. *Nucl. Instrum. Meth.*, A518:532–536, 2004. FERMILAB-CONF-03/198-E.

[34] W.Ashmanskas et al. Performance of the CDF Online Silicon Vertex Tracker. *IEEE Trans. Nucl. Sci.*, 49:1177–1184, 2002.

[35] J. Antos et al. *Nucl. Instrum. Meth.*, A360:118, 1995.

[36] J. Antos et al. *Nucl. Instrum. Meth.*, A383:13, 1996.

[37] M. Garcia-Sciveres et al. *Nucl. Instrum. Meth.*, A435:58, 1999.

[38] P. Gatti. *Performance of the New Tracking System at CDF II.* 2001. FERMILAB-THESIS-2001-23. February 2001. PhD Thesis.

[39] K. A. Bloom et al. (the CDF-II Collaboration). Track reconstruction for the CDF Silicon tracking system. 1999. FERMILAB-CONF-98–370-E, Apr 1999. 8pp. Also published in the proceedings of International Conference on Computing in High-Energy Physics (CHEP 98), Chicago, IL, 31 Aug - 4 Sep 1998.

[40] T. Dorigo (for the CDF collaboration). The muon system upgrade for the CDF II experiment. *Nucl. Instrum. Meth.*, A461:560–562, 2001. FERMILAB-CONF-00-176-E, Aug 2000. 3pp. Prepared for 8th Pisa Meeting on Advanced Detector: Frontier Detectors for Frontier Physics, La Biodola, Isola d'Elba, Italy, 21-25 May 2000.

[41] C. Grozis et al. The Time-of-Flight detector for CDF. *Nucl. Phys. Proc. Suppl.*, 47:344, 2001.

[42] C. Grozis et al. The Time-of-Flight detector for CDF. *Int. J. Mod. Phys.*, A16S1C, 2001. 1119.

[43] G. Gomez for the TOF group. The CDF-II Time of Flight detector. *CDF internal Note 6258*, 2003.

[44] S. Klimenko et. al. CDF Run II luminosity monitor. *CDF internal note 5119*, 1999.

[45] David Saltzberg. Run-II trigger table and datasets plan. *CDF internal Note 4718*, 1998.

[46] J. Gartner et al. The performance of the XFT in run II. *CDF internal Note 6766*, 2003.

[47] E. J. Thompson et. al. Online track processor for the CDF upgrade. *IEEE Transactions on Nuclear Science*, 49 No. 3, 2002.

[48] L. Hollaway and J. Strologas. Determination of the XTRP maps. *CDF internal Note 4146*, 1997.

[49] J. Boudreau, K. Carrell, J. Cranshaw, J. P. Fernandez, L. Flores-Castillo, V. Papadimitriou, and M. Spezziga. Measurement of exclusive B lifetimes in the modes: Jpsi K+, Jpsi K*+, Jpsi K*0, and Jpsi Kshort. *CDF internal note 6387*, 2004.

[50] J. Antos et al. The CDF Run II offline computer farms. 2001. CHEP 2001, Beijing, September 2001.

[51] M.Bishai et. al. COT and SVX II tracking performance. *CDF internal note 5391*, 2000.

[52] R. Madrak. Correcting the COT covariance matrix. 2002. Presentation in the CDF Tracking Group meeting. http://harv11.fnal.gov/∼madrak/trk.ps.

[53] J. Marriner. Secondary vertex fit with mass and pointing constraints (CTVMFT). *CDF internal note 2724*, 1994.

[54] J. Boudreau, K. Carrell, L. R. Flores-Castillo, J. Cranshaw, V. Papadimitriou, and M. Spezziga. Yields and cut optimization for $B \to J/\psi\, X$, $J/\psi \to \mu^+\mu^-$. *CDF internal note 6226*, 2003.

[55] R. Madrak et al. Measurement of the Λ_b lifetime in the decay mode $J/\psi\,\Lambda$. *CDF internal note 6393*, 2003.

[56] P. Sphicas. A $b\bar{b}$ Monte Carlo generator. *CDF internal note 2655*, 1994. CDF/DOC/BOTTOM/CDFR/2655.

[57] K. Anikeev, Christoph Paus, and Pasha Murat. Description of Bgenerator II. *CDF internal note 5092*, 1999. CDF/DOC/BOTTOM/CDFR/5092.

[58] J. D. Lewis and P. Avery. Cleomc: The CDF interface to the CLEO Monte Carlo (QQ). *CDF internal note 2724*, 1994. CDF/DOC/MONTECARLO/PUBLIC/2724.

[59] K. Anikeev, A. Korn, and Ch. Paus. Generator level dimuon trigger simulation. *CDF internal note 5686*, 2001. CDF/DOC/BOTTOM/CDFR/5686.

[60] L. Demortier and L. Lyons. Everything you always wanted to know about pulls. *CDF note 5776*, 2002.

[61] R. Culbertson. 2004. Private Communication.