Measurement of the B_d^0 lifetime using $B_d^0 \to J/\psi K_S^0$ decays at DØ

The cover illustration is a graphical representation of a particle collision in the DØ detector, viewed from the side of the detector. Special thanks go out to Peter Balm and Nico Keizer, for the cover design and realization.

Measurement of the B_d^0 lifetime using $B_d^0 \to J/\psi K_S^0$ decays at DØ

ACADEMISCH PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE UNIVERSITEIT VAN AMSTERDAM
OP GEZAG VAN DE RECTOR MAGNIFICUS
PROF. MR. P.F. VAN DER HEIJDEN
TEN OVERSTAAN VAN
EEN DOOR HET COLLEGE VOOR PROMOTIES INGESTELDE COMMISSIE,
IN HET OPENBAAR TE VERDEDIGEN
IN DE AULA DER UNIVERSITEIT
OP WOENSDAG 8 DECEMBER 2004
OM 14:00 UUR

door

Paul Wijnand Balm

geboren te Haarlem

Promotores:	Prof. dr. M. W. J. M. Demarteau Prof. dr. F. L. Linde
Co-promotor:	Dr. ir. P. J. de Jong

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The work described in this thesis is part of the research programme of the ‘Nationaal Instituut voor Kernfysica en Hoge-Energie Fysica’ (NIKHEF) in Amsterdam, the Netherlands. The author was financially supported by the ‘Stichting voor Fundamenteel Onderzoek der Materie’ (FOM).

Contents

Introduction

One of the first modern particle physics experiments was performed by Rutherford in 1911. Rutherford discovered that atoms are not the indivisible units that they are named after, but that they are composed of different building blocks. An atom turned out to be an electron cloud containing a tiny kernel, like a fly in a cathedral. In the years after this discovery, a new model of fundamental matter was developed. This postulated protons and neutrons as the fundamental building blocks of the kernel (which Rutherford had called the *nucleus*). Our current understanding goes one step beyond this model. A number of theories have been grouped into what is called the Standard Model. The Standard Model states that protons and neutrons are composed of fundamental components, which have been named quarks. The Standard Model also provides a detailed description of the forces, or interactions, between the (postulated) fundamental particles. This description can be used to make predictions on what happens in particle collisions. The past thirty years, physicists have been performing measurements to test these predictions, because historically, when nature turns out to deviate from model predictions, this provides a clue to a more fundamental theory. While there are good arguments to believe that such a theory exists, these arguments are currently only theoretical. No experimental measurement has been performed thus far, that beyond doubt shows a deviation from Standard Model predictions[1].

The Tevatron accelerator at the Fermi National Accelerator Laboratory in Chicago, U.S.A., has recently been upgraded to increase its collision energy and rate. The experiments at the Tevatron, DØ and CDF, have also been upgraded, to enable them to deal with the higher collision rates. Additionally, the capabilities of the DØ detector have been extended with the addition of a new inner tracking detector. The DØ experiment is currently performing a broad range of measurements, which include measurements of the properties of B-mesons. These measurements have only become possible at DØ after the accelerator and detector upgrade. They are interesting from a theoretical point of view, because they allow the determination of parameters of the Standard Model, as well as the verification, or falsification, of models predicting B-meson lifetimes. This thesis presents a measurement of the lifetime of the B_d^0 meson in the decay to $(J/\psi K_S^0)$ at DØ. In recent years, accurate measurements of this quantity have been performed by other experiments, such as Belle and BABAR. These experiments, however, can only cover a fraction of the measurements that are possible at DØ. In the collection of measurements possible at DØ, the measurement presented in this thesis will serve as an important benchmark.

The presented measurement is the first measurement of the B_d^0 lifetime at DØ in this channel,

[1]The recent observation that neutrino's are massive is considered an extension to the Standard Model, that does not invalidate it.

and this in itself makes the measurement experimentally challenging. Further experimental challenges are posed by the fact that only one in a few million collisions at the Tevatron produces a $B_d^0 \to J/\psi K_S^0$ decay and by the complicated structure of the background events passing through the event selection procedure.

Chapter 1 of this thesis describes in more detail the theory motivating a measurement of the B_d^0 lifetime in the decay to $(J/\psi K_S^0)$ at DØ. The experimental set-up consists of the accelerator complex and the DØ detector, of which chapter 2 provides a description. The vast amounts of data recorded by the DØ detector — from October 2002 to June 2003 approximately 4.3 TB raw data has been recorded, enough to fill more than 5000 cd's — are reconstructed into meaningful objects, like particle trajectories, by the algorithms described in chapter 3. To obtain a more manageable amount of data, a selection of the interesting data is developed in chapter 4. Chapter 5 describes the measurement of the B_d^0 lifetime. In chapter 6 we draw conclusions and provide an outlook to the future.

Chapter 1

Theoretical motivation

We begin by giving a brief description of the theory that motivates the measurement described in this thesis. Our current understanding of fundamental particles and their interactions is described by the Standard Model. The Standard Model was developed in the second half of the twentieth century, as a theoretical conclusion after a large number of experimental discoveries in the decades before. This "conclusion" cannot be regarded as a final answer, though. For example, the Standard Model does not include gravity. Also, it has many external parameters, for which the theory does not provide a value or even motivates a particular range and which have to be determined by experiment. Provided with these parameters though, the Standard Model is extremely accurate. Ever since the conception of the Standard Model, its predictions have been tested in attempts to invalidate it, and to uncover clues of an even more fundamental theory, that does not have the problems that the Standard Model has. So far, no significant discrepancy between Standard Model predictions and nature has been found, but the testing continues, and the work described in this thesis is part of this effort. A more detailed description of the Standard Model is given by Ref. [1].

The next section discusses the features of the Standard Model. The following sections will describe some of the tests of the Standard Model, for which an accurate measurement of the lifetime of B-mesons is a necessary ingredient.

1.1 The Standard Model

According to the Standard Model, matter at its most fundamental level is composed of half-integer spin particles called fermions. There are two types of fermions, quarks and leptons, and each type comes in three families. The fermions contained in the Standard Model are indicated in Table 1.1. Fermions interact through the exchange of force particles, called bosons, which have integer spin. We know of the existence of four fundamental forces: The strong force, the electromagnetic force, the weak force and the gravitational force. Conceptually, the Standard Model is a collection of gauge theories, each theory describing one type of interaction. Gravitation has not yet been written as a gauge theory, and is therefore not included in the Standard Model.

Only the quarks participate in the strong interaction, which confines them to bound states,

family	1	2	3	charge [e]
quarks	u (up)	c (charm)	t (top)	$+\frac{2}{3}$
	d (down)	s (strange)	b (bottom)	$-\frac{1}{3}$
leptons	e (electron)	μ (muon)	τ (tau)	-1
	ν_e	ν_μ	ν_τ	0

Table 1.1: The matter particles, or fermions, in the Standard Model, as they are distributed over the three families. The neutrino's, ν_e, ν_μ and ν_τ, are named after their partners: electron neutrino, muon neutrino and tau neutrino.

called hadrons. Hadrons are composed of a quark-antiquark pair (mesons), or a combination of three quarks or antiquarks (baryons). Recently the LEPS Collaboration reported experimental results regarding the observation of hadrons containing four quarks and one antiquark (called pentaquarks) [2]. The existence of pentaquarks has been confirmed by a multitude of other experiments. The different allowed combinations of quarks and antiquarks have led to the concept of quarks carrying color charge, or just color. There are three different colors (which we label red, green and blue) and hadrons are required to be colorless objects. A quark can combine with an antiquark of the conjugate color (labeled anti-red, etc.) into a colorless hadron, and so can three quarks of three different colors. The strong interaction is mediated by gluons, transferring the color charge.

The electromagnetic and the weak force are unified in a single theory. The electromagnetic force works through the exchange of photons, which are massless and electrically neutral. The weak force is mediated by massive bosons, of which one is electrically neutral, the Z^0, (we speak of "neutral current" interactions) and two are charged, the W^+ and W^- ("charged current" interactions). Charged current interactions can change the flavor of quarks. The strength with which the charged current manifests itself is characterized by, apart from an overall coupling constant, the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3]. The CKM matrix relates the weak eigenstates of quarks, that participate in the weak interaction, to the mass eigenstates, of which observable hadrons are formed. By convention the matrix operates on the charge $-\frac{1}{3}$ quark states:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

Each charge $-\frac{1}{3}$ quark couples predominantly via the charged current interaction to one of the charge $+\frac{2}{3}$ quarks. This is the motivation for grouping these quarks in families, as indicated in Table 1.1. This is also reflected in the magnitude of the CKM matrix elements. The norm of the elements on the diagonal is close to 1, indicating the relatively strong coupling between quarks of the same family. The off-diagonal elements are smaller. In its most general form, the CKM matrix has 9 complex parameters, or 18 real ones. The CKM matrix is required to be unitary: $(V_{CKM} \cdot V_{CKM}^\dagger) = \mathbb{1}$. For three families, this yields nine constraints:

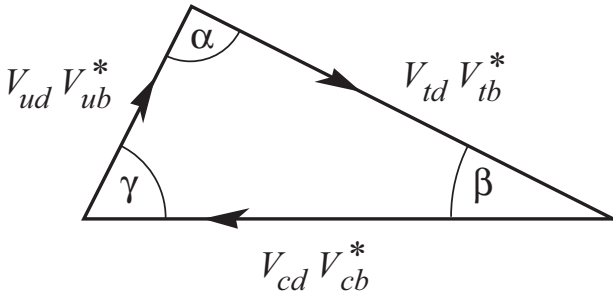


Figure 1.1: The unitarity triangle, resulting from equation 1.2, with $q_2 = d$, $q_3 = b$.

$$\sum_{q_1=\{u,c,t\}} V^*_{q_1 q_2} V_{q_1 q_3} = \delta_{q_2 q_3}; \quad q_2, q_3 \in d, s, b. \tag{1.2}$$

These nine restrictions leave three free real parameters, and six phases. We have the freedom, however, to absorb a phase into each of the six quark states. This would imply six more restrictions, but the CKM matrix is unchanged when all quark states are transformed, so there is one restriction less. The result is three real parameters, and one free phase. Such a limited number of independent parameters calls for a parametrization of the CKM matrix. Widely used is the Wolfenstein parametrization [4]:

$$V_{CKM} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & \lambda^3 A(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & \lambda^2 A \\ \lambda^3 A(1-\rho-i\eta) & -\lambda^2 A & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.3}$$

This parametrization introduces the real parameters λ, A and ρ, and it accommodates the complex phase with an imaginary parameter $i\eta$. The parameter λ is set equal to the element V_{us}, which is, to order λ^3 equal to the sine of the Cabibbo angle, $\sin(\theta_C)$, and has a value of ~ 0.22. The expansion in λ therefore shows clearly, that the elements on the diagonal are close to 1, and that the mixing between neighboring families is suppressed by a factor λ. Finally, the parametrization shows that transitions from family 3 to family 1 are suppressed by even higher orders of λ.

We can interpret the elements of the CKM matrix as vectors in the complex plane. The six equations following from equation 1.2 with $\delta_{q_2 q_3} = 0$ can then be visualized as triangles, yielding three different triangles, and their complex conjugates. The most interesting triangle corresponds to $q_2 = d$ and $q_3 = b$, because it is the only one that has three sides of similar length. This triangle is therefore usually referred to as *the* unitarity triangle. It is shown in Fig. 1.1. If we divide the length of the sides of this triangle by $|V_{cd}V_{cb}^*|$, then in terms of the Wolfenstein parameters, the coordinates of the triangle apex will be $\left(\bar{\rho}, \bar{\eta}\right) = \left(\rho(1-\frac{\lambda^2}{2}),\ \eta(1-\frac{\lambda^2}{2})\right)$.

The observable complex phase in the CKM matrix is associated with CP-violation in the Standard Model, as will be explained in section 1.3. If there is no CP-violation in the Standard Model, this phase will be zero, and so will the area of the unitarity triangle: The angle α will be π, and the other two angles will be zero.

1.1.1 Relation to macroscopic phenomena

The world that we live in is the macroscopic result of the fundamental fermions and forces, as postulated by the Standard Model. All matter that we see around us consists of molecules. A molecule is the smallest amount of a pure substance that still has all the chemical properties of that substance. Molecules are built of atoms, that contain a small nucleus surrounded by a cloud of electrons. For instance, water molecules are built of two hydrogen atoms and one oxygen atom. A hydrogen atom is the simplest of atoms. Its nucleus consists of a single proton, and its electron cloud of a single electron. (Due to the uncertainty principle, a single electron is still a cloud.) The most important reason liquid water sticks together is the electromagnetic interaction between the water molecules. The hydrogen nucleus, a proton, is composed of two up quarks and one down quark, and the reason it stays in one piece is the attraction between these quarks, which is due to the strong force. Finally, the weak force is known in the macroscopic world through for instance radio-activity. Radio-activity results from the transformation of e.g. a down-quark into an up-quark through a weak interaction. This results in a neutron consisting of two d's and one u being transformed into a proton, which contains two u's and one d.

1.2 Production and decay of B-mesons

Using the framework of particles and interactions given by the Standard Model, this section describes the mechanisms responsible for the production and decay of B-mesons.

The Tevatron collides protons and antiprotons at a center of mass energy of 1.96 TeV. B-meson production at the Tevatron is dominated by two strong interaction processes, called "flavor creation" and "flavor excitation". Flavor creation refers to fusion of a quark from the proton and an antiquark from the antiproton, and fusion of two gluons, radiated from the proton and antiproton. Flavor excitation requires the presence of a b- or $\bar{b}$-quark in the initial state, i.e. in the quark sea of the (anti-)proton. Feynman diagrams are shown in Fig. 1.2. The B-mesons are formed by binding each b with a quark or antiquark from the proton remnant, or with a quark drawn from the vacuum. This mechanism explains the high production rates of the B_d^0 and B^+, compared to the B_s^0. The B_d^0 and B^+ are composed of a $\bar{b}$-quark and a d or u, respectively. The B_s contains $\bar{b}s$. The u- and d-quarks are available in the proton remnant, and moreover they can be drawn more easily from the vacuum, due to their lower mass. The production rates of the B_d^0 and B^+ each cover about 40% of the total b-quark production, and the B_s accounts for about 10%. Most of the remaining b-quarks hadronize into Λ_b baryons, composed of a b-, u-, and d-quark.

The decay of B-mesons proceeds through the weak, charged current interaction. To first order, i.e. if the quark masses are neglected, the relative contribution of the different possible processes, $b \to tW$, $b \to cW$, and $b \to uW$ is given by the magnitude of the CKM matrix elements, but the process $b \to tW$ is kinematically forbidden by the mass of the t. Therefore, only processes associated with off-diagonal CKM matrix elements are allowed, and the decay process of B-mesons is suppressed, leading to lifetimes of the order of 1 ps. Different models of B-meson decay exist, some allowing a precise calculation of the lifetime, as will be discussed in

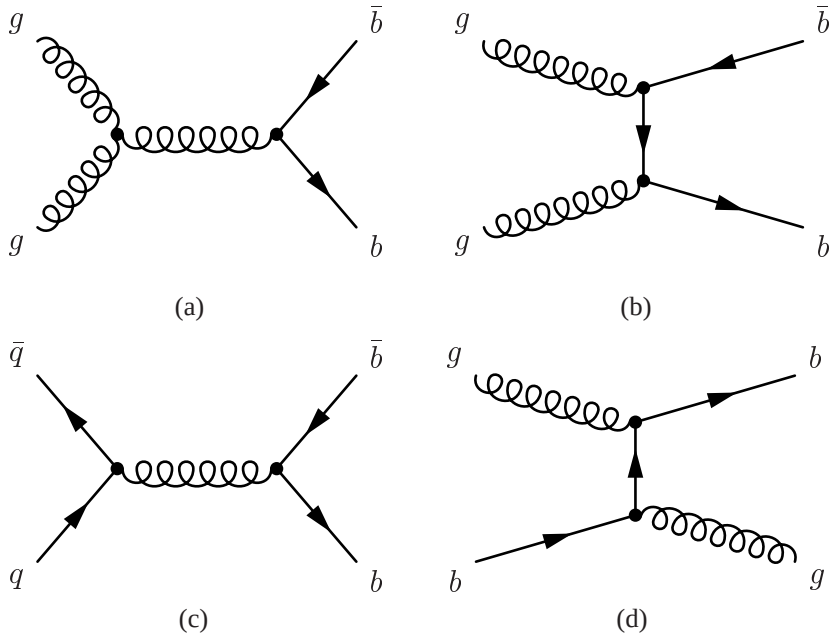


Figure 1.2: Feynman diagrams of the dominant b-quark production processes at the Tevatron. The processes (a), (b) and (c) are referred to as "flavor creation". Process (d) is called "flavor excitation".

the next section.

1.2.1 B-meson lifetimes

The simplest model of B-meson decays assumes the b-quark is entirely undisturbed by the accompanying lighter quark, as if the lighter quark is a passive spectator in the decay. This model is called the "spectator model" (see Fig. 1.3 (a)). According to this model, the lifetime of all B-mesons are equal, as the quark type of the spectator is irrelevant. For all B-mesons except the B_c, this approximation is accurate to the $\approx 10\%$ level (see Table 1.2). The B_c has a significantly shorter lifetime, because for this meson, a weak decay of the c-quark is also possible.

A more accurate description of B-meson decays is given by the Heavy Quark Effective Theory (HQET). The gluons exchanged between the two quarks in the meson and between the final state quarks as shown in Fig. 1.3 (b), which are ignored in the spectator model, have momenta of order $\Lambda_{QCD} \approx 200$ MeV). If the mass of the heavy quark is much larger than this (as is the case with b-quarks, having a mass of ≈ 5 GeV), its momentum is only marginally disturbed by the soft gluon interactions. HQET uses this fact to simplify the description of the meson, as has been described by Neubert [5]. In HQET, the Lagrangian is expressed as an expansion in powers of $1/m_Q$ (with m_Q the mass of the heavy quark). This is commonly referred to as "Operator Product Expansion" (OPE). Using HQET, one can obtain an expansion for the inclusive B-meson decay rate Γ [6]:

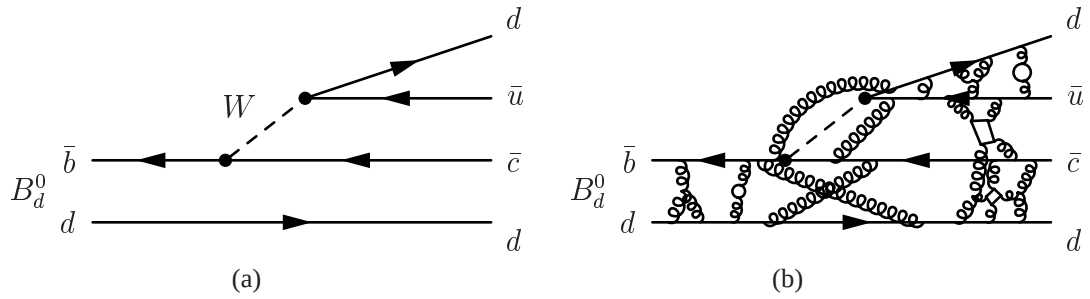


Figure 1.3: Two different pictures of B_d^0 meson decay.

Spectator model	1	
HQET	1.06 ± 0.02	[8]
DØ, March 2004	1.093 ± 0.029	[9]
CDF, February 2004	1.119 ± 0.048	[10]
World average	1.085 ± 0.017	[11]

Table 1.2: Experimental and theoretical results on the lifetime ratio $\tau(B^+)/\tau(B_d^0)$.

$$\Gamma = \Gamma_0 + \frac{1}{m_Q}\Gamma_1 + \frac{1}{m_Q^2}\Gamma_2 + \frac{1}{m_Q^3}\Gamma_3 + \cdots \tag{1.4}$$

The leading term in this expansion, Γ_0, is the decay rate of a free b-quark:

$$\Gamma_0 = \frac{G_F^2 m_b^5}{192\pi^3} \left|V_{ub} + V_{cb}\right|^2, \tag{1.5}$$

with G_F the Fermi constant.Thus, to first order, HQET reproduces the spectator model. The higher-order terms are corrections due to gluon exchange between the quarks. The advantage of this expansion is that the first few terms can be calculated, and their accuracy can be estimated based on the uncertainty on the hadronic parameters that enter it. The same is true for the higher-order terms. Their magnitude can be estimated, so that it is possible to make an estimate of the uncertainty of the HQET prediction of B-hadron lifetimes. This uncertainty estimate, in turn, allows for a meaningful test of the theoretical prediction by comparing it to experimental results. This thesis describes an absolute determination of the lifetime of the B_d^0 meson. The ratio of the lifetime of the B^+ to the lifetime of the B_d^0, $\tau(B^+)/\tau(B_d^0)$, is less prone to systematic experimental and theoretical uncertainties. Table 1.2 lists some recent experimental and theoretical results on $\tau(B^+)/\tau(B_d^0)$. These results are in good agreement. Improved precision will provide a further test of HQET. On the experimental side improvement is expected soon. DØ expects to achieve an accuracy better than 0.5%, based on an integrated luminosity of 2 fb^{-1} [7], which it expects to accumulate over the next few years.

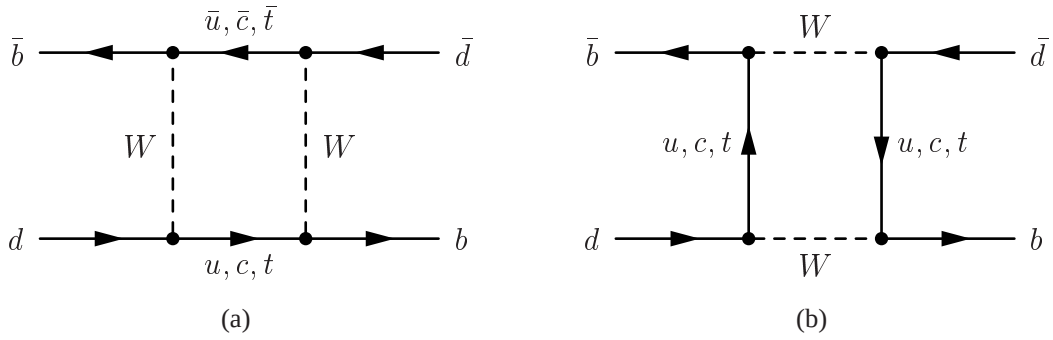


Figure 1.4: Box diagrams, depicting the dominant contributions to mixing of neutral B-mesons.

1.2.2 Mixing in the neutral B-meson system

The weak interaction gives the neutral B-mesons, B_d^0 and B_s^0, an interesting capability beyond simple decay. The exchange of two W-bosons allows these mesons to transform into their own antiparticle. This process is referred to as mixing, because the mesons, initially produced as pure flavor states, will become a mixture of the flavor states B^0 and $\bar{B}^0$ as time elapses. This section will explain the mechanism of mixing, and we will introduce some of the formalism used to describe it. The mixing process is equivalent for the B_d^0 and B_s^0 mesons, but for simplicity here we will describe the B_d^0 meson case.

The most important contributions to the mixing process come from the Feynman diagrams shown in Fig. 1.4, the so-called box diagrams. In these processes, in addition to the two W-bosons, a u-, c-, or t-quark is exchanged, though the t is strongly favored. When a B_d^0 meson is produced, it is a pure $(\bar{b}d)$ flavor state, which we label $|B^0\rangle$. The fact that B_d^0 mesons can mix to $\bar{B}_d^0$ means that this initially pure flavor state will evolve over time, to become a linear superposition of the $(\bar{b}d)$ and $(b\bar{d})$ flavor states, labeled $|B^0(t)\rangle$. The time-evolution of this superposition is described by a time dependent Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} = \mathcal{H}\begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} \equiv \begin{pmatrix} M - \frac{1}{2}i\Gamma & M_{12} - \frac{1}{2}i\Gamma_{12} \\ M_{12}^* - \frac{1}{2}i\Gamma_{12}^* & M - \frac{1}{2}i\Gamma \end{pmatrix}\begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix}. \tag{1.6}$$

We have written M for $M_{11} = M_{22}$ and Γ for $\Gamma_{11} = \Gamma_{22}$. The mass-eigenstates can be found by diagonalizing $\mathcal{H}$ and extracting the eigenvectors. The mass-eigenstates are B_H and B_L (subscripts for heavy and light):

$$\begin{aligned} |B_H\rangle &= p\,|B^0\rangle - q\,|\bar{B}^0\rangle\,, \\ |B_L\rangle &= p\,|B^0\rangle + q\,|\bar{B}^0\rangle\,. \end{aligned} \tag{1.7}$$

These two eigenstates correspond to the eigenvalues:

$$\lambda_\pm = \left(M - \frac{i}{2}\Gamma\right) \pm \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right). \tag{1.8}$$

For the parameters p and q, the normalization requirement applies ($|p|^2 + |q|^2 = 1$), and for the ratio the following holds:

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{1.9}$$

An alternative notation to equation 1.7 is:

$$\begin{aligned} |B_H\rangle &= \frac{(1+\epsilon_B)\,|B^0\rangle - (1-\epsilon_B)\,|\bar{B}^0\rangle}{\sqrt{2(1+|\epsilon_B|^2)}}, \\ |B_L\rangle &= \frac{(1+\epsilon_B)\,|B^0\rangle + (1-\epsilon_B)\,|\bar{B}^0\rangle}{\sqrt{2(1+|\epsilon_B|^2)}}. \end{aligned} \tag{1.10}$$

The relation between q/p and ϵ_B is:

$$\frac{q}{p} = \frac{1-\epsilon_B}{1+\epsilon_B}. \tag{1.11}$$

The mass of the states $|B_L\rangle$ and $|B_H\rangle$ is $\mathrm{Re}(\lambda_\pm)$, and their decay widths are $-2\,\mathrm{Im}(\lambda_\pm)$. Any $|B\rangle$ state can be written as an admixture of the states $|B_L\rangle$ and $|B_H\rangle$, whose time evolution is given by the solution of the Schrödinger equation $e^{-i\lambda_\pm t}$:

$$|B_L\rangle(t) = |B_L\rangle(0)e^{-i\lambda_- t}; \qquad |B_H\rangle(t) = |B_H\rangle(0)e^{-i\lambda_+ t}. \tag{1.12}$$

A state which is initially pure $|B^0\rangle$ or $|\bar{B}^0\rangle$ determines the values for the amplitudes at $t = 0$. The time evolution of initially pure $|B^0\rangle$ or $|\bar{B}^0\rangle$ states is then given by:

$$|B^0(t)\rangle = I_+(t)\,|B^0\rangle + \frac{q}{p} I_-(t)\,|\bar{B}^0\rangle\,, \tag{1.13}$$

$$|\bar{B}^0(t)\rangle = \frac{p}{q} I_-(t)\,|B^0\rangle + I_+(t)\,|\bar{B}^0\rangle\,, \tag{1.14}$$

with $I_\pm(t)$ time-dependent functions:

$$I_\pm(t) = \frac{1}{2}\left(e^{-it\lambda_+} \pm e^{-it\lambda_-}\right). \tag{1.15}$$

The functions $I_\pm(t)$ show that an initially pure $|B^0\rangle$ will obtain a $|\bar{B}^0\rangle$ component as time evolves. This is an oscillatory behavior, as becomes clear if we write $I_+(t)$ and $I_-((t)$ as follows:

$$I_+(t) = e^{-iMt}e^{-\Gamma t/2}\left[\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right], \tag{1.16}$$

$$I_-(t) = e^{-iMt}e^{-\Gamma t/2}\left[\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right]. \tag{1.17}$$

Δm and $\Delta\Gamma$ have been defined as the mass and decay width difference between the states $|B_L\rangle$ and $|B_H\rangle$. The mass difference Δm is the oscillation frequency.

1.3 CP-violation

It has long been thought that parity is a symmetry of nature. The parity operation is equivalent to mirror reflection, reversing the sign of space coordinates: $\vec{x} \to -\vec{x}$. Parity symmetry means that all physical processes are invariant under a change of parity. Indeed, parity is conserved in all electromagnetic and strong interactions, but in 1956 C.S. Wu found parity to be violated in weak interactions [12]. Neutrinos provide a clear example of parity violation. In the limit of vanishing neutrino masses, all neutrinos are left-handed and all anti-neutrino's are right-handed[1]. Parity reverses the spin of a particle, turning a left-handed neutrino (ν_L) into a right-handed neutrino (ν_R), which doesn't exist in this limit. This fact is exhibited in the charged pion decay, when we compare it to the decay with opposite parity:

$$BR(\pi^+ \to \mu^+ \nu_L) = 99.98\% \neq BR(\pi^+ \to \mu^+ \nu_R) = 0\%. \tag{1.18}$$

This constitutes parity violation.

The operator C is called "charge conjugation", and changes a particle into its anti-particle. It changes the sign of all internal quantum numbers, such as charge, baryon number and strangeness or bottomness, while leaving mass, momentum, and the direction of spin unchanged. For example, the charge conjugated decay of $\pi^+ \to \mu^+ \nu_L$ is $\pi^- \to \mu^- \bar{\nu}_L$. The branching ratios for these processes are vastly different, 99.98% vs. 0%, constituting C violation in weak decays.

If we apply *both* C and P to the decay $\pi^+ \to \mu^+ \nu_L$, we find the CP conjugated decay $\pi^- \to \mu^- \bar{\nu}_R$. The decay rate and branching ratio of these decays are identical:

$$BR(\pi^+ \to \mu^+ \nu_L) = BR(\pi^- \to \mu^- \bar{\nu}_R), \tag{1.19}$$

meaning that CP is conserved, at least in this process.

In 1964 however, a small CP-violating effect was found in neutral kaon decays. Kaons are produced by the strong interaction as flavor eigenstates, i.e. as states with a well-defined flavor: $d\bar{s}$ or $\bar{d}s$. However, in the same fashion as neutral B-mesons, they can mix to their own antiparticle, so over time they will become a superposition of the two flavor states $d\bar{s}$ and $\bar{d}s$. Neutral kaons typically decay to two or three pions, through a weak interaction. By coincidence, the two-pion final state has CP eigenvalue +1 and the three-pion final state has CP eigenvalue -1. If we assume that CP is conserved in the weak interaction (which it is to good approximation), it is interesting to view the superposition of neutral kaon flavor states as a combination of CP eigenstates. Because the energy release in the decay to two pions is larger, this decay is faster than the decay to three pions. This means that the $CP = +1$ component will quickly fade away, and a relatively pure $CP = -1$ state remains, which can only decay to three pions. Therefore, if we observe a neutral kaon decay to two pions at large distance from the point of production, we shall know that CP has been violated. Christenson *et al.* [13] found such decays, establishing the existence of CP-violation in the neutral kaon system.

After this discovery, it was realized that the complex phase in the CKM matrix, related to the angle β of Fig. 1.1, allows CP-violation in the Standard Model. The next section describes how

[1] Right-handed (left-handed) means the spin-vector is parallel (anti-parallel) to the direction of motion.

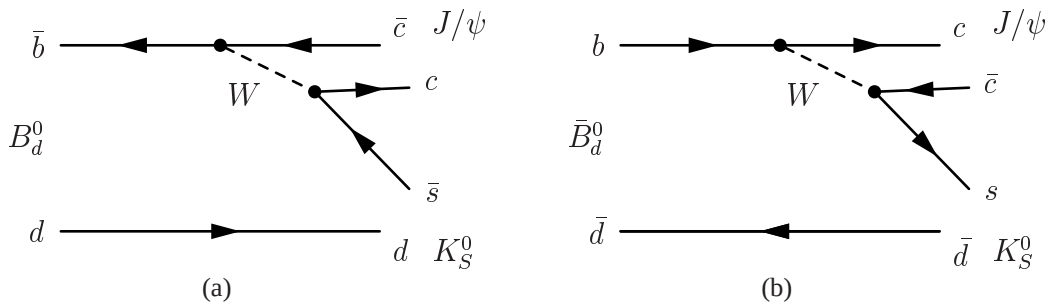


Figure 1.5: Dominant process contributing to (a) the decay $B_d^0 \to J/\psi K_S^0$ and (b) the decay $\bar{B}_d^0 \to J/\psi K_S^0$.

a non-zero angle β leads to CP-violation in $B_d^0 \to J/\psi K_S^0$ decays, and how these decays can be used to measure β.

1.3.1 CP-violation in $B_d^0 \to J/\psi K_S^0$

In section 1.2.2 we introduced some of the formalism to describe the effects of mixing of neutral B-mesons. We will now apply this to the case of the decay mode $B_d^0 \to J/\psi K_S^0$ and explain how this decay allows the extraction of the CP-violation parameter β.

We define the amplitudes for decay to a final state f as:

$$A_f = \langle f \mid B^0 \rangle , \tag{1.20}$$
$$\bar{A}_f = \langle f \mid \bar{B}^0 \rangle . \tag{1.21}$$

Note that both A_f and $\bar{A}_f$ will be non-zero only if f is a final state accessible by both B^0 and $\bar{B}^0$. This is the case for B_d^0's and $\bar{B}_d^0$'s decaying to $(J/\psi K_S^0)$. The dominant (tree-) diagrams contributing to these decays are shown in Fig. 1.5.

The time-dependent decay rate of a $|B^0\rangle$ to a final state f is:

$$\Gamma(B^0 \to f) = \left| \langle f \left| H \right| B^0 \rangle \right|^2 . \tag{1.22}$$

We can now use equations 1.13 and 1.14, in combination with equations 1.16 and 1.17, to find the time-dependent decay rates:

$$\Gamma\left(B^0(t) \to f\right) = \left|A_f\right|^2 e^{-\Gamma t} \cdot \left[\frac{1+|\zeta|^2}{2} + \frac{1-|\zeta|^2}{2} \cos\left(\Delta m t\right) - \mathrm{Im}\zeta \sin\left(\Delta m t\right) \right] , \tag{1.23}$$

$$\Gamma\left(\bar{B}^0(t) \to f\right) = \left|A_f\right|^2 \left|\frac{p}{q}\right|^2 e^{-\Gamma t} \cdot \left[\frac{1+|\zeta|^2}{2} - \frac{1-|\zeta|^2}{2} \cos\left(\Delta m t\right) + \mathrm{Im}\zeta \sin\left(\Delta m t\right) \right] . \tag{1.24}$$

For convenience we have defined:

$$\zeta = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \tag{1.25}$$

and we have also used the fact that in the B_d^0 system, $\Delta\Gamma$ is close to 0 [11].

The Standard Model predicts the mixing in the B_d^0 system to proceed predominantly through the exchange of virtual t-quarks, and the CKM matrix elements involved are V_{tb} and V_{td}. To order $\mathcal{O}(10^{-3})$ we can write p/q as:

$$\frac{q}{p} = \frac{V_{tb}V_{td}^*}{V_{tb}^*V_{td}} = e^{2i\beta}. \tag{1.26}$$

Here β is one of the angles in the unitarity triangle (see Fig. 1.1). The ratio of amplitudes $\bar{A}_f/A_f$ can be written as:

$$\frac{\bar{A}_f}{A_f} = \pm\left|\frac{\bar{A}_f}{A_f}\right| \cdot e^{2i\phi}. \tag{1.27}$$

The ratio $|\bar{A}_f/A_f|$ is 1 to good accuracy if only one diagram contributes to the decay, or if the phase ϕ of each contributing diagram is the same. The higher-order diagrams contributing to $B_d^0 \to J/\psi K_S^0$ are called penguin diagrams. The penguin diagrams are strongly suppressed in this decay, which implies that to good approximation only the tree diagram shown in Fig. 1.5 (a) contributes to the decay. Moreover, the dominant penguin diagram has the same phase as the tree diagram [14]. Therefore, for the decay $B_d^0 \to J/\psi K_S^0$, we can write:

$$\zeta = \pm e^{2i(\beta+\phi)}. \tag{1.28}$$

The sign in equation 1.28 is positive if f is a CP eigenstate with CP eigenvalue $+1$, or negative for a CP eigenvalue of -1, but in any case equation 1.28 implies $|\zeta| = 1$. This simplifies equations 1.23 and 1.24 to:

$$\Gamma\left(B_d^0(t) \to J/\psi K_S^0\right) = |A_f|^2 e^{-\Gamma t} \cdot \left[1 - \mathrm{Im}\zeta \sin\left(\Delta M t\right)\right], \tag{1.29}$$

$$\Gamma\left(\bar{B}_d^0(t) \to J/\psi K_S^0\right) = |A_f|^2 e^{-\Gamma t} \cdot \left[1 + \mathrm{Im}\zeta \sin\left(\Delta M t\right)\right]. \tag{1.30}$$

We define the time-dependent CP-asymmetry as:

$$a_{CP}(t) \equiv \frac{\Gamma\left(B^0(t) \to f\right) - \Gamma\left(\bar{B}^0(t) \to \bar{f}\right)}{\Gamma\left(B^0(t) \to f\right) + \Gamma\left(\bar{B}^0(t) \to f\right)}. \tag{1.31}$$

To describe the CP-asymmetry in the decay $B_d^0 \to J/\psi K_S^0$, we can use equations 1.29 and 1.30. Equation 1.31 then becomes

$$a_{CP}(t) = -\mathrm{Im}(\zeta) \sin\left(\Delta M t\right), \tag{1.32}$$

or, using equation 1.28,

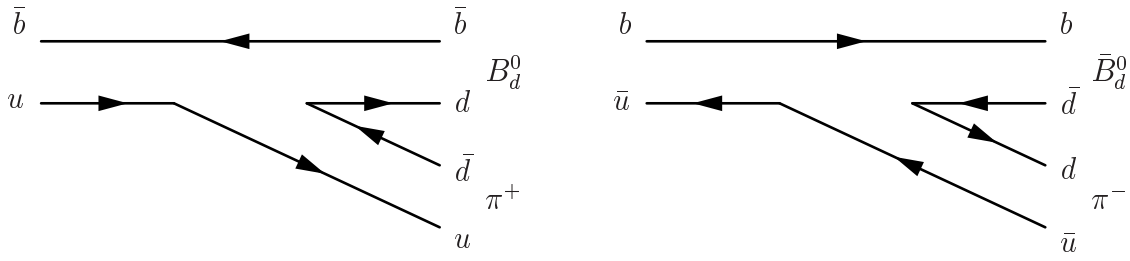


Figure 1.6: The principle of same-side tagging. The charge of the pion created in the fragmentation of the B-meson is correlated with the B-meson flavor.

$$a_{CP}(t) = \sin\left(2(\beta+\phi)\right)\sin\left(\Delta M t\right). \tag{1.33}$$

The sum $\beta+\phi$ is independent of the chosen parametrization of the CKM matrix. In the Wolfenstein parametrization β is the phase of the product $V_{tb}V_{td}^*$, as defined by equation 1.26. The angle ϕ is associated with the $b \to c\bar{c}s$ decay amplitude, which is proportional to the matrix elements $V_{cb}V_{cs}^*$. In the Wolfenstein parametrization $\phi = 0$, as is clear from equation 1.3. Therefore, a measurement of the CP-asymmetry $a_{CP}(t)$ using $B_d^0 \to J/\psi K_S^0$ decays provides a clean measurement of $\sin(2\beta)$.

1.3.2 Measurement of the angle β at DØ

This thesis describes the selection of B_d^0's and $\bar{B}_d^0$'s decaying to the $(J/\psi K_S^0)$ final state and a detailed study of the relevant backgrounds. Since we do not discriminate between a B_d^0 and a $\bar{B}_d^0$ in the initial state, we determine the sum of the time evolutions $\Gamma(B_d^0 \to J/\psi K_S^0)(t)$ and $\Gamma(\bar{B}_d^0 \to J/\psi K_S^0)(t)$. For a $\sin(2\beta)$ measurement, it is necessary to separately determine the two contributions to this sum. That is, for each event, one needs to know if the initial state is B_d^0 or $\bar{B}_d^0$. The determination of the initial state is called flavor tagging.

In general, the approach to flavor tagging can be to either determine the B-meson flavor (whether it contains a b- or $\bar{b}$-quark) on the side of the fully reconstructed decay, or one can attempt to determine the flavor of the opposite B-hadron in the event. These techniques are referred to as same-side tagging or opposite-side tagging, respectively. The same-side tagging technique recently developed at DØ [15] was first proposed by Gronau, Nippe and Rosner [16]. The fragmentation of the B-meson, where the b-quark binds with another quark from the event, generally has an associated pion. The charge of this pion is correlated with the B-meson flavor, as indicated in Fig. 1.6. If the B-meson decay produces a single lepton, then the charge of this lepton can also be used as a tag of the B-meson flavor, but this tag is not usable in the decay $B_d^0 \to J/\psi K_S^0$. The opposite-side tagging technique that has recently become available at DØ requires the other-side B-meson to decay to hadrons and a muon. As in the same-side tagging procedure, the charge of this muon is correlated with the B-meson flavor.

1.3.3 Recent results on the angle β

The first measurement of the CP-violating parameter $\sin(2\beta)$, using $B^0_d \to J/\psi K^0_S$ decays, has been published in 1998 by the CDF experiment [17]. This result was later updated to $\sin(2\beta) = 0.79^{+0.41}_{-0.44}$, using approximately 400 reconstructed $B^0_d \to J/\psi K^0_S$ decays [18]. Around the same time, the OPAL experiment at LEP published a measurement [19], but the statistics available (24 $B^0_d \to J/\psi K^0_S$ candidates, with a purity of 60%) is a severe challenge to the analysis. The OPAL result is $3.2^{+1.8}_{-2.0}(\text{stat.}) \pm 0.5(\text{syst.})$. Obviously, if a value of $\sin(2\beta)$ larger than 1 is reported, a special interpretation of the result is required. The OPAL result is interpreted as an indication of the consistency of the data with a certain value of $\sin(2\beta)$. The consistency of the result with a positive value of $\sin(2\beta)$ is found to be 68.5%.

Also in 1998, the PEP-II [20] asymmetric B-factory came into operation at the Stanford Linear Accelerator Center. This e^+e^- accelerator and its accompanying detector BABAR have specifically been designed to study B^0_d and B^+ decays. BABAR and PEP-II are still operating today. In 1999 a second asymmetric B-factory called KEKB was commissioned in Japan at the KEK laboratory. The experiment at this accelerator, called Belle, is currently still running as well. These two experiments, BABAR and Belle, have produced the most accurate measurements of $\sin(2\beta)$ to date. Using the decay modes: $B^0_d \to J/\psi K^0_S, \psi(2S)K^0_S, \chi_{c1}K^0_S, \eta_c K^0_S, J/\psi K^0_L$, BABAR reports a value of $\sin(2\beta)$ of $0.741 \pm 0.067(\text{stat.}) \pm 0.034(\text{syst.})$ [21]. The most recent result reported by Belle is $0.733 \pm 0.057(\text{stat.}) \pm 0.028(\text{syst.})$, using the same decay modes [22]. The result from Belle is shown in Fig. 1.7, showing a_{CP} as a function of proper time of the decaying B-meson. The asymmetry in the $J/\psi K^0_L$ channel is separated from the other channels, because it has the opposite CP eigenvalue. The Heavy Flavor Averaging Group has averaged these and other results on the angle β, resulting in the world average of [23]:

$$\sin(2\beta)_{WA} = 0.692 \pm 0.047. \tag{1.34}$$

A measurement of $\sin(2\beta)$ provides a four-fold ambiguity of the angle β. In Fig. 1.8 this is visualized as four cones protruding from the bottom-right corner of the unitarity triangle, in which the apex of the unitarity triangle is allowed. Other measurements can also be translated into constraints on the position of the apex of the triangle. The most stringent results are shown in Fig. 1.8. CP-violation in kaon decays can give a measure of the parameter ϵ_K, the analog in the neutral kaon system of ϵ_B, defined in equation 1.10. Measurements of ϵ_K constrain the unitarity triangle, as indicated in Fig. 1.8. The mass difference between the B^0_d and $\bar{B}^0_d$ states, Δm_d, and the mass difference between the B^0_s and $\bar{B}^0_s$ states, which is labeled Δm_s, pose a constraint on the length of one of the sides of the triangle. Information on these mass differences is obtained from B^0_d–$\bar{B}^0_d$ and B^0_s–$\bar{B}^0_s$ mixing studies. A final constraint is imposed by knowledge of the ratio $|V_{ub}/V_{cb}|$. This ratio can be measured by studying the end-point spectrum in semileptonic B-decays. The two coordinates of the apex are overconstrained by these results, constituting a test of the Standard Model. A fit of the triangle is shown in Fig. 1.8. All measurements show excellent consistency, thereby confirming the Standard Model.

Interestingly, both BABAR and Belle have measured $\sin(2\beta)$ in B^0_d decays to ϕK^0_S, $K^+K^-K^0_S$, and $\eta' K^0_S$ as well [24, 25]. All measurements are consistent with the Standard Model predictions, with one exception. Belle finds a 3.5σ deviation in the $B^0_d \to \phi K^0_S$ channel: $-0.96 \pm 0.5^{+0.09}_{-0.11}$.

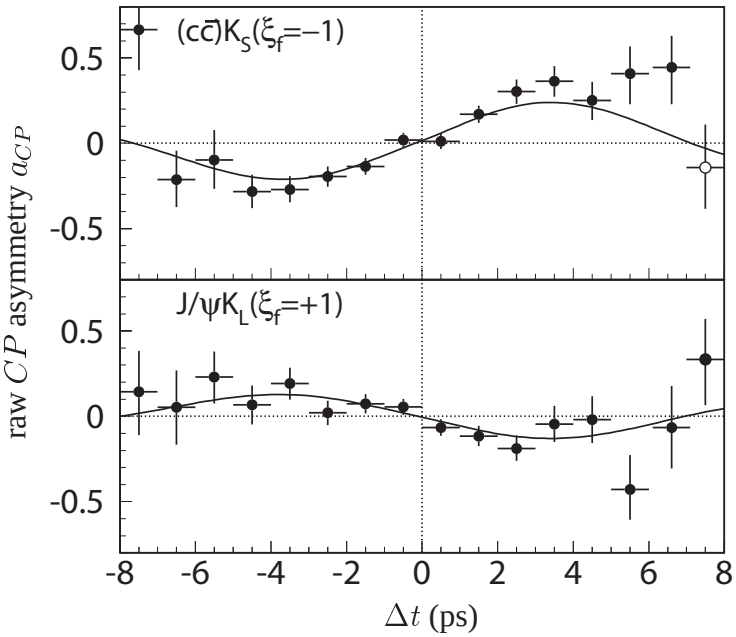


Figure 1.7: Raw, time-dependent CP asymmetry in the B^0_d decay modes to $c\bar{c}K^0_S$ (top) and $J/\psi K^0_L$ (bottom) [22]. In the formalism used at the B-factories, the proper time difference of the decaying B-mesons is labeled Δt. The amplitude of the fit would be $\sin(2\beta)$, if the mistagging rate were zero.

The BABAR result for $\sin(2\beta)$ in this channel is $0.45 \pm 0.43 \pm 0.07$, which is consistent with the current world average. In the Standard Model interpretation, the CP asymmetry in all of the modes discussed above yields $\sin(2\beta)$ to good accuracy. Already in 1996 Grossman and Worah pointed out that regardless of what the Standard Model prediction is, if the result in one or more of the decay modes listed above deviates from the other results, this is a sensitive test of physics beyond the Standard Model [26]. This is an interesting interpretation of the Belle result in the $B^0_d \to \phi K^0_S$ mode, though it should be noted that the BABAR measurement in this mode is consistent with the Standard Model.

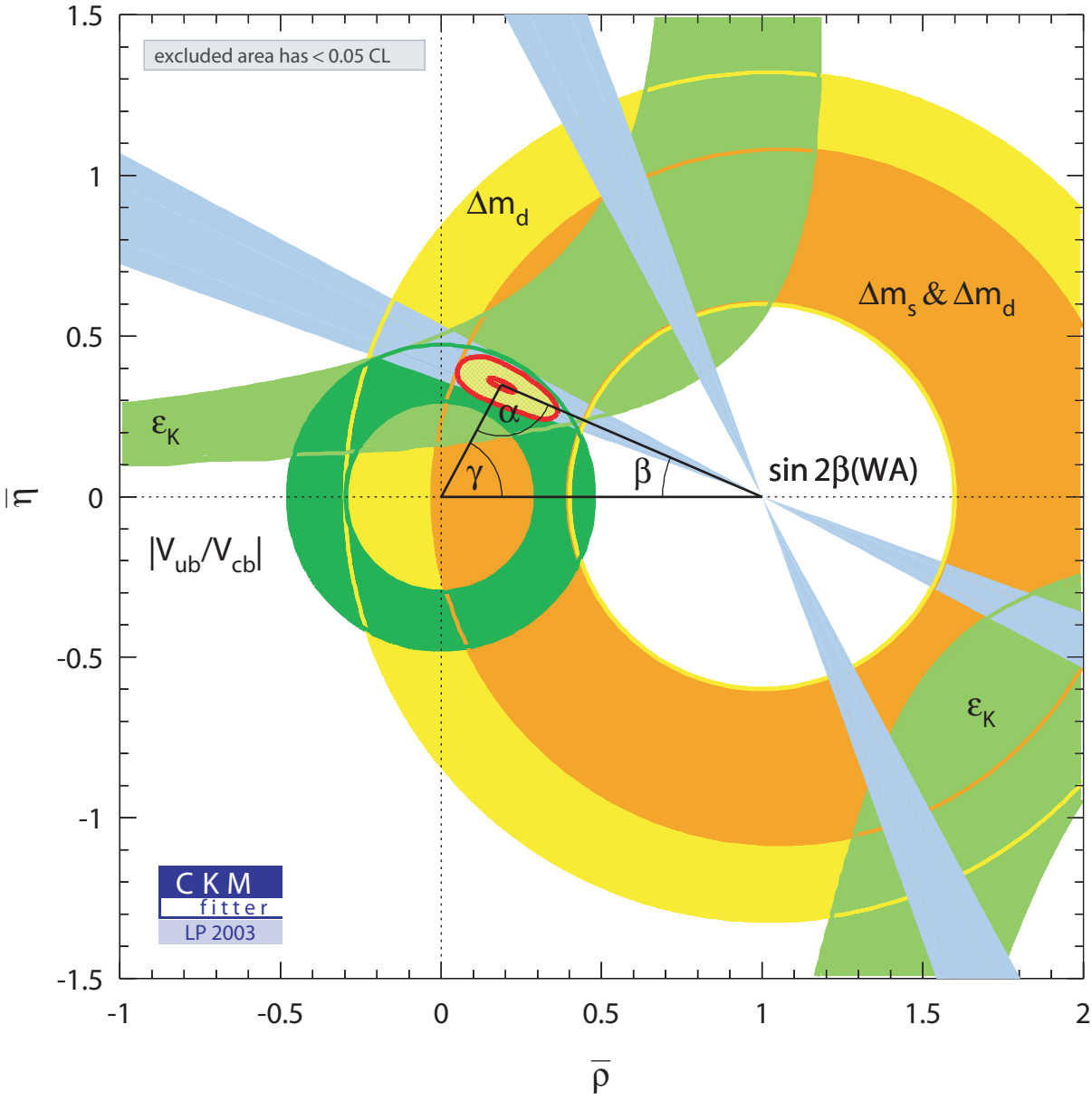


Figure 1.8: The unitarity triangle in the $(\bar{\rho}, \bar{\eta})$ plane, which is equivalent to the plane in which the unitarity triangle of Fig. 1.1 is drawn.

Chapter 2

The DØ experiment

The DØ experiment has been constructed to study proton-antiproton collisions at the Tevatron collider at Fermilab near Chicago, Illinois, U.S.A. It was originally designed to study high-mass states and high-p_T phenomena. It has already taken data during Run I of the Tevatron from 1992 to 1996. Using this data, the DØ experiment has discovered the top quark [27], together with the CDF experiment [28, 29, 30], which is also located at the Tevatron. The DØ experiment was named after the interaction point in the Tevatron ring where it is located.

After 1996, both the DØ and the CDF detectors, as well as the Tevatron, have been upgraded. The upgrade of the DØ detector has extended in particular the capabilities to study low-p_T phenomena, reflecting increased attention for B-physics and soft QCD. The upgrades were completed in March 2001, which marked the starting point of Tevatron Run II. In this chapter a description will be given of the Tevatron and its pre-accelerators, as well as the DØ detector.

2.1 The Tevatron

The Tevatron is a circular accelerator with a circumference of about 4 miles. It accelerates beams of protons and antiprotons to an energy of 980 GeV, and collides them at two interaction points at a center of mass energy of 1.96 TeV. It is the first accelerator to provide collisions with a center of mass energy greater than 1 TeV. Until the start-up of the Large Hadron Collider (LHC), which is planned for 2007, the Tevatron will be the highest-energy collider in the world. The Tevatron was originally designed to provide a luminosity of $1 \times 10^{31}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$. After the upgrades of the past years, a luminosity of $29.4 \times 10^{31}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ is now possible, and an integrated luminosity exceeding 2 fb^{-1} is foreseen before the start-up of the LHC in 2007. Some parameters outlining the performance of the accelerator are given in Table 2.1. We will describe the Tevatron operation below.

The proton beam for the Tevatron is generated in the pre-accelerator. Here negatively charged hydrogen ions are accelerated to 750 keV in a Cockroft-Walton accelerator. From there, the hydrogen atoms are bunched and fed into a 150 meter long linear accelerator, called the Linac. The Linac accelerates the ions to an energy of 400 MeV, after which they are led through a carbon foil. This foil strips both electrons off the hydrogen ion, leaving only the proton. The protons are

	Run IB (1993 – 1995)	Typical for April 2003	Typical for May 2004	Run II (projected)
Protons per bunch	2.32×10^{11}	2.0×10^{11}	2.4×10^{11}	2.7×10^{11}
Anti-protons per bunch	0.55×10^{11}	0.22×10^{11}	0.27×10^{11}	1.35×10^{11}
Number of bunches	6	36	36	36
Energy [GeV]	900	980	980	980
Bunch length [cm]	60	60	50	50
Initial luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	1.6×10^{31}	3.5×10^{31}	6.0×10^{31}	29.4×10^{31}
Integrated lumin. [pb^{-1}/week]	3.2	5.2	9.0	53.4
Bunch spacing [ns]	3500	396	396	396
Interactions per crossing	2.5	0.9	1.6	7.8

Table 2.1: Tevatron parameters.

then led into the Booster, a circular accelerator, which accelerates the protons to an energy of 8 GeV. At this point, the protons are ready to be transferred to the Main Injector.

The Main Injector has the purpose of accelerating the protons to the right energy for injection into either the Tevatron, or the antiproton source to produce antiprotons. To produce the antiprotons, a beam of protons is generated in the Main Injector. This beam of protons is accelerated to 120 GeV and led to the antiproton source, where it is collided with a nickel target. The antiprotons produced at the target are temporarily stored in the Accumulator. When enough antiprotons have been produced to fill the Tevatron (about 10^{12}), they are transferred to the Main Injector, which accelerates them to 150 GeV for injection into the Tevatron.

When both the proton and the antiproton beams are in the Tevatron, they will be accelerated from 150 to 980 GeV. In order to keep the particle beams in their orbit, 4.2 Tesla superconducting dipole magnets are employed along the 4 mile long circumference of the Tevatron. The state of the accelerator in which both beams are in the machine, is called a store.

Finally, the beams are brought into collision in the Tevatron. The instantaneous luminosity will slowly decrease from its initial peak value, due to the interaction of the beams with each other, bunch emittance growth and other effects [31]. After a number of hours, the point is reached where it is more beneficial to dump the beams and start with new beams, than to continue using the existing ones. At this point new beams will be injected into the Tevatron. The lay-out of the accelerator complex is depicted in Fig. 2.1.

2.2 The DØ detector

The DØ detector has a layered structure, typical for general purpose collider detectors. From the interaction point in the center of the detector a particle encounters the central tracking detectors first. These trackers are focused on accurate and efficient measurement of tracks, as well as good separation of tracks when many tracks appear close to each other. Moreover, the design of trackers is such, that the disturbance of paths of passing particles, as well as the energy-loss,

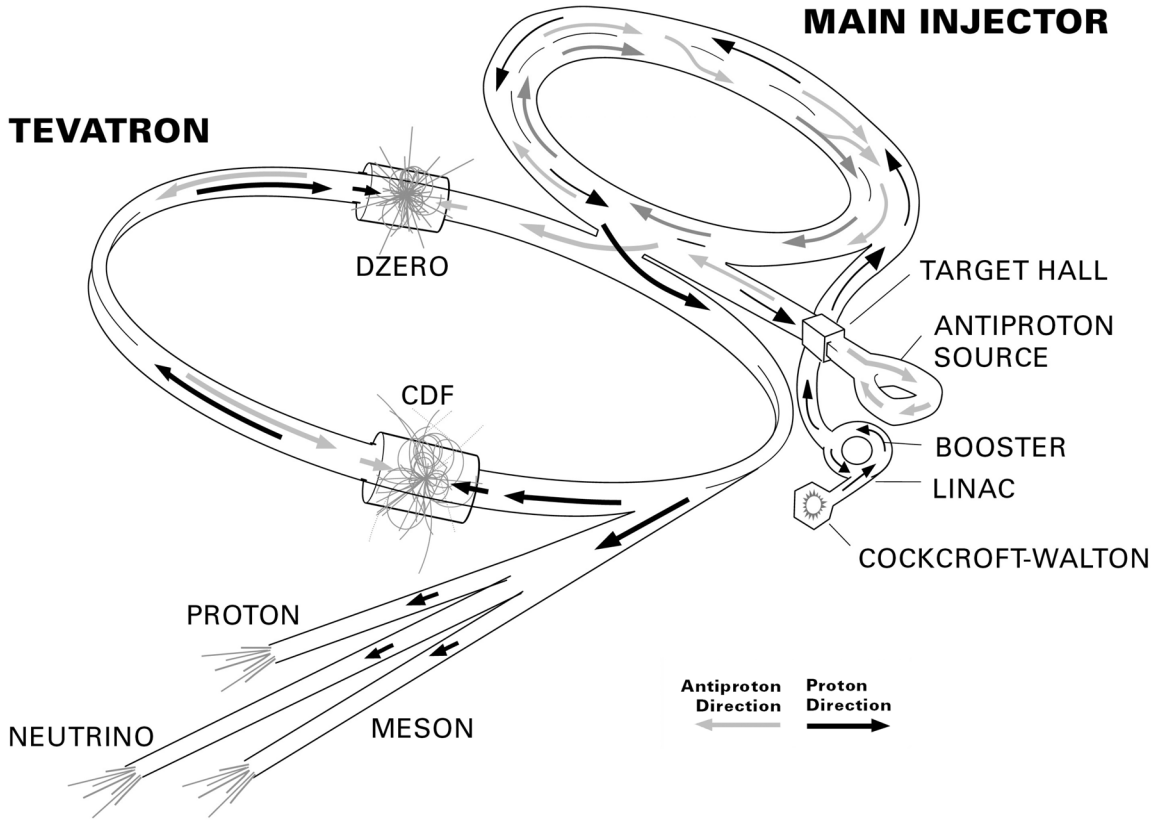


Figure 2.1: The Fermilab accelerator complex, showing the DØ and CDF experimental areas as well as the fixed target lines.

is minimized. To enable the tracker to measure the momenta of tracks, a 2 Tesla magnetic field is generated in the tracking volume by a superconducting solenoid, which surrounds the tracker [32]. The calorimeter, located outside the solenoid, absorbs hadrons, electrons and photons and accurately measures their energy and direction. The outer layer of the DØ detector is the muon system, measuring the tracks and time of flight of passing charged particles. Generally, the only particles that pass through the calorimeters and leave traces in this outer layer are muons. The lay-out of different subdetectors is shown in Fig. 2.2. The sequence of subdetectors as described, makes it possible to identify many of the particles in the final state, because different particles leave different signatures, as is shown in Fig. 2.3.

The next sections will describe the subdetectors in more detail, starting with central tracking detectors. This is followed by a section on the calorimeter, and finally the muon system will be discussed. For convenience, we will begin with the definition of coordinate systems used in DØ.

2.2.1 Coordinate system

We use a right-handed coordinate system, in which the positive z-axis is aligned with the proton beam direction. The y-axis points vertically upward, and the x-axis is horizontal and points away from the center of the ring. The *transverse plane* is defined as the (x, y) plane. A more useful set of coordinates are the polar coordinates (r, ϕ, θ), in which r and ϕ are defined as:

$$r = \sqrt{x^2 + y^2} \quad ; \quad \phi = \arctan \frac{y}{x}, \tag{2.1}$$

and θ is the angle with the beam axis:

$$\theta = \arctan \frac{r}{z}. \tag{2.2}$$

The angle θ is often replaced by the *pseudo-rapidity* η:

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right). \tag{2.3}$$

The pseudo-rapidity is a convenient choice at a hadron collider as the multiplicity of high-energy particles is roughly constant as a function of η.

2.2.2 Central tracking detectors

The most important goal of the central tracking system is to efficiently measure tracks and vertices over the large η coverage of the DØ detector (see Fig. 2.4). Efficient track reconstruction in the Tevatron's high-particle density environment puts stringent requirements on the tracker's pattern recognition capabilities. Accurate vertex reconstruction requires the detector to be located sufficiently close to the interaction region, which is complicated by the high radiation levels in the environment at the Tevatron. Finally, it's necessary to be able to read out the tracking system at the 2.5 MHz bunch crossing rate of the Tevatron. These considerations have led to a tracker design consisting of two different technologies.

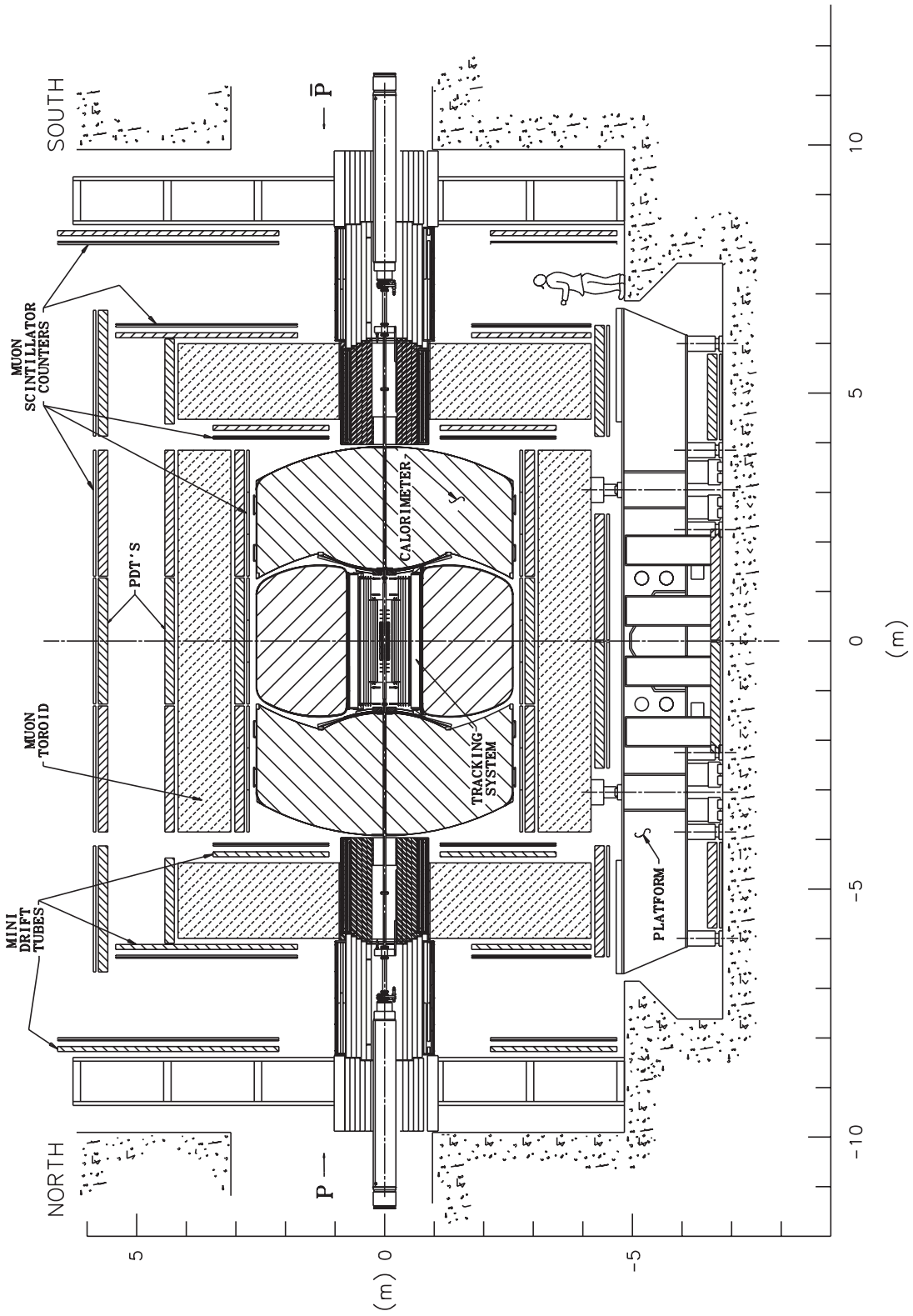


Figure 2.2: Cross-sectional view of the DØ detector.

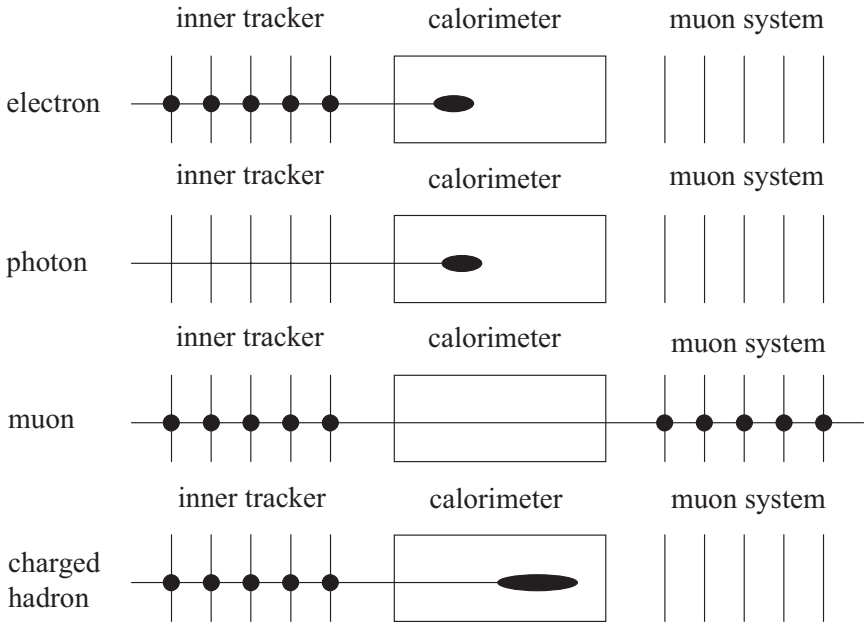


Figure 2.3: Signatures of different particles in the subdetectors of the DØ detector.

For the region close to the beam pipe a silicon strip detector was chosen for its radiation hardness and good granularity, in order to achieve accurate vertex reconstruction. For the region further away from the beam pipe, a scintillating fiber tracker was chosen. A fiber tracker allows the measurement of many points on a track, without disturbing the track too much. This is necessary for proper pattern recognition. A fiber tracker can also satisfy the less stringent, but still demanding requirements on radiation hardness at this somewhat larger radius.

Silicon microstrip tracker

The overall design of the silicon microstrip tracker (SMT) is motivated by the extended interaction regions of the Tevatron, which are about 60 cm long. A silicon strip detector provides the most accurate measurements when the tracks intersect the detector planes at normal incidence. This leads to a hybrid system with barrel detectors, sandwiched between disk detectors. The detector lay-out is depicted in Fig. 2.4. The SMT consists of six 12.4 cm long barrel detectors, containing eight layers of rectangular silicon microstrip detectors, referred to as ladders. Almost all ladders consist of double-sided silicon with axial strips on one side of the sensor and strips with a small (stereo-) angle on the other side of the sensor. This allows the determination of the third coordinate, in the direction of the strip. In layers 1, 2, 5 and 6 of the four innermost barrels the stereo-angle is 90˚, in layer 3, 4, 7 and 8 the stereo-angle is 2˚. On the two outermost barrels, layers 1, 2, 5 and 6 are single-sided. These ladders do not provide stereo information. A view of the ladder arrangement is shown in Fig. 2.5. The innermost layer is at a radius of 2.7 cm, the outermost layer at 9.7 cm. The disks between the barrel sections of the detector are referred

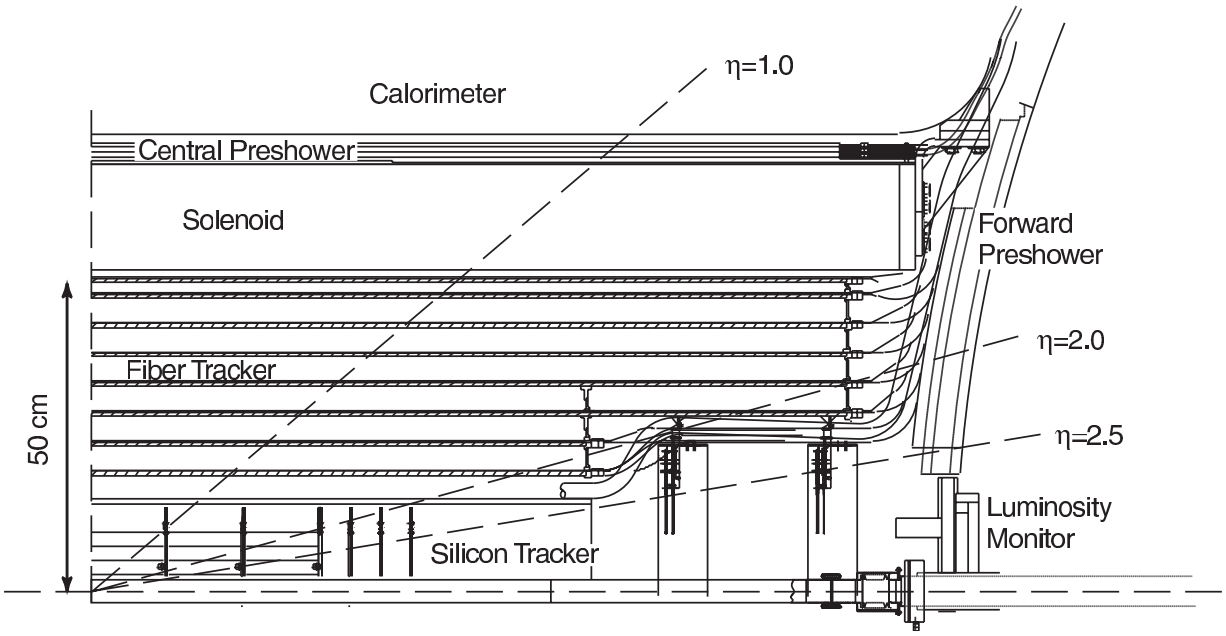


Figure 2.4: Configuration of one quadrant of the tracking system. The bottom of this graph is formed by the beam pipe, with the center of the detector at the bottom left.

to as F-disks, which hold detectors with a stereo-angle of 30˚. To improve the coverage of the SMT to $|\eta| < 3$, four F-disks are placed outside the barrel sections at larger $|z|$ and two disks of a different type, called H-disks, are positioned even further away on each side of the interaction point. The H-disk detectors have a stereo angle of 15˚.

The pitch of the strips on the different sensors varies across the different types (see Table 2.2). All silicon sensors are read out with 128 channel SVX-IIe chips [33]. The whole SMT comprises about $793,000$ read-out channels.

The F-disks and the H-disks are equipped with silicon diodes that measure the instantaneous and integrated radiation dose to which the detector is exposed [34]. These radiation monitors were contributed to the experiment by NIKHEF.

Central fiber tracker

The central fiber tracker (CFT) [35] consists of 32 concentric barrel-shaped layers of scintillating fibers, surrounding the SMT. These 32 layers are arranged in 16 doublet layers, which are then grouped together in eight "superlayers" at radii from 19.5 cm to 51.4 cm as shown in Fig. 2.4. The CFT covers the central region ($|\eta| < 2$). In addition to the measurement of tracks of charged particles, the purpose of the CFT is to provide a level 1 trigger.

Each superlayer consist of two doublet layers. A doublet layer is composed of two layers of scintillating fibers, which are half a fiber diameter offset with respect to each other, to improve coverage. The inner doublet layer in each superlayer is parallel to the beam axis, the outer doublet

	pitch [μm]
Barrel sensors	
Axial strips	50
2˚ stereo strips	62.5
90˚ stereo strips	156
Disk sensors	
F-disks, p-side	52
F-disks, n-side	65
H-disks	81

Table 2.2: Strip pitch for the different silicon sensors.

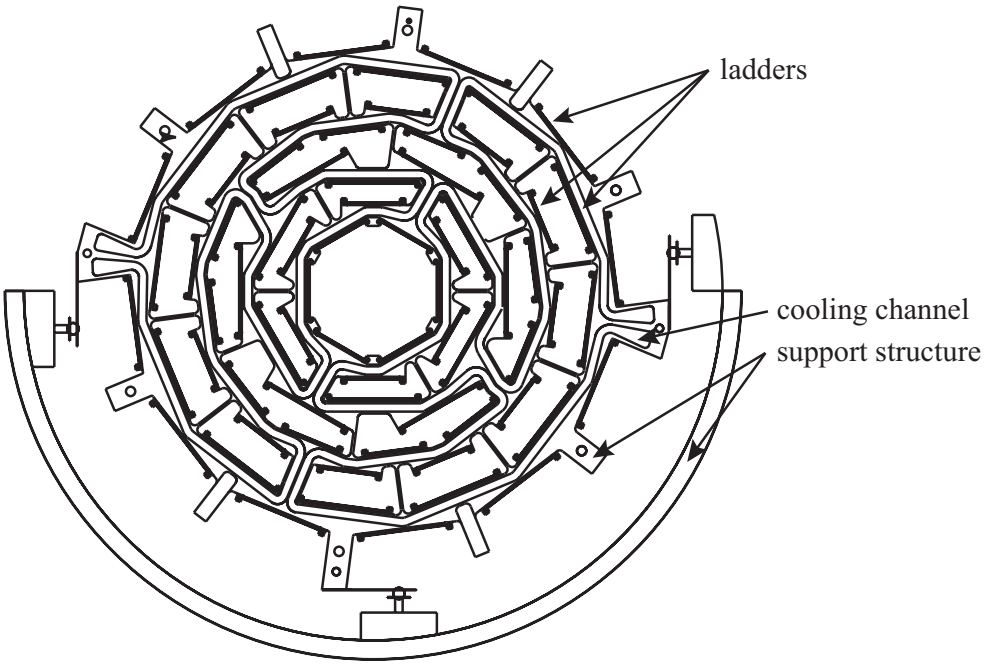


Figure 2.5: Arrangement of the ladders in the DØ SMT barrel detectors.

layer is at an alternating stereo angle of $\pm 3^{\circ}$, to obtain information about the z-coordinate.

The fibers have a diameter of 835 μm. The peak emission wavelength of these fibers is around 530 nm (green visible light). Every fiber is mated to a 7–11 m long waveguide which pipes the scintillation light to a visible light photon counter (VLPC). VLPC's are solid state photon detectors, based on silicon diodes with an operating temperature around 10 K. They have a high gain (50,000 electrons per converted photon) and their quantum efficiency for visible light is about 80%.

Solenoid magnet

The central tracking systems are surrounded by a superconducting solenoid, enabling momentum measurement of charged particles. The solenoid is a 2.8 m long two-layer coil, with a mean radius of 60 cm and a field strength of 2 Tesla. The stored energy is 5 MJ. There is no specific flux return: Most of the field returns in the space between the solenoid cryostat and the muon system, which is taken up by the calorimeters. A field uniformity of $\sin\theta \int B_z dl$ better than 0.5% is achieved by winding the coil such that an increased current density is obtained near the ends of the coil.

The field is monitored by 36 three-dimensional magnetic field sensors, that have been installed in the small gap between the inner wall of the solenoid cryostat and the outer radius of the CFT [36]. These sensors have been designed and manufactured at NIKHEF.

2.2.3 Calorimeter

The liquid-argon calorimeter [37] provides energy measurements for electrons, photons and jets. It remains unchanged from Run I, but the reduced bunch crossing interval of 396 ns in Run II (from 3.5 μs in Run I) has required an upgrade of its front-end electronics. The former peak sampling time of 2.2 μs was reduced to 400 ns, mainly by decreasing the effective integration time and reducing the intrinsic noise of the pre-amplifier [38].

The calorimeter, as shown in Fig. 2.6, is divided into a central calorimeter covering $|\eta| < 1$ and two end calorimeters to cover the region $1 < |\eta| < 4$. Both calorimeters contain three basic types of modules: An electromagnetic section with thin nearly pure depleted uranium plates (thickness 3 or 4 mm), fine hadronic modules with 6 mm thick uranium-niobium(2%) alloy plates and coarse hadronic sections. These latter sections consist of 46.5 mm thick plates, made of copper for the central calorimeter and stainless steel in the end calorimeters. There are four separate layers for the electromagnetic modules, three for the fine hadronic and one or three for the coarse hadronic modules. These layers are used to measure the longitudinal shower shape to distinguish between electrons and hadrons.

The smallest unit of the calorimeter is a read-out cell. The read-out cells are organized in semi-projective towers, pointing toward the interaction region. Semi-projective means that a straight line can be drawn through the interaction point and the cells of one tower, but that the sides of the cells are not aligned toward the interaction point. The segmentation of the read-out cells is (0.1×0.1) in (η, ϕ), except for the third layer in the electromagnetic modules, where cells cover (0.05×0.05) in (η, ϕ).

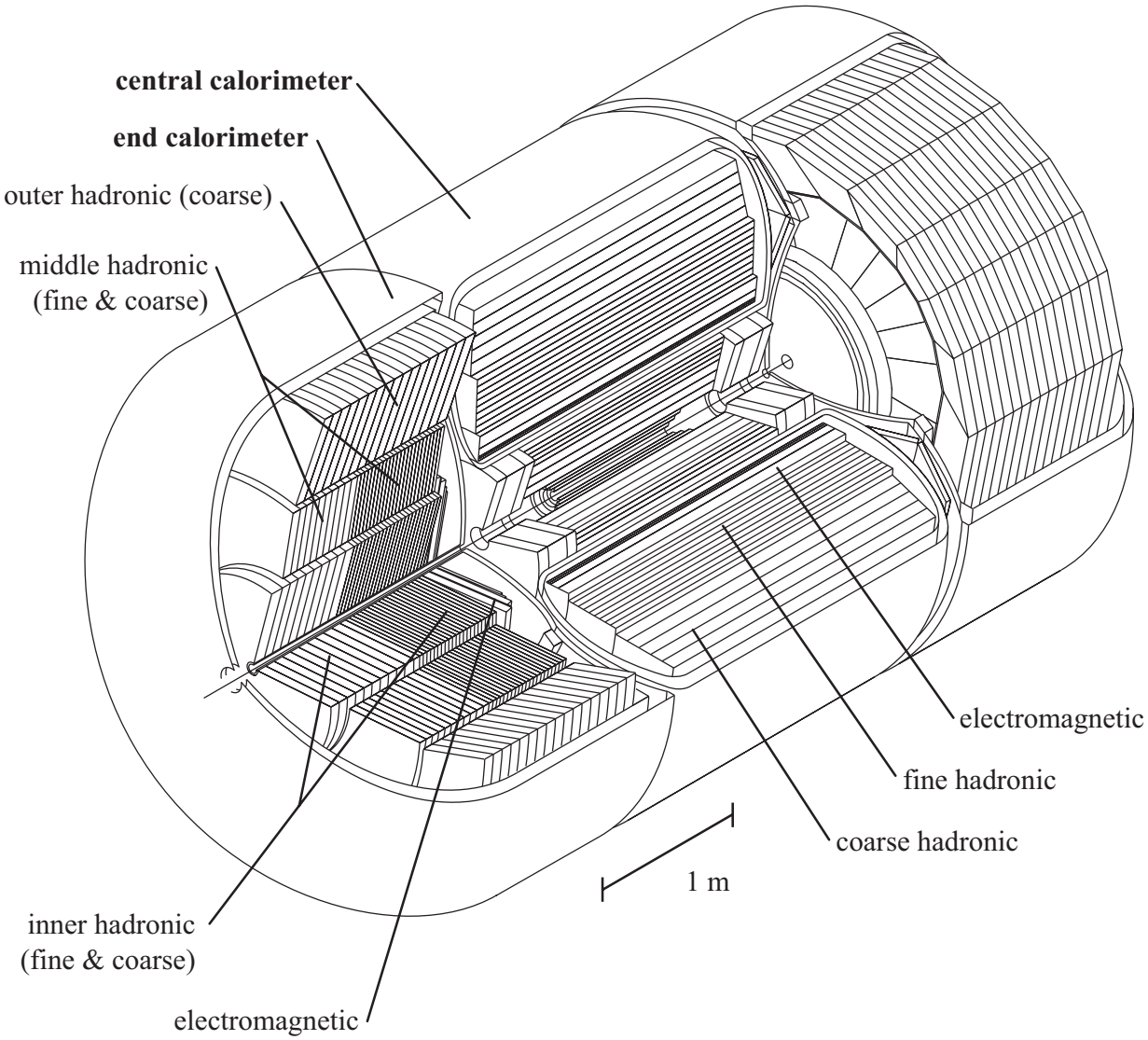


Figure 2.6: Three-dimensional view of the calorimeter.

Intercryostat detector

The region between the central and end calorimeters ($0.8 < \eta < 1.4$) contains large amounts of inactive material, mostly the walls of the calorimeter cryostats, light guides coming from the CFT and read-out electronics connected to the SMT and the magnetic field monitors for the solenoid. To supplement the energy measurement for particles in this region, an inter-cryostat detector has been installed. This detector consists of a single layer of 384 scintillating tiles, covering (0.1×0.1) in (η, ϕ), to match the calorimeter cells. The tiles are mounted on the front surface of the end calorimeters. The light signals are picked up by wavelength shifting fibers in the tiles and transported via clear fiber ribbon cables toward photo-multiplier tubes outside the magnetic field.

Preshower detectors

With the addition of the solenoid after Run I, the electromagnetic energy resolution has suffered. This is due to energy loss of particles in the solenoid and central tracker material. To remedy this loss, preshower detectors have been added in the central, as well as in the forward region.

The central preshower detector consists of triangular scintillating strips, arranged in one inner axial layer and two outer stereo layers at an angle of about $\pm 23^{\circ}$. Each strip has a base of 7.1 mm, with a 1 mm hole in the center. The read-out occurs via wavelength shifting fibers passing through this hole. Near the end of the strips, the wavelength shifting fibers are connected to clear waveguide fibers. These lead to VLPC's under the detector. A layer of lead absorber has been installed in front of the detector, so that the total material between the detector and the interaction region corresponds to approximately 2 radiation lengths of material.

Two forward preshower detectors are mounted on the front surface of each end calorimeter. They cover the $|\eta|$ range from 1.4 to 2.5. Their design is very similar to that of the central preshowers. A layer of lead absorber is sandwiched between two scintillator planes. The inner scintillators act as detectors for minimum ionizing particles, the outer scintillators detect the (possible) particle showers generated in the absorber.

2.2.4 Muon system

The muon detectors are based on two technologies: Drift chambers and scintillation counters. Both technologies are low-cost relative to the area that can be covered. This is important, because the muon system is located at the outer radius of the detector and therefore large area coverage is needed to obtain the necessary angular coverage. A drawback of drift chambers is their poor measurement of the time when a particle passes through. Therefore the drift chambers are supplemented with scintillation counters, that do not provide accurate tracking information, but have very good timing resolution. The timing resolution of the scintillator counters is of the order of 1 ns [39, 40]. Finally, a magnet system is provided, to enable momentum measurement in the muon system. The DØ muon system consists of three layers of drift-chambers and scintillation counters, called the A-, B- and C-layer, labeled starting at the inner layer, and a toroid magnet system. It is divided into central and forward systems, covering the range $|\eta| < 1$ and

the range $1 < |\eta| < 2$, respectively. The central muon system [41] consists of the WAMUS drift chambers, the Cosmic cap and bottom scintillation counters, and the A-phi scintillation counters [39]. The forward muon system [42] consists of the FAMUS drift chambers, as well as the pixel scintillation counters [40]. The design of these detectors will be described in the following paragraphs. We will start by describing the muon magnet system.

The muon magnet system

The magnet for the muon system is an iron-core toroid. The iron core helps reaching a high field value, at modest current (and operational cost). As for the muon system itself, it consists of a central and a forward section. The central toroid magnet, visible in Fig. 2.7, is a square annulus, with a thickness of 109 cm and a weight of 1973 metric tons. The forward toroid core is made of single large weldments of plate steel. The central and forward toroids are currently operated in series at 1250 Ampère, resulting in a magnetic field of approximately 1.8 Tesla.

WAMUS drift chambers

WAMUS is an acronym for Wide Angle Muon System. It consists of proportional drift tube chambers (PDT's). Its purpose is to provide muon identification and an independent momentum measurement. The PDT's are typically 250×560 cm^2. They are made of rectangular aluminum tubes, 10.1 cm wide. These aluminum tubes are the drift-cells. The PDT's outside the magnet have three decks of drift-cells. The PDT's inside the magnet (the A-layer) have four decks of drift cells, with the exception of the bottom PDT's inside the magnet, which have three decks. Figure 2.8 shows a cut-away view of the three-deck and four-deck PDT's. Each driftcell is equipped with Vernier pads, as shown in Fig. 2.8. The Vernier pads provide an additional measurement of the hit-coordinate along the signal wire. The pad read-out however is only instrumented in the layer inside the toroid and in about 10% of the other two layers, because with the addition of scintillators in Run II, the Vernier pad measurement only represents a minor resolution improvement.

The signal from the PDT wires is the time measured on the anode wire in the center of the cell. The two wires of neighboring cells are connected with a delay, so that two times are measured for each hit. Knowing the propagation times of the signal along the wire and the delay-time, it is possible to calculate the drift-time and axial time (time of propagation along the wire). After calibration, these times can be converted into distances. The resolution on the drift distance is of the order of 500 microns. The resolution on the axial distance varies with the location of the hit along the wire. If the hit occurs far from the electronics, the resolution is approximately 10 cm. If the hit occurs near the read-out, the signal has to propagate two wirelengths, and due to dispersion of the signal the resolution on the axial distance will be approximately 50 cm.

Cosmic cap and bottom scintillation counters

The "Cosmic Cap" scintillation counters cover the top, sides and bottom of the PDT C-layer. They are used to determine the timing of muon tracks in the PDT chambers relative to the bunch

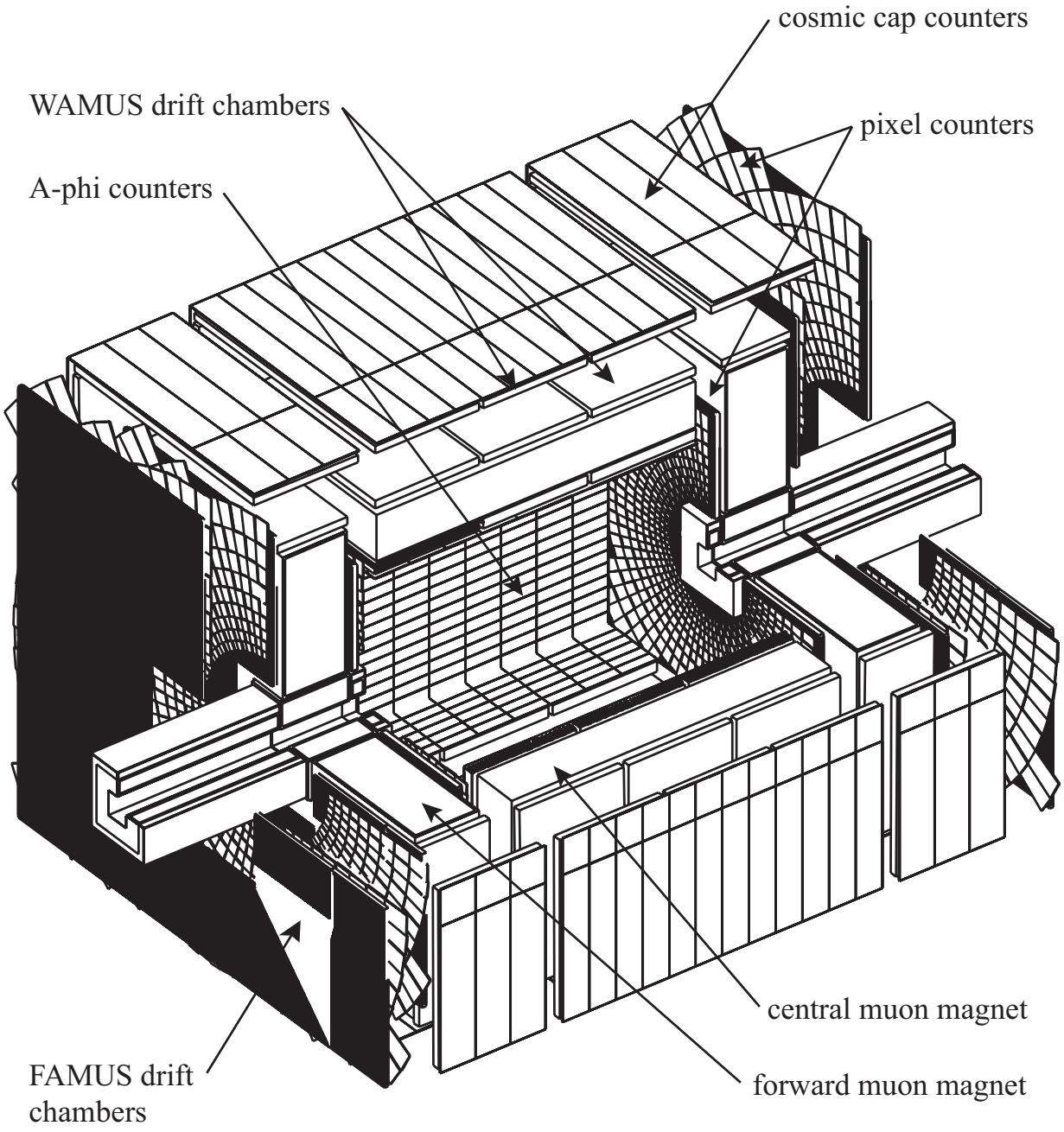


Figure 2.7: The muon system.

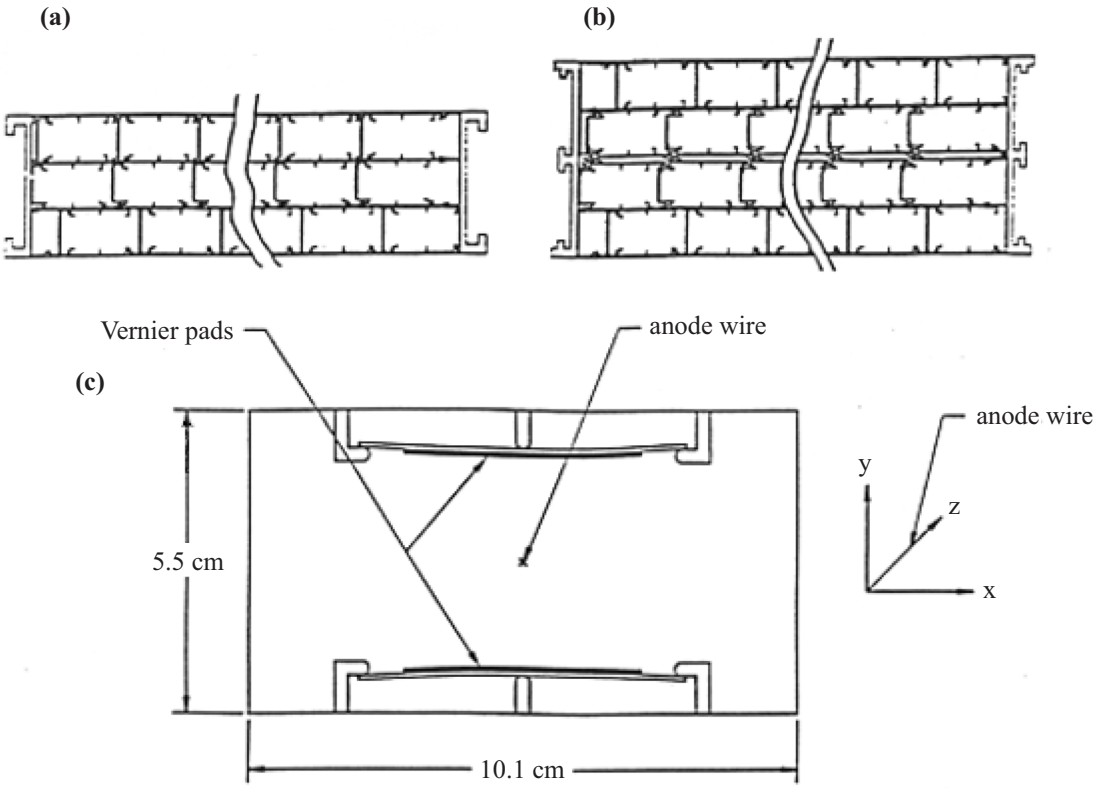


Figure 2.8: Cut-away sketch of PDT chambers. Two types of chambers are shown: The three-deck type (a) and the four-deck type (b). Also a view of a drift tube is shown (c), indicating the signal wire and Vernier pads.

crossing and thereby determine from which bunch crossing they originate. Similarly they are used to reject cosmic rays, by determining that the timing of the track is inconsistent with any bunch crossing. The limited coverage of the counters underneath the bottom C-layer, due to the detector support structure, is supplemented with counters underneath the PDT B-layer.

A-phi scintillation counters

The A-phi counters have been mounted on the inside of the A-layer PDT's. Similarly to the Cosmic Cap counters, the A-phi counters measure the timing of muon tracks in the PDT's they are mounted on, to determine the bunch crossing of the track. They are also used to reject out-of-time particles. For the A-phi counters, however, these are mostly not cosmic particles, but backscatter from the forward direction. The A-phi counters are 4.5° wide, matching the expected multiple scattering for high-p_T muons. One of the arguments for the construction and installation of the A-phi counters, was to use them in a level 1 $J/\psi \rightarrow \mu^+\mu^-$ trigger [39]. The p_T threshold for muon identification in Run I was around 4 GeV, corresponding to the energy needed to penetrate both the calorimeter and the muon toroid. Tracks that were stopped in the muon toroid were not usable (so-called A-stubs), due to the high background of out-of-time hits in the PDT A-layer. With the addition of the A-phi counters, these background hits have been reduced, thereby improving the signal to noise ratio of the A-stubs to a useful level. The threshold for muon identification could therefore be lowered to 1.1 – 1.4 GeV.

FAMUS drift chambers

The Forward Angle Muon System, or FAMUS, is made of 3 layers of mini drift tube detectors (MDT's), and covers the range $1 < |\eta| < 2$, on both sides of the interaction point. An MDT is rectangular in shape and contains a stack of 8 drift-cells. "Iarocci tubes" have been chosen as the technology for the drift-cell, because this technology is known to be robust, reliable, and efficient [43]. To increase efficiency, the FAMUS A-layer contains four layers of MDT tubes, and the B- and C-layers each contain three layers. The Iarocci drift-cells have an efficiency close to 100%. However, each 9.4 mm wide cell is separated from the next by a 0.6 mm thick wall, which introduces an inefficiency of approximately 6%. Also, the efficiency is reduced near the wire support in each cell. Furthermore, there are dead areas between octants, for gas, high voltage and signal connectors. Overall, the efficiency for finding two hits on a muon track in one layer of the MDT detector (out of three expected) is around 90%. The momentum resolution of the MDT detector is limited by the coordinate resolution of the detector, as well as multiple scattering in the toroid. The drift distance measurement by the Iarocci tubes has an accuracy of around 0.7 mm, and after digitization in 18.8 ns bins, the accuracy is about 0.9 mm. With 3 or 4 planes, the resolution is effectively about 0.7 mm. Taking into account this resolution, the multiple scattering in the toroid and the lay-out of the planes, the forward muon system has a momentum resolution of around 20% for low momentum muons [44].

	event rate	time available
level 1	2.5 MHz	4.2 μs
level 2	6 kHz	100 μs
level 3	1 kHz	150 ms

Table 2.3: Input event rates at different levels of the trigger system. The level 3 accept rate equals the event rate to storage and lies between 20 and 50 Hz.

Pixel scintillation counters

The purpose of the forward trigger counters, or pixel scintillation counters, is to provide accurate time information to supplement the measurements from the MDT's. This allows substantial reduction of the background rates in the MDT's. The layout of the pixel counters in the A-layer is shown in Fig. 2.7, the B- and C-layer are analogous. The ϕ segmentation of the counters is 4.5°, and the segmentation in η is 0.1. This segmentation was chosen to reasonably limit the total number of counters over the total area to be covered. The area of the C-layer is 12×10 m^2, and there is a total of almost 5000 counters in this system.

2.2.5 Luminosity monitor

The device measuring the instantaneous luminosity at DØ is visible at the bottom right in Fig. 2.4. The luminosity is determined from the total event rate for $p\bar{p}$ inelastic interactions. This rate is measured with scintillator tiles mounted close to the beam pipe on the front surfaces of the end calorimeters. From the event rate, the luminosity can be calculated using the total cross-section for $p\bar{p}$ inelastic interactions and the known acceptance of the luminosity monitor [45].

For the lifetime measurement presented in this thesis, the luminosity is not an important quantity. It will be used to calculate the expected event yield in section 4.5.2.

2.2.6 Trigger

The overwhelming majority of proton-antiproton encounters results in events of little interest. To accumulate a large sample of events of interest without having to store and reconstruct a staggering number of uninteresting events, DØ employs an event trigger which decides whether to store an event or to disregard it. The trigger system is organized in three major levels, the hardware based level 1 and level 2, and the software based level 3. In short, these levels are referred to as L1, L2 and L3. Each level examines the event in more detail than lower tiers and restricts the rate of events to higher tiers. This is reflected in the amount of decision time per event that each level has available, as indicated in Table 2.3.

The hardware L1 trigger takes input from the calorimeter, the muon system and the luminosity monitor, with an event rate of 2.5 MHz. As output, it produces a list of so-called trigger terms. The L1 framework, employing a series of field programmable gate arrays (FPGA's), examines these trigger terms and issues an L1 accept when certain criteria are met.

The level 1 muon subsystem takes input from the drift chambers and the muon scintillators. Until recently only scintillator information has been available in the level 1 muon trigger. To record the data used in this thesis, a level 1 dimuon trigger was used, requiring coincidence between a scintillator in the A-layer of the muon system and a scintillator in the B- or C-layer for each muon separately. A line through the two scintillators should approximately point back to the interaction point. This is a very efficient trigger for high-p_T muons, but low-p_T muons can be stopped either in the calorimeter or in the iron magnet core between the A- and B-layers of the muon system. On average an energy of around 4 GeV is required to pass through the calorimeter and muon toroid (see Fig. 2.9). For this trigger, no further requirements at higher trigger levels are necessary, because the accept rate is low enough.

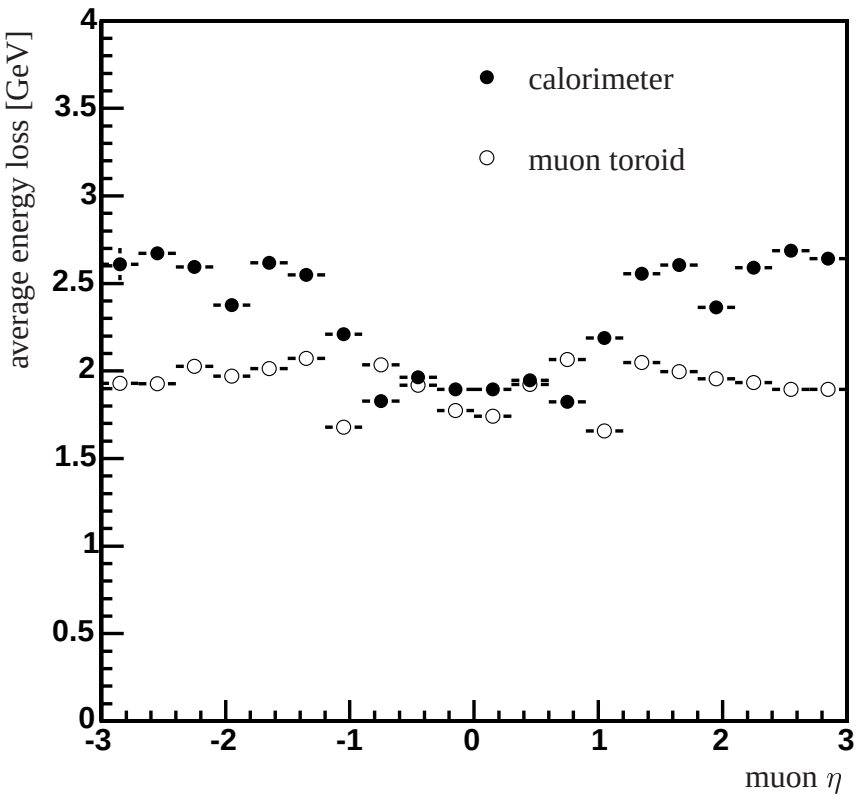


Figure 2.9: The average energy lost by a muon when passing through the calorimeter and muon toroid, as a function of η. The values shown here are from a parametrization based on the position where the muon strikes the muon system. Given an assumption for the primary vertex, the amount of absorber that has been traversed can be calculated.

The L2 trigger is the first trigger to match information from different subdetectors. It is comprised of two stages, a preprocessor stage and a global processor stage. At the preprocessor stage, each individual subdetector forms physics objects. At the global processor stage, physics objects from the different preprocessors can be combined to make the final L2 trigger decision.

Once an L2 accept has occurred, digitized data is loaded onto the so-called single board computers (SBC's). Data from the SBC's are transferred to the L3 farm. The farm presently

contains 150 processors, and can easily be expanded to extend the allowed processing time. Each processor of the farm, called a “level 3 node”, runs an independent instance of the L3 filtering software and processes a complete event and makes the trigger decision. Events that are passed, are transmitted to the datalogger, from where they are copied to tape.

The basic elements of the L3 filtering software are the *tools*, which contain the L3 reconstruction algorithms. These algorithms mimic the reconstruction algorithms used offline, but have improved processing speed at the cost of some precision. The reconstructed objects provided by the tools (e.g. tracks, muons, jets) are passed to *filters*, which apply quality criteria. The *filter scripts* are easily modifiable lists defining the trigger parameters.

Chapter 3

Event reconstruction

This thesis presents a measurement of the B_d^0 lifetime, using decays to $J/\psi K_S^0 \to \mu^+\mu^-\pi^+\pi^-$. This measurement requires the decay length and the p_T of the B_d^0 meson to be reconstructed, to determine its proper decay time. This chapter describes the procedures used to extract these quantities from the data: Reconstruction of tracks, primary and secondary vertices, and the identification of tracks as muons. We will also discuss the measurements of the decay length and p_T.

3.1 Track reconstruction

For this analysis, the Alternative Algorithm (AA) for finding tracks has been used [46]. It has the lowest p_T cut-off of the algorithms available, and the highest efficiency for high-impact parameter tracks, which is important for K_S^0 reconstruction. Compared to other algorithms, it is somewhat less tolerant of misalignment of detector elements, and this makes it slightly less efficient for straight, high-p_T tracks.

The general scheme of the track reconstruction in this algorithm, as indicated in Fig. 3.1, begins with the construction of track hypotheses. If the track hypotheses fulfill a number of requirements, they are stored in an intermediate pool of track hypotheses. The hypotheses in this pool are then filtered to remove the fakes and select the actual tracks.

Construction of track hypotheses

The basic read-out elements of the SMT and CFT are single strips or fibers. Charged particles can fire multiple strips, especially if they cross the detector at an angle different from 90˚. Therefore,

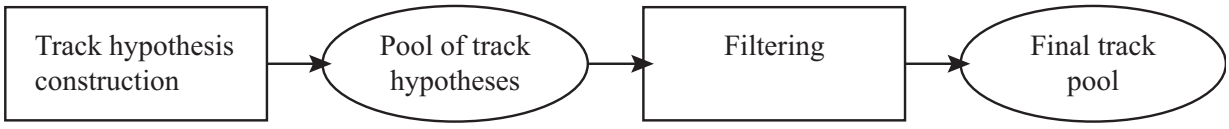


Figure 3.1: The general scheme of track reconstruction in the AA algorithm.

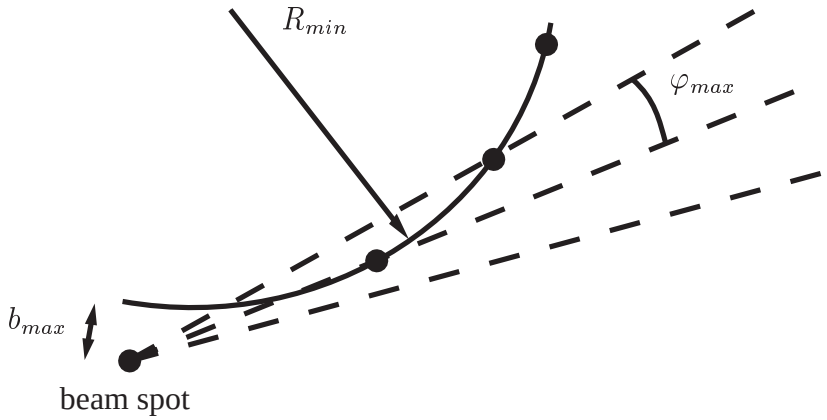


Figure 3.2: Parameters defining the acceptance of the AA tracking algorithm.

φ_{max}	0.2 rad
R_{min}	30 cm
b_{max}	5 cm

Table 3.1: Setting of parameters defining the AA acceptance.

if the signals in neighboring strips or fibers are over threshold, they are merged into a cluster of strips. These clusters are referred to as "hits".

A track hypothesis is constructed from three two-dimensional hits, which can be CFT clusters or SMT clusters in barrels or disks. The hits that are added after the first one are on layers beyond the layer having the first hit, i.e. on layers at larger radii. The second hit is added, provided that the φ-difference in the transverse plane is less than φ_{max} (see Fig. 3.2). The third measurement is selected from a layer beyond the layer with the second hit. We can now draw a circle through the first three hits. The third hit is only accepted if the radius of this circle is larger than R_{min}, and if the "impact parameter" of the circle with respect to the beam spot is smaller than b_{max}.

The parameters φ_{max}, R_{min} and b_{max} define the acceptance of the tracking algorithm. Their values are listed in Table 3.1. The setting $R_{min} = 30$ cm corresponds to a p_T cut-off of 180 MeV. This parameter set is labeled "AA extended", because it is looser than the default cuts used in the AA algorithm. These cuts make the algorithm run about four times slower than the default cuts, but the efficiency for finding high-impact parameter tracks is much improved, resulting in 80% more reconstructed K^0_S's, depending on selection cuts [47].

Stereo information

Almost all SMT layers can measure the z-coordinate of a hit as well, using stereo strips. Because each strip crosses multiple stereo strips, and given a typical hit occupancy in the SMT of 1%, each axial hit typically has multiple associated stereo hits. Each track hypothesis therefore contains one axial projection, and one or more stereo projections, using all possible combinations of stereo hits. With the addition to the track hypothesis of more axial hits, the number of stereo projections is generally reduced, because the stereo projections that are incompatible with the stereo hits associated with the new axial hit are deleted when the hit is added.

Completion of track hypotheses

Based on the initial three hits, a window is defined on the detector layer beyond the third hit, in which an additional hit is expected. If a hit is found in this window, it is added if it is compatible with the track, based on a χ^2 requirement.

The concept of missing hits is very important to both the tracking efficiency and fake rejection. If the window is completely covered by an active area of the detector, and no hit in the window is compatible with the track hypothesis, the track hypothesis is said to have a "miss". This ensures that there is no missing hit due to inactive detector areas. We define three types of misses. "Inside misses" are misses in between hits on the track. "Backward misses" are misses on layers inside the innermost hit on the track. This can happen with decay products of long-lived neutral particles, such as charged pions from a K^0_S decay inside the SMT. Finally, "forward misses" are misses on layers outside the outermost hit on the track.

We consider a track hypothesis complete if the last three extrapolations to detector layers yielded misses, or if all hits have been used. The complete track hypothesis is saved to the track hypothesis pool if it has:

- at least 4 stereo hits;
- less than 4 inside misses overall, and less than 3 in the SMT;
- less than 6 forward and backward misses;
- the number of hits is at least five times the number of misses.

We repeat this process for every allowed initial combination of three hits.

Filtering

To obtain the final track sample, we have to select the real tracks from the pool of track hypotheses, and remove the fakes. This procedure is called filtering. The first step is to order the tracks in decreasing number of track hits. The tracks with the same number of hits are ordered based on the number of misses (fewest misses first). The track with the same number of hits and misses are ordered in increasing χ^2.

A hypothesis is copied to the final track pool if it satisfies the "shared hits criterion". A hit is shared, if it has already been used on a track in the final track pool. We count the total number of hits N_{tot} and the number of shared hits N_{shared}. The "shared hits criterion" is satisfied if:

- $N_{shared} \leq \frac{2}{3} N_{tot}$, and
- $N_{shared} \leq \frac{1}{5} N_{tot}$ or $N_{tot} - N_{shared} > 3$.

To reduce the fake rate, we determine the primary vertex (PV) using this track sample. The algorithm for determining the PV is explained in the next section. For each track that has a small impact parameter with respect to this PV, we increase the hit count N_{tot} by 2. Using this updated

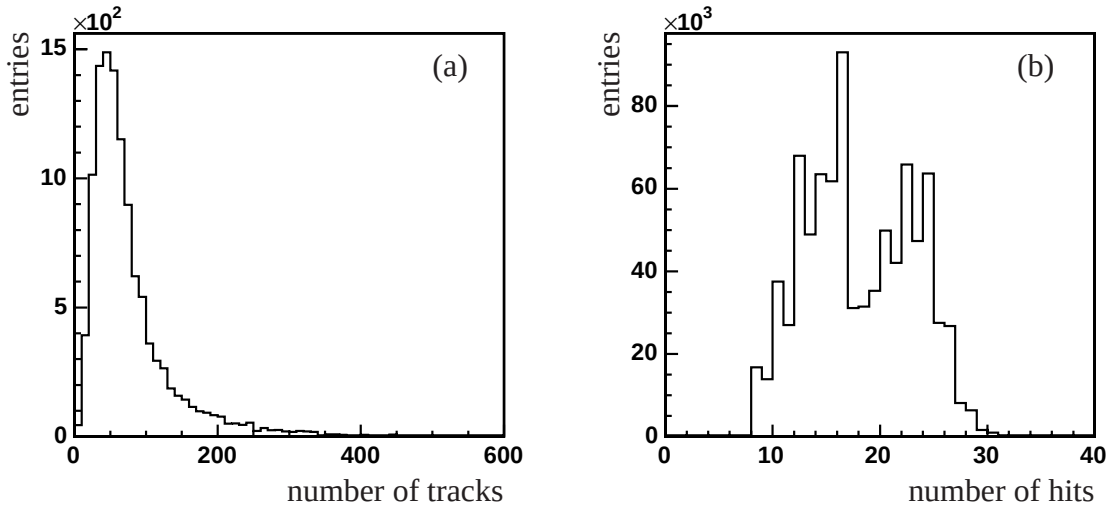


Figure 3.3: The AA track reconstruction algorithm on real data: (a) The number of tracks per event and (b) the number of hits per track.

N_{tot} the tracks are filtered a second time: They are ordered, and tested using the shared hits criterion. If they pass they are kept in the pool of final tracks. In this fake reduction procedure, tracks from the PV are considered before high-impact parameter tracks with the same number of hits. But this procedure does not eliminate high-impact parameter tracks, if they have a sufficient number of unique (not shared) hits.

3.1.1 Performance

The resulting number of tracks per event is indicated in Fig. 3.3 (a). We have used a subset of the data sample described in section 4.1. Figure 3.3 (b) shows the distribution of the number of (axial) hits on the tracks. The selection requirement of at least 4 stereo hits corresponds to a minimum of 8 axial hits. A typical track with $\eta = 0$ will cross 8 SMT layers and 16 CFT layers, which results in 24 hits. Figure 2.5 indicates that there is overlap between some ladders, allowing a track to acquire more than 24 hits. Figures 3.4 (a) and (b) display the number of missing hits and the number of shared hits per track. The AA algorithm was found to be 85 – 90 % efficient on a $Z \rightarrow e^+e^-$ real data sample, using standard cuts [48]. Standard cuts are $\varphi_{max} = 0.08$ rad, $R_{min} = 30$ cm and $b_{max} = 2.5$ cm. The extended cuts used for this thesis are listed in Table 3.1, and the efficiency for the high-p_T, isolated tracks in the $Z \rightarrow e^+e^-$ sample is expected to be identical when using extended cuts. The efficiency that is lost is attributed to the hit efficiency in the tracking detectors.

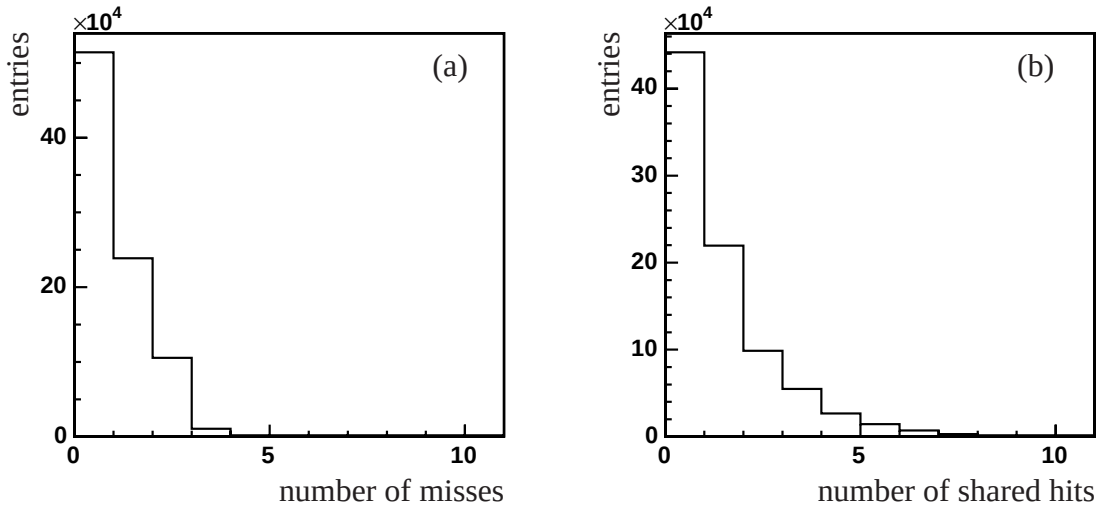


Figure 3.4: The AA track reconstruction algorithm on real data: (a) The number of "misses" per track and (b) the number of shared hits per track.

3.2 Primary vertex reconstruction

The primary vertex (PV) reconstruction is a two-pass procedure. On the first pass, an initial selection of tracks close to the beam spot is made, to prevent the PV from being biased by particles with significant lifetime. The beam spot is a run-by-run average of the PV. The requirement on the tracks for inclusion in the PV determination is a distance of closest approach (DCA) smaller than 2 mm. The vertex is fit, using the beam spot as a constraint. On the second pass, an attempt is made to include additional tracks in the PV to improve its accuracy. The difference with the first pass is that tracks are selected, based on their DCA to the PV from the first pass. The original implementation of this algorithm is described in Ref. [49]. From the reconstructed primary vertex candidates, we select the candidate with the highest track multiplicity. A minimum track multiplicity of 5 is required. If no PV with 5 tracks or more is found, the event is abandoned. Less than 0.5% of the events in our data sample, that will be described in section 4.1, fails this requirement.

3.2.1 Resolution

We have studied the accuracy of the determination of the PV position using a Monte Carlo sample of $B^0_d \to J/\psi K^0_S$ decays. The technical details of this sample are described in section A.1. A comparison of the reconstructed PV to the true PV is shown in Figs. 3.5 and 3.6. Figure 3.5 shows the distribution of the residuals for (x, y, z) coordinates. Figure 3.6 shows the distribution of the residuals for the direction of flight of the B^0_d (labeled L), and the residual in the coordinate transverse to this direction (labeled T, see Fig. 3.7).

The distributions in Fig. 3.5 are fit using two Gaussians with different widths but identical

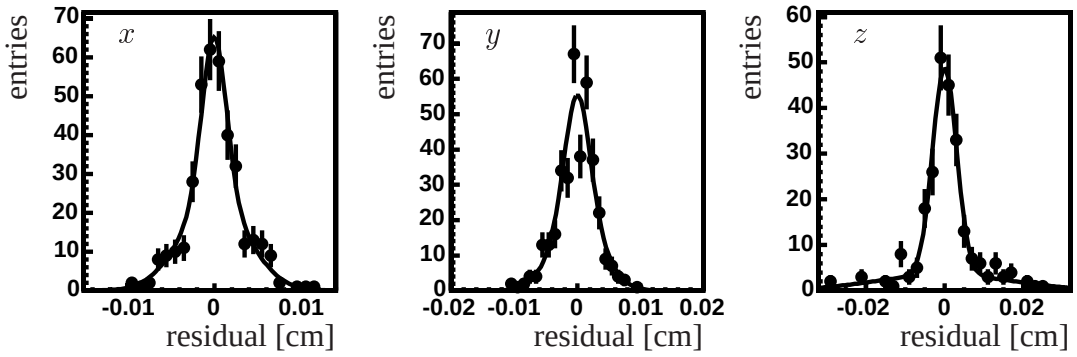


Figure 3.5: Distance from reconstructed PV to MC true PV in the x, y and z coordinate.

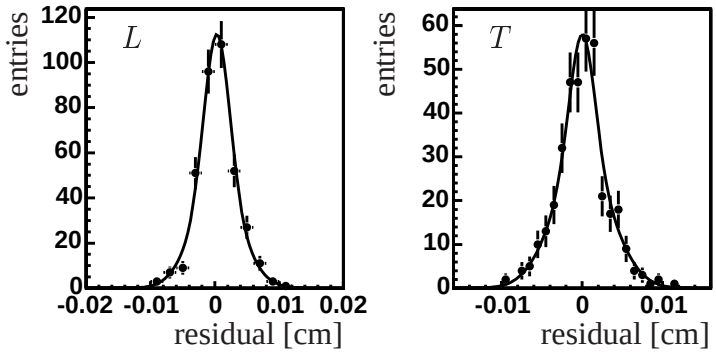


Figure 3.6: Distance from reconstructed PV to the MC true PV in the L and T coordinate (see Fig. 3.7).

Coordinate	mean [μm]	std.dev. [μm]	σ_1 [μm]	σ_2 [μm]	α
x	0.7 ± 1.02	25.7 ± 1.79	13.2 ± 1.94	33.1 ± 2.54	0.47
y	1.5 ± 0.92	24.6 ± 2.24	14.5 ± 1.76	38.2 ± 4.33	0.68
z	1.6 ± 1.69	54.0 ± 5.42	27.5 ± 2.77	107.3 ± 13.29	0.80
L	0.8 ± 0.99	23.4 ± 2.18	17.0 ± 1.74	36.1 ± 5.00	0.75
T	0.0 ± 1.07	27.1 ± 2.13	17.2 ± 1.49	41.2 ± 4.24	0.69

Table 3.2: Standard deviation of PV residual distributions, in different coordinates.

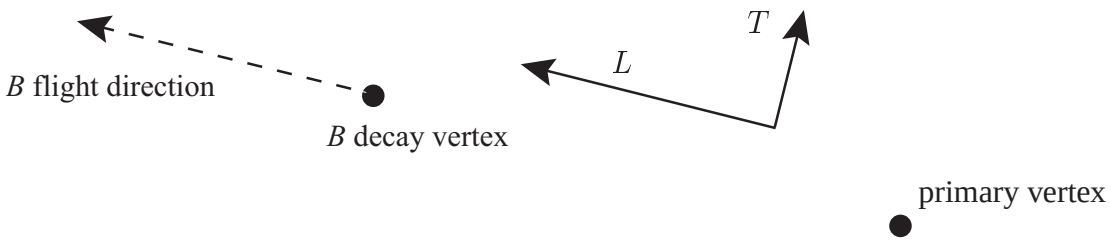


Figure 3.7: The definition of L and T coordinates, as used in describing the vertexing resolution.

means:

$$f(x) = C \cdot \left(\frac{\alpha}{\sigma_1\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma_2^2}} \right), \tag{3.1}$$

and the results are given in Table 3.2. Due to the cylindrical symmetry of the tracking detectors, we do not expect the distributions in x and y to differ significantly. The distribution along z, however, is expected to be wider, because the hit resolution in x and y is better than in z. This is the case for both the SMT and the CFT.

A different effect plays a role in the L distribution. If the event contains one or more long-lived particles (such as a B-meson), the inclusion of tracks from the decay of such a particle can bias the PV toward its decay vertex. In the case of B-mesons, this effect can also cause the PV to be biased away from its decay vertex, because B-mesons are typically produced in back-to-back pairs. A bias toward one B will therefore generally mean a bias away from the other B's vertex. This effect will result in a degradation of the PV resolution in the L direction. Table 3.2 shows that our reconstruction of the primary vertex is not biased in the L-direction, nor is the resolution degraded.

3.3 Secondary vertex reconstruction

In this section we will describe the calculation of vertices of two oppositely charged tracks. It can easily be extended to using additional tracks. The secondary vertex finding algorithm, called the "impact parameter technique", is described in more detail in Ref. [50]. It is an extension of an algorithm originally implemented by experiments at LEP, as described in e.g. Ref. [49]. The algorithm was designed to make optimal use of the fact that the vertex resolution of the detector is significantly better in the plane perpendicular to the beam-line, than in the coordinate along the beam-line. Therefore, the algorithm deals with the two impact parameters of the track in these two orientations separately.

A charged particle in a constant magnetic field, as long as it doesn't scatter, follows a helical path. We approximate the field in the DØ central tracking volume with a constant field, parallel to the beam line. That is, we assume the tracks follow helices. The projections of these helices on the transverse plane are circles. When we have two circles, two basic configurations are possible: Either they cross each other (in two points), or they are separated, as is shown in Fig. 3.8 (a). Of course, if the tracks originate from a common parent particle, the tracks will cross (or at least *touch*). However, given two tracks, without applying a vertexing algorithm we cannot know whether the two tracks come from one parent. Also, if two tracks come from one parent, due to finite measurement accuracy, the tracks can appear as being separated.

In case the tracks are separated, the first estimate for the vertex position is the point directly in between the two circles. In case the tracks cross, there are always two crossing points. The coordinates of both crossing points are calculated as follows. From the tracks, we know their radius of curvature r_a and the center of the circle (x_a, y_a), with a equal to 1 or 2 for the two tracks respectively. We can now calculate d, which is $\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$ (see Fig. 3.8 b).

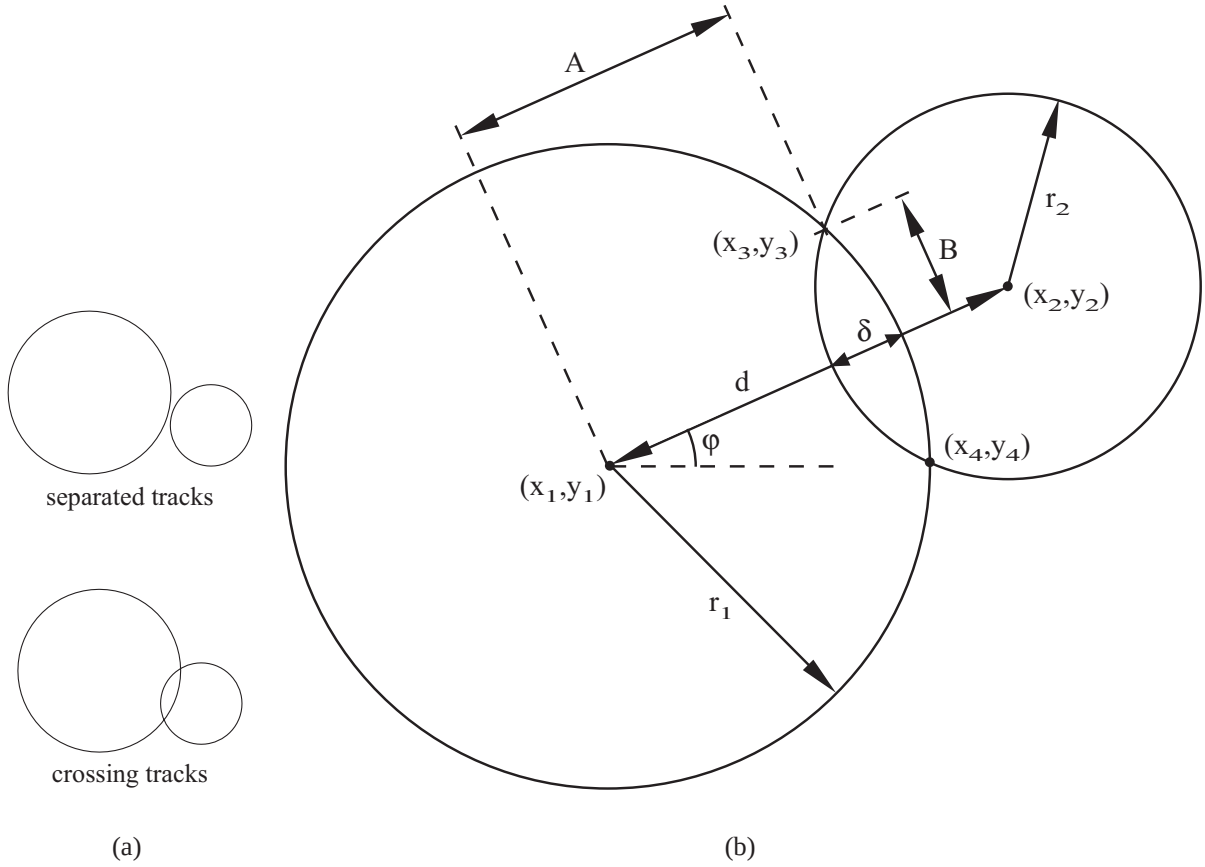


Figure 3.8: Vertexing two oppositely charged tracks in the transverse plane.

Forming two triangles from the two circle centers and one of the crossing points allows us to find A and B:

$$A = \frac{r_1^2 - r_2^2 + d^2}{2d} \qquad B = \sqrt{r_1^2 - A^2}. \tag{3.2}$$

The angle φ in Fig. 3.8 (b) is $\arctan\left[(y_2 - y_1)/(x_2 - x_1)\right]$. Using A and B we can now calculate the first crossing point, by rotating the vector (A, B) by φ and translating (x_1, y_1) with the result:

$$x_3 = x_1 + A\cos(\varphi) - B\sin(\varphi), \tag{3.3}$$
$$y_3 = y_1 + A\sin(\varphi) + B\cos(\varphi). \tag{3.4}$$

Analogously for the other crossing point, by taking the opposite sign of B:

$$x_4 = x_1 + A\cos(\varphi) + B\sin(\varphi), \tag{3.5}$$
$$y_4 = y_1 + A\sin(\varphi) - B\cos(\varphi). \tag{3.6}$$

We now have to decide which of the two crossing points to use as first assumption of the vertex position. We select the crossing point, for which the distance along z between the two tracks, Δz, is smallest. As an initial estimate of the z-coordinate of the vertex position we use the midpoint between the two points of closest approach (PCA) of the two tracks to the original crossing point. Referring back to Fig. 3.8 (b), the x and y coordinates of this first estimate are either (x_3, y_3) or (x_4, y_4).

We are now able to make a preselection on the found vertices, before continuing with the more precise, but also more time-consuming calculation of the final vertex position. In case the tracks cross, we abandon the vertex if Δz is greater than 5 cm, and δ is larger than 3 mm (see Fig. 3.8 b). We also reject this vertex if the tracks do not cross and δ is larger than 3 mm.

From this point on the algorithm is identical for both possible configurations of tracks. The residual in the transverse plane d_1 of each track with respect to the final vertex position $\vec{r}$, can be written as:

$$d_1 = \left| \vec{E} \cdot (\vec{r} - \vec{P}) \right|, \tag{3.7}$$

where we defined $\vec{P}$ as the vector from the origin to the PCA of the track and $\vec{E}$ as the unit vector pointing from the PCA of the track to the final vertex. We define η, starting from equation 3.7:

$$d_1 = \left| \vec{E} \cdot \vec{r} - \vec{E} \cdot \vec{P} \right| = \left| \vec{E} \cdot \vec{r} - \eta \right|. \tag{3.8}$$

Figure 3.9 illustrates these definitions.

In addition we introduce the impact parameter in z, labeled d_2. To calculate d_2, a small extrapolation is necessary. In this extrapolation, we treat the track as a straight line, so that we can write d_2 as:

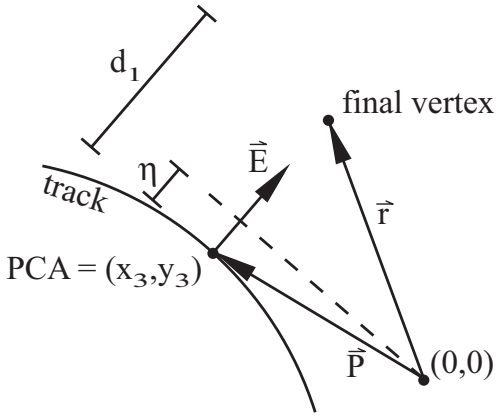


Figure 3.9: Definition of the impact parameter in the transverse plane. See text for explanation.

$$\begin{aligned} d_2 &= |r_z - P_z + \cot(\theta) \cdot [(r_x - P_x)\sin(\phi_{per}) - (r_y - P_y)\cos(\phi_{per})]| \\ &= |\cot(\theta)\sin(\phi_{per}) \cdot (r_x - P_x) - \cot(\theta)\cos(\phi_{per}) \cdot (r_y - P_y) + r_z - P_z| \,. \end{aligned} \quad (3.9)$$

Here ϕ_{per} is the angle of the track in the transverse plane at the point $\vec{P}$, and θ is the angle of the track with the beam axis. If we define the vector $\vec{L} = (\cot(\theta)\sin(\phi_{per}), -\cot(\theta)\cos(\phi_{per}), 1)$, the expression for d_2 becomes:

$$d_2 = \left|\vec{L} \cdot (\vec{r} - \vec{P})\right| \quad (3.10)$$

$$= \left|\vec{L} \cdot \vec{r} - \vec{L} \cdot \vec{P}\right| = \left|\vec{L} \cdot \vec{r} - \xi\right| . \quad (3.11)$$

The final vertex position $\vec{r}$ is obtained by minimizing a χ^2 function, representing the compatibility of the residuals of all tracks with zero:

$$\chi^2(\vec{r}) = \sum_a \vec{d_a}^T (V_a^{-1}) \vec{d_a} = \sum_a \chi_a^2(\vec{r}), \quad (3.12)$$

where the index a runs over all tracks, and χ_a^2 is the contribution of each track to the total χ^2 of the vertex. $\vec{d_a}$ is the 2-dimensional vector with components d_1 and d_2 for track a, given in equations 3.8 and 3.11 resp. V_a is the (2×2) covariance matrix of $\vec{d_a}$:

$$V = \begin{pmatrix} \sigma_{d_1}^2 & \sigma_{d_{12}}^2 \\ \sigma_{d_{12}}^2 & \sigma_{d_2}^2 \end{pmatrix} . \quad (3.13)$$

Here $\sigma_{d_{12}}^2$ is the covariance of d_1 and d_2. The determinant of V is $\sigma_{d_1}^2\sigma_{d_2}^2 - \sigma_{d_{12}}^4$. Because $\sigma_{d_{12}}$ is expected to be small with respect to σ_{d_1} and σ_{d_2}, we drop the $\sigma_{d_{12}}^4$ term from the determinant. This simplifies the inverse of V to:

$$(V^{-1}) = \begin{pmatrix} \frac{1}{\sigma_{d_1}^2} & \frac{-\sigma_{d_{12}}^2}{\sigma_{d_1}^2 \sigma_{d_2}^2} \\ \frac{-\sigma_{d_{12}}^2}{\sigma_{d_1}^2 \sigma_{d_2}^2} & \frac{1}{\sigma_{d_2}^2} \end{pmatrix}. \tag{3.14}$$

Substitution of equations 3.8, 3.11 and 3.14 into equation 3.12 yields (summation over indices implied):

$$\chi_a^2 = \mathcal{A} - 2 \cdot \mathcal{B}_i \cdot r_i + \mathcal{C}_{ij} \cdot r_i \cdot r_j, \tag{3.15}$$

with $\mathcal{A}$, $\vec{\mathcal{B}}$, and $\mathcal{C}$ given by:

$$\mathcal{A} = \left(\frac{\eta}{\sigma_{d_1}}\right)^2 + \left(\frac{\xi}{\sigma_{d_2}}\right)^2 + 2 \cdot \frac{\eta \cdot \xi}{\sigma_{d_{12}}^2}, \tag{3.16}$$

$$\mathcal{B}_i = \frac{\eta \cdot E_i}{\sigma_{d_1}^2} + \frac{\xi \cdot L_i}{\sigma_{d_2}^2} + \frac{\xi \cdot E_i + \eta \cdot L_i}{\sigma_{d_{12}}^2}, \tag{3.17}$$

$$\mathcal{C}_{ij} = \frac{E_i \cdot E_j}{\sigma_{d_1}^2} + \frac{L_i \cdot L_j}{\sigma_{d_2}^2} + 2 \cdot \frac{L_i \cdot E_j}{\sigma_{d_{12}}^2}. \tag{3.18}$$

The vector $\vec{r}$ that minimizes χ^2 can now easily be found by taking the derivative of equation 3.15 to $\vec{r}$ and solving for $\vec{r}$:

$$\vec{r} = (C^{-1}) \cdot \vec{B}. \tag{3.19}$$

The matrix (C^{-1}) is the covariance matrix of the vertex position $\vec{r}$.

3.3.1 Mass-constrained fit

Some vertices can be identified as the decay vertex of a certain particle. Identifying properties are the mass, track multiplicity and decay length. If a vertex has been identified, we can improve the resolution of the vertex parameters by constraining the mass of the vertex to the world average of the parent particle's mass, and refitting the vertex. This procedure is called a mass-constrained fit. Vertex parameters that are typically improved after such a refit are obviously the mass, but also the position and momentum.

In our implementation of the mass-constrained fit, the p_T of the tracks is varied until the invariant mass is "correct". The variation of the p_T of each track is weighted with the inverse of the error on the p_T: The larger the error, the larger the allowed variation. Generally, mass-constrained fit implementations allow all track parameters to be varied, instead of just the p_T. The variations on all track parameters will then have to be weighted with their inverse errors. In our case however, the errors on the other track parameters were not sufficiently understood yet. Allowing them to contribute in the mass-constrained vertex fit does not further improve the resolution on the vertex parameters.

3.3.2 Resolution

Using our Monte Carlo sample of $B_d^0 \to J/\psi K_S^0$ events, we have compared the reconstructed B_d^0 vertex to the true B_d^0 vertex. The procedure and cuts used to select $B_d^0 \to J/\psi K_S^0$ decays will be discussed in the next chapter. To fit the distributions, we have again used two Gaussians with identical means, as defined in equation 3.1. The results are presented in Figs. 3.10 and 3.11 and Table 3.3.

Contrary to the case of the primary vertex, in the case of a B_d^0 vertex we do expect a different resolution in the L and T directions. The B_d^0 is required to have a boost to be reconstructed and selected. Therefore, the tracks of the decay products from the J/ψ and the K_S^0 will lie in a narrow cone. This means the opening angle between the tracks is small, and the resolution on the vertex position in the direction of flight of the B_d^0 will be degraded. The distribution in T on the other hand, will be narrow. The result in Table 3.3 follows this expectation. Moreover, the means of the residual distributions are consistent with zero, indicating no bias is present.

3.4 Decay length measurement

The direction of flight of the B_d^0 mesons can be measured by comparing the production and decay vertices, or it can be taken from the reconstructed B_d^0 momentum. These two methods are compared in Fig. 3.12. The direction determined from the vertex positions has an RMS of 93 mrad in the transverse plane, which improves to 24 mrad if a decay length of at least 1 mm is required. The resolution on the direction using the B_d^0 momentum is 7 mrad. Clearly the direction can be determined much more accurately using the momentum, especially at shorter decay lengths.

We can define the decay length of the B_d^0 mesons as simply the length of the vector from the production to the decay vertex, which we label $\vec{r}$. It is also possible to define the B_d^0 decay length as the projection of $\vec{r}$ onto the B_d^0 momentum vector, called the *projected* decay length. In Fig. 3.13 the measured decay length is compared to the true decay length. The residual of the projected decay length is much better behaved around zero and displays a Gaussian behavior. An added benefit is that the RMS of the distribution is reduced.

Throughout this thesis, when referring to the decay length, we will be referring to the measured transverse decay length, projected onto the B_d^0 direction of flight, as determined by its measured momentum.

3.5 Determination of $p_T(B_d^0)$

The measurement of the p_T of the B_d^0's relies on the result of the vertex fits of the J/ψ and the K_S^0. In fitting the J/ψ and K_S^0 vertices, a mass constraint is applied, which results in an improved momentum determination, both in terms of direction and p_T. The p_T of the B_d^0 is simply the sum of the J/ψ and K_S^0 momenta.

In MC data we can compare the measured p_T to the MC true p_T. The p_T spectrum of the B_d^0's is an exponentially falling distribution. As a selection cut, we have required the p_T of the

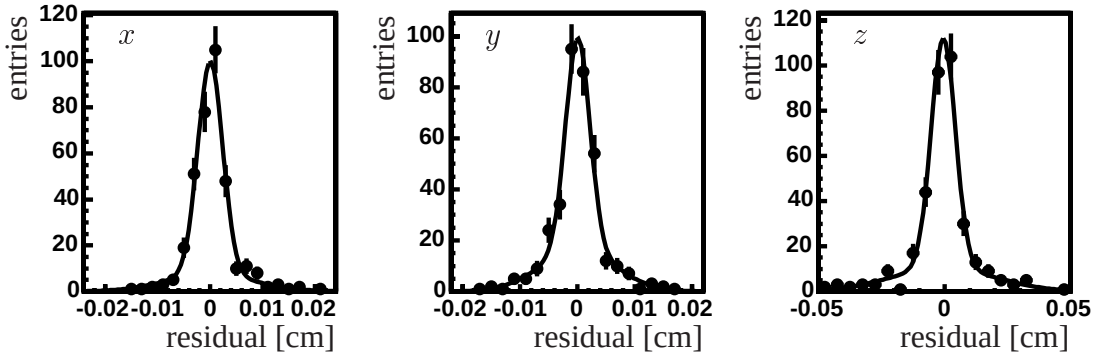


Figure 3.10: Distance from reconstructed B vertex to MC true B vertex in the x, y and z coordinate.

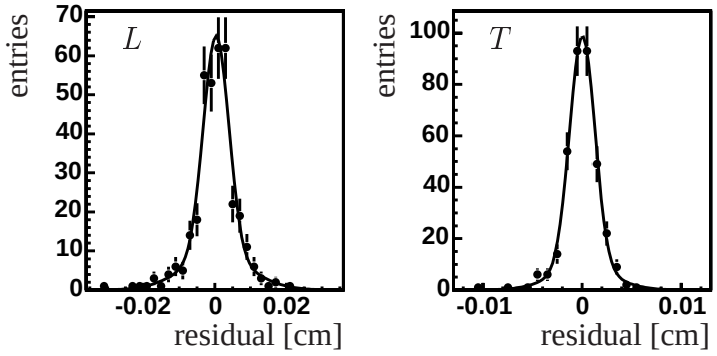


Figure 3.11: Distance from reconstructed B vertex to MC true B vertex in the L and T coordinate (see Fig. 3.7).

Coordinate	mean [μm]	std.dev. [μm]	σ_1 [μm]	σ_2 [μm]	α
x	0.5 ± 1.63	44.7 ± 4.05	23.7 ± 1.84	84.1 ± 9.56	0.78
y	0.9 ± 1.72	43.5 ± 3.50	21.1 ± 2.34	65.6 ± 6.06	0.63
z	-4.7 ± 3.58	120.2 ± 9.33	48.1 ± 3.87	205.9 ± 17.87	0.70
L	3.6 ± 2.55	59.3 ± 5.61	35.5 ± 3.77	96.0 ± 11.72	0.72
T	0.2 ± 0.83	17.5 ± 1.91	12.7 ± 1.26	31.1 ± 5.50	0.82

Table 3.3: Standard deviation of secondary vertex residual distributions, in different coordinates. The Cartesian coordinates (x, y, z) are defined in section 2.2.1. Coordinates L and T are defined in Fig. 3.7.

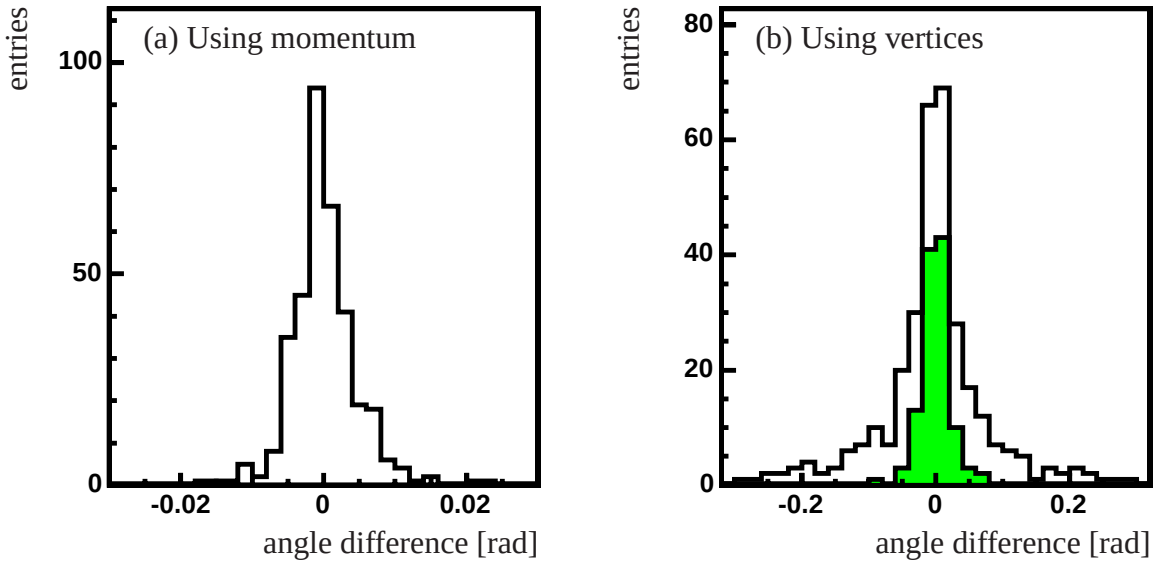


Figure 3.12: The angle in the transverse plane between the reconstructed and true B_d^0 flight direction. Figure (a) shows the result when the direction is determined from the reconstructed momentum vector. In Fig. (b) the direction has been determined from the production and decay vertices. The filled histogram shows the angle difference for B_d^0's with a decay length greater than 1 mm.

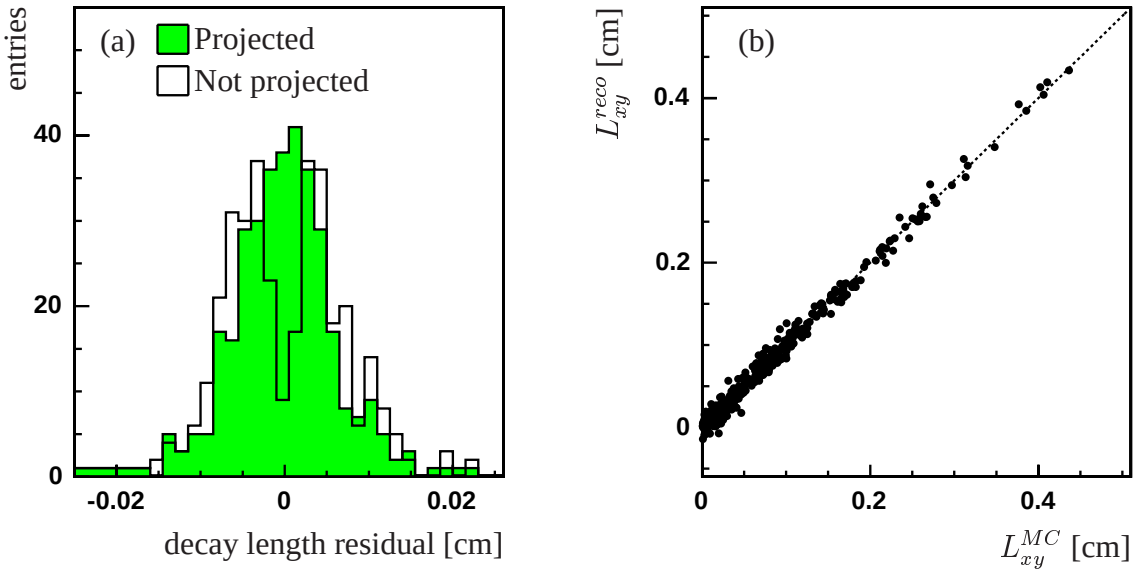


Figure 3.13: The difference between reconstructed and true transverse decay length. In (a), the open histogram is the result when we use the transverse component of the vector $\vec{r}$, connecting the production and the decay vertex. The filled histogram is the result after projecting $\vec{r}$ onto the momentum vector of the B_d^0. Figure (b) shows reconstructed L_{xy} as a function of MC true L_{xy}. The dashed line indicates $L_{xy}^{reco} = L_{xy}^{MC}$.

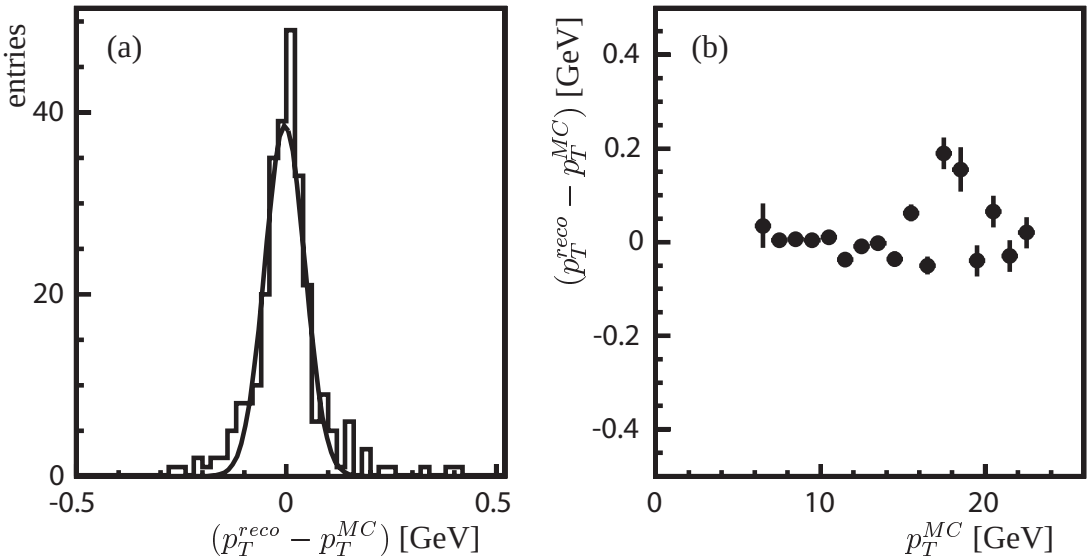


Figure 3.14: The difference between reconstructed p_T and MC true p_T of B_d^0's.

B_d^0's to be larger than 7 GeV (see section 4.5.3). The distribution of the difference between true and measured p_T is shown in Fig. 3.14 (a). To estimate the resolution, we have fit a Gaussian distribution. This fit yields a standard deviation of 42 ± 3 MeV. Strictly speaking, we don't expect the p_T resolution to be Gaussian. The quantity that is Gaussian distributed is q/p_T, with q the charge of the track. Fitting a Gaussian can only give a reasonable impression of the p_T resolution, averaged over all tracks. The mean of the Gaussian is consistent with zero, indicating that no bias is present. In Fig. 3.14 (b) we demonstrate that there is no bias as function of MC true p_T.

3.6 Muon identification

The identification of tracks found in the central tracking systems (SMT, CFT) as muons proceeds in two distinct steps: The reconstruction of tracks in the muon system (local track reconstruction), and the matching of the reconstructed objects to central tracks. We will describe the algorithms in the sections that follow.

3.6.1 Local muon reconstruction

The reconstruction of tracks in the muon system begins with the reconstruction of hits in the scintillators and wire chambers (PDT's and MDT's). Because the wire chambers are outside of the toroidal field, straight track segments can be formed within each chamber. This gives a measure of the direction of the track through the chamber, and the position in the plane perpendicular to the wires (the drift plane). However, the coordinate along the wire (the axial coordinate) is not well measured using the wire hits. To improve this, segments consisting of wire hits are

matched to scintillator hits, which have a finer granularity than the wire hit resolution in the axial coordinate. When the track segments have been formed, they are fit into a track to determine the momentum. A picture of a muon system track with associated wire and scintillator hits is shown in Fig. 3.15. Finally, a spatial match with a track in the inner tracker is made, and a global refit of the track parameters is performed, if possible. We will describe these steps below.

Hit reconstruction

The scintillators provide accurate timing information. This can be used to reject cosmic ray particles, because cosmic rays have no time relation to the bunch crossing. As explained above, scintillator hits are also used to improve the accuracy of the axial position of wire segments. This is most important in the forward system, for two reasons. First, the scintillators in the forward system (pixel scintillators) are smaller in size. Secondly, while the PDT's in the central system have a dual time read-out providing some axial position information, the MDT's do not have such functionality. In the forward system, the scintillators are the only measurement of the coordinate along the wire.

The PDT cells are much wider than the MDT cells (10 cm vs. 1 cm cell width). Drift times in the PDT's are therefore larger, and the conversion of the drift time to drift distance is a more important issue. The time-to-distance relation has been derived using a test-stand setup consisting of four layers of PDT chambers plus a layer of scintillators, providing a trigger [51]. The time-to-distance relation for the MDT's has been measured using a similar setup.

Straight track segment reconstruction

The algorithm for the reconstruction of straight line segments is described in detail in [52]. The algorithm operates in the drift plane. It assigns two-dimensional hits to each drift circle, to account for the left-right ambiguity from the circles. If we consider a muon detector layer in the central top part of the detector, the track direction will be roughly upward. In this case, the two-dimensional hits are placed directly to the left and right of the wire. All possible straight track segments within a layer are constructed. The resulting segment candidates are fit to determine the optimal values for the segment parameter and to allow a χ^2 quality cut to reduce the fake rate.

After forming track segments based on wire hits as described above, the segment will be matched to a scintillator hit, if present. After matching a scintillator hit, the axial segment position will have improved, and therefore the drift distances of the wire hits can be calculated more accurately. After this has been done, the segment is refit.

There is only very little material and negligible magnetic field between the B- and C-layers in the muon system. The segments in the B- and C-layer are therefore expected to lie on a straight line. Compatible B- and C-layer segments are merged, and refit to further improve the resolution.

Muon track fit

We define a BC-segment as a merged B- and C-layer segment, or a single B- or C-layer segment. We further define the center of gravity (c.o.g.) of a segment to be the c.o.g. of the hits it uses.

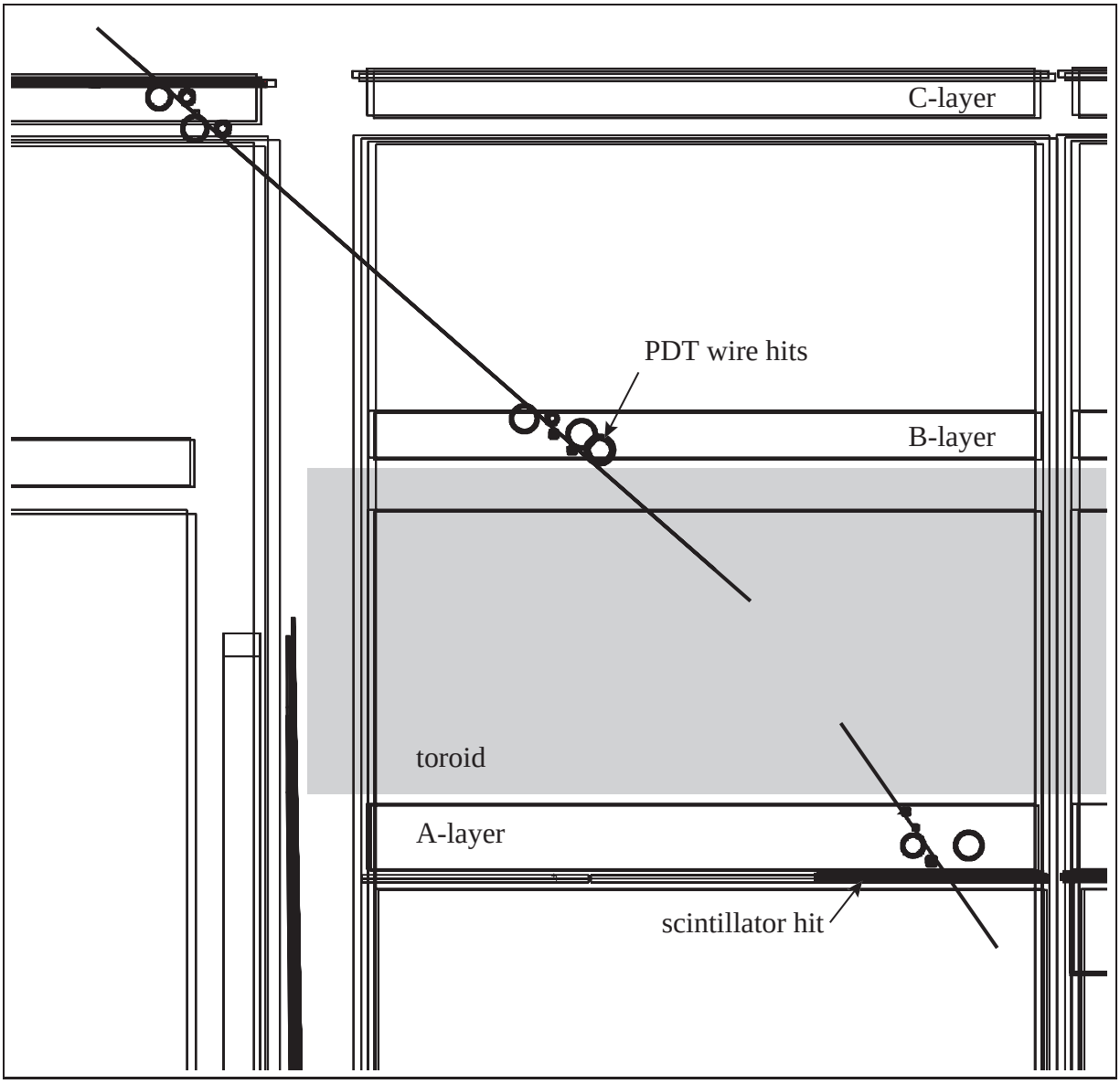


Figure 3.15: A local muon track with associated PDT wire hits and scintillator hits.

To fit the reconstructed track segments into tracks and determine the momenta, BC-segments are matched to A-layer segments. A fit is then performed to extract the track parameters [53]. This fit starts from an analytically calculated first estimate of the particle momentum and trajectory through the toroid, using a simple toroid geometry and magnetic field map.

In order to determine the final track parameters, we propagate the track from the c.o.g. of the BC-segment back to the c.o.g. of the A-layer segment, assuming the track parameters from the first estimate. At each propagation step, a momentum-dependent energy loss correction is applied, and the track direction is adjusted for the bend from the magnetic field, where applicable. The local field is taken from a detailed, calculated magnetic field map [54]. To account for Coulomb scattering in the iron core toroid magnet, we allow for a scattering angle at two planes in the toroid (see Fig. 3.16). This propagation allows a fit to be performed in the bending plane, with five free parameters:

- the B-layer z-coordinate;
- the muon track momentum;
- the bending angle θ;
- two Coulomb scattering angles θ_{S1} and θ_{S2}.

The fit results in the values for these parameters, plus a (5×5) error matrix.

3.6.2 Central track matching

The ultimate goal of muon identification is to be able to tag tracks in the inner tracker as muon tracks. This is implemented by making a spatial match between a fully fit local muon track and a central track. Using the track parameters and the (5×5) error matrix from both the muon track and the central track, we can form a χ^2 and minimize it by varying the parameters according to their errors. The resolution of the central track can be improved using this technique, but only marginally, due to the large effects of Coulomb scattering in the calorimeter, especially for the low-p_T muons relevant for the analysis presented in this thesis.

It is only possible to apply the above matching algorithm, if a muon track has track segments on both sides of the toroid, because otherwise the local track fit cannot be performed and the error matrix will not be available. In practice this means the muon must have had enough momentum to penetrate the iron core toroid. In case the muon was stopped in the toroid, we use an alternative algorithm to make a match between a central track and the muon A-layer segment [55]. This alternative algorithm starts from the central track and extrapolates outward to the A-layer of the muon system. Here a match χ^2 is formed, based on the position accuracy of the A-layer segment and the accuracy of the track extrapolation. The accuracy of the extrapolation is most importantly given by the Geant-motivated accounting for scattering in the calorimeter absorber. The muons from J/ψ's, used in this thesis, typically have low p_T. We have therefore opted to use the alternative algorithm to tag muon tracks, even though this algorithm is more susceptible

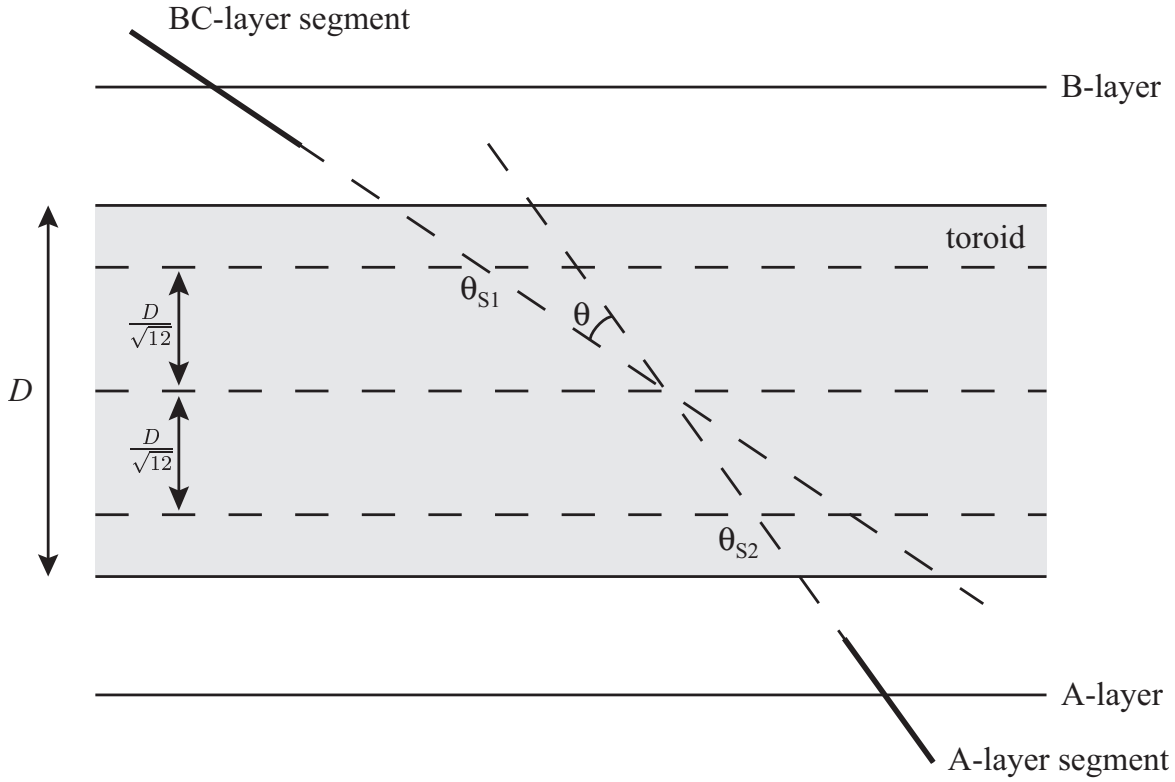


Figure 3.16: Variables relevant to the local muon track fit. The drawing shows the bending plane.

to background. In section 4.3 we will demonstrate that in our sample, even using the alternative algorithm, the background level is very low.

The efficiency of the local muon reconstruction is typically 90% for muons from J/ψ's, with a track matching efficiency of 92%, yielding a muon identification efficiency of around 83% [56]. The efficiency is lower in the bottom part of the detector, where the support structure is located (see Fig. 2.2).

Chapter 4

Selection of $B_d^0 \to J/\psi K_S^0$ decays

This chapter describes the selection of B_d^0-mesons, in the decay to a J/ψ and a K_S^0. The branching fraction for this decay mode is 4.25×10^{-4} [11]. Additionally, the J/ψ is required to decay to two muons, which happens with a branching fraction of 0.06, and the K_S^0 is required to decay to two charged pions, which happens with a branching fraction of 0.69. As a consequence, on average only one in approximately $50,000$ B_d^0's will decay in the required mode. Still, this decay channel is considered the "golden decay mode", because the event signature with two muons and two high-impact parameter tracks is relatively easily identifiable, with low background. In addition a measurement of the CP-violation in this decay is a clean measurement of $\sin(2\beta)$, due to small theoretical uncertainties.

Using the event selection procedure, a typical $B_d^0 \to J/\psi K_S^0$ decay candidate was found in event 73437230 from run 167664, shown in Fig. 4.1. This event was recorded on November 8, 2002. The view in Fig. 4.1 is along the beam line, showing the SMT barrel detectors with hits. In gray all reconstructed tracks are shown, and in black we have indicated the decay products of the B_d^0. Marked are the charged pions from the K_S^0 and the muons from the J/ψ. The primary vertex is shown as a cross. The J/ψ vertex can be identified with the B_d^0 decay vertex due to the small J/ψ lifetime. It has a separation from the primary vertex of 1.2 mm in the transverse plane. The K_S^0 has a decay length of 18.6 mm. A zoomed view of the primary vertex region is shown in Fig. 4.2, indicating a clear separation between the primary vertex and the J/ψ vertex.

To select B_d^0 candidates, first J/ψ and K_S^0 signals need to be reconstructed. We will begin with the definition of simple cuts to select clean J/ψ and K_S^0 samples, before we move on to use these samples to form B_d^0 candidates. When we have the initial sample of B_d^0 candidates, we will revisit the J/ψ and K_S^0 selection, to optimize these for the B_d^0 signal. We begin this chapter describing the selection of the initial dataset.

4.1 Data sample selection

For this analysis we have used data that was taken between October 15, 2002 and June 10, 2003. This dataset corresponds to an integrated luminosity of 114 pb^{-1} and holds approximately 276 million events. A small fraction of the runs in this sample are affected by detector problems [57].

Figure 4.1: Graphical display in the transverse plane of event 73437230 from run 167664. The SMT barrel detectors are shown with hits. All reconstructed tracks are shown in gray, and the tracks that were identified as B_d^0 decay products are shown in black. The charged pions originate from the K_S^0 in the $B_d^0 \to J/\psi K_S^0$ decays, and the muons stem from the J/ψ.

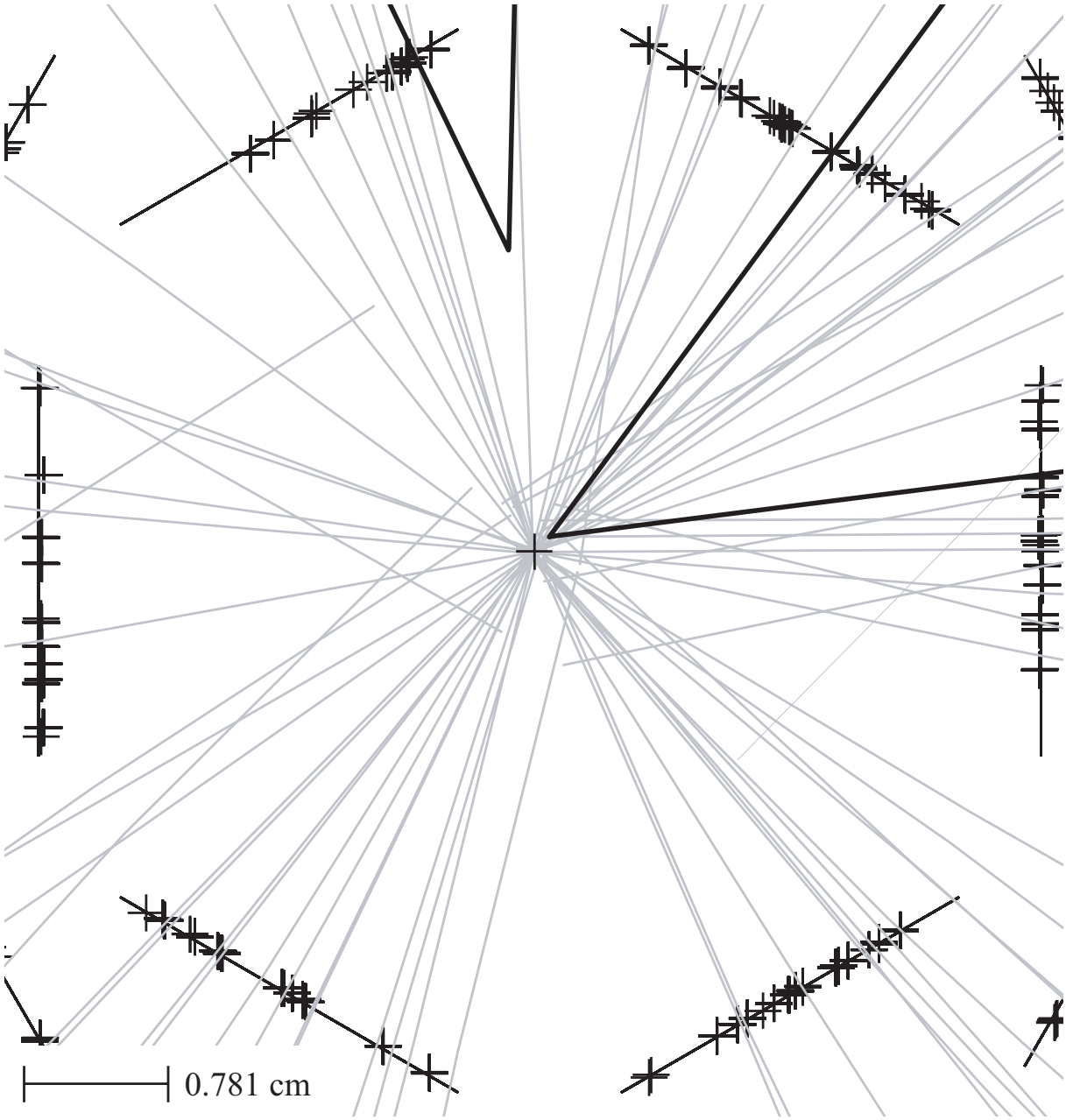


Figure 4.2: A zoomed view of the primary vertex region of the event shown in Fig. 4.1. A clear separation between the primary vertex and the dimuon vertex is visible.

These runs have been removed for this analysis.

To obtain a sample that is more easily handled, we have first selected the events passing the Level 1 dimuon trigger described in section 2.2.6. The second stage in the event selection is a selection of opposite-sign dimuons. The requirement is two central tracks of opposite charge, with a p_T greater than 1.5 GeV, that are matched to a track (-segment) in the muon system. Additionally, the invariant mass of this track pair should be in the range from 2.6 to 3.5 GeV. The dimuon sample resulting from this selection contains approximately 700,000 events.

4.2 Selection cut optimization procedure

An issue common to J/ψ, K_S^0 and B_d^0 selection is the optimization of selection cuts. The goal of optimization is to reduce the relative error on the measurement as much as possible. For the measurement of the lifetime of a B-meson we have to measure the number of B-mesons in each proper-time bin. In order to determine the number of signal particles N_s (the number of B_d^0 mesons) as accurately as possible, we have to minimize $\sigma(N_s)/N_s$, where $\sigma(N_s)$ is the error on N_s. We measure N_s and N_b by fitting a Gaussian peak plus a background function to the mass spectrum. N_s is calculated as the integral of the Gaussian peak, and N_b is the integral of the background function in a given mass window around the mean of the Gaussian. The statistical error on the total number of events, $N_s + N_b$, equals $\sqrt{N_s + N_b}$, if $N_s + N_b$ is not too small. After subtraction of the background N_b, assumed to be accurately known, also N_s has an error $\sqrt{N_s + N_b}$.

We define the *signal significance* $\mathcal{S}$ as the reciprocal of the relative error:

$$\mathcal{S} = \frac{N_s}{\sigma(N_s)} = \frac{N_s}{\sqrt{N_s + N_b}}. \tag{4.1}$$

The procedure of optimizing $\mathcal{S}$ is as follows. We select J/ψ, K_S^0 or B_d^0 candidates using an initial set of cuts. We do this a number of times for different settings of the first cut that we want to optimize. After finding the setting that yields the highest $\mathcal{S}$, we move on to the next cut. When we have done this for all cuts, it is not clear yet that we have the *overall* optimal set of cuts. The reason for this is that cuts can be correlated. If cuts A and B are correlated, the optimal value of cut A depends on the value of cut B. Varying cut B may shift the optimum of cut A. Therefore the optimization is an iterative procedure that has to be repeated until the optimum for each cut is stable.

The cuts are optimized using real data. One disadvantage of using real data is that it has low signal statistics. The optimization of $\mathcal{S}$ is therefore sensitive to statistical fluctuations. To counter this problem, one can estimate the amount of background from the data, and the amount of signal from a high-statistics Monte Carlo sample. The Monte Carlo sample would have to be scaled down to approximate the number of signal events in data, but the statistical fluctuations would be reduced. We have chosen to use real data, however, to avoid the risk that parameters that are being cut on are not properly described by the Monte Carlo.

When the same data is used for optimization of the selection cuts, as well as for the final measurement, it is inevitable that a bias is introduced in the event selection. The bias is intro-

duced in the parameter that is optimized, in our case the signal significance $\mathcal{S}$. In a cross-section measurement, one counts the number of signal events, N_s, and expect this to follow a Poisson distribution. Our selection biases N_s, because it is correlated with $\mathcal{S}$, and can therefore not be used for a cross-section measurement. A bias of $\mathcal{S}$ and correlated quantities is not problematic for a lifetime measurement, however.

4.3 J/ψ selection

In this section we will describe the selection of a clean sample of J/ψ's. J/ψ particles are identified by two opposite-sign tracks coming from a common vertex. Additionally, in the case of the decay to muons, one can require that the tracks leave traces in the muon system. This procedure is described in section 3.6. Because we have already required the presence of two opposite-sign muons in the initial selection of the dimuon sample, we require both tracks that form a J/ψ candidate to be confirmed as a muon.

Determining that both tracks come from the same vertex is done using the vertex fitting algorithm described in section 3.3, and placing a cut on the χ^2 of the vertex. Figure 4.3 (a) shows the χ^2 distributions of two-track vertices, with both tracks confirmed by the muon system. The entries in the distributions have passed all J/ψ identification cuts (see summary at the end of this section), but no cut on the vertex χ^2 has been made. We show the distribution of the entries from the "J/ψ signal mass region", which ranges from 2.9 to 3.3 GeV, as well as the distribution in the "J/ψ sidebands", for which the invariant mass of the vertex is required to lie between 2.0 and 2.7 GeV or between 3.8 and 5.0 GeV. The longer tail for the entries from the J/ψ sidebands in the χ^2 distribution in Fig. 4.3 (a) indicates that these muons sometimes do not originate from a single vertex. To select J/ψ's we are cutting on a vertex χ^2 of 20, which is a cut safe on the plateau where the signal significance is maximized (see Fig. 4.3 (b)).

To further improve the J/ψ signal significance, we will require in addition a minimum p_T for each track. Due to the thick calorimeter at DØ, significant energy is required for the muons to reach the muon system. Therefore, the lower the p_T of the track, the higher the chance that the particle that created it was stopped in the calorimeter, and that the muon system track was actually created by a different particle. Figure 4.4 (a) shows the distribution of the p_T of the lowest-p_T track on the J/ψ. Figure 4.4 (b) shows that the significance is maximized at a value close to 1.5 GeV. We cannot lower our cut beyond 1.5 GeV, because this is the cut-off in the pre-selection described in section 4.1. The pre-selection was devised to select J/ψ's from B-decays, which on average produce higher-p_T muons than general J/ψ's, as we will see later.

Finally, a cut on the p_T of the J/ψ itself is made. The p_T of J/ψ's from the J/ψ signal and sideband regions is shown in Fig. 4.5 (a). In Fig. 4.5 (b), the signal significance is shown. We employed a cut on the p_T of the J/ψ of 2 GeV.

Using the cuts described above, a clean sample of J/ψ's can be selected. However, in order for the J/ψ's to be useful for a lifetime measurement, a reasonable resolution on the J/ψ vertex is required. Therefore, even though it slightly decreases the eventual B_d^0 signal significance, we require the tracks of the J/ψ's to have at least three hits in the SMT.

In summary, in the selection of J/ψ's, we require:

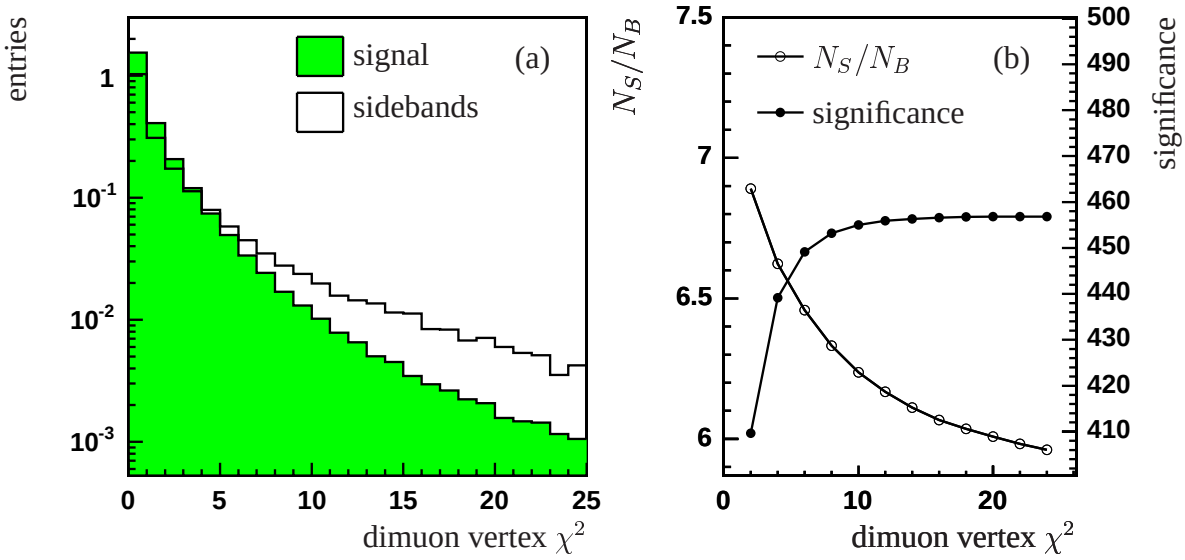


Figure 4.3: (a) The vertex χ^2 distribution of J/ψ signal and background, and (b) N_S/N_B for the J/ψ's, as well as the significance.

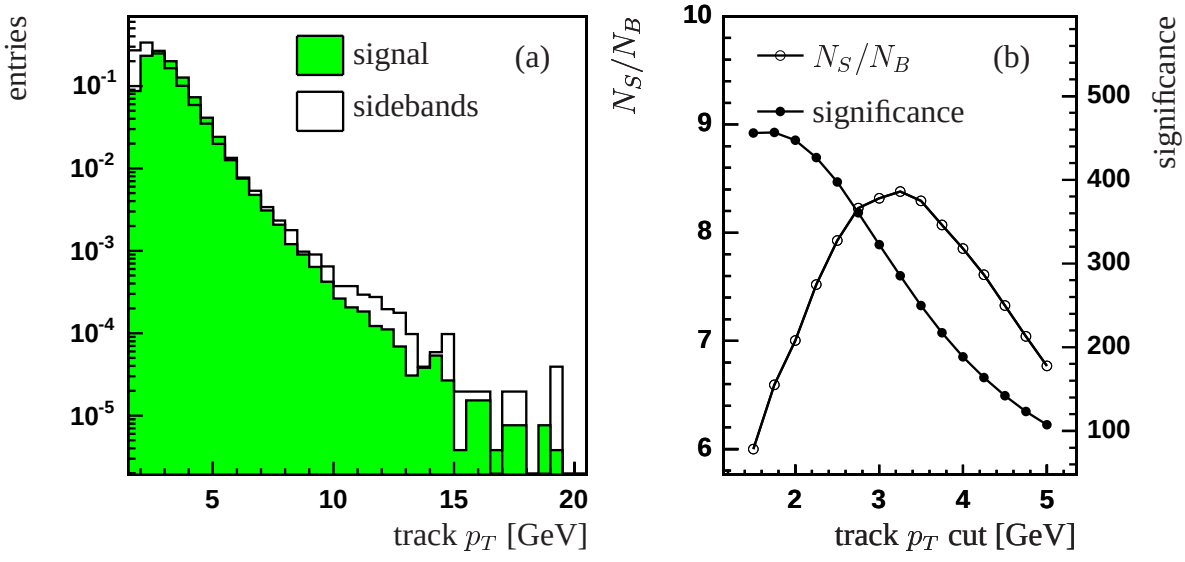


Figure 4.4: (a) The p_T distribution of the lowest-p_T track of J/ψ signal and background, and (b) N_S/N_B for the B_d^0's, as well as the significance.

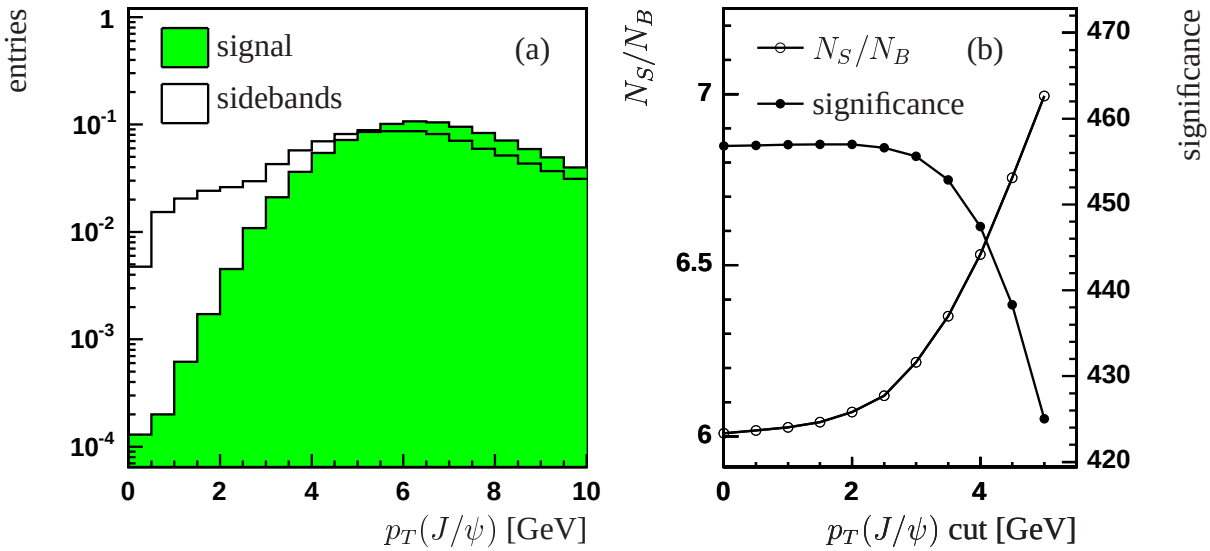


Figure 4.5: (a) The p_T distributions of the J/ψ candidates from the J/ψ signal and sideband regions, and (b) N_S/N_B and the J/ψ signal significance for the cut on the p_T of the J/ψ.

- confirmation by the muon system of both tracks;
- track $p_T > 1.5$ GeV;
- χ^2 of the two-track vertex < 20;
- each track has at least three SMT hits;
- p_T of the $J/\psi > 2$ GeV.

The mass spectrum resulting from this selection is shown in Fig. 4.6 (a). A straight line was used to fit the background, and we have fit two Gaussians to the signal peak, centered at the same mass, representing the fact that tracks with different p_T resolution contribute to the peak. The fit yields a J/ψ mass of 3070.8 $\pm$ 0.2 MeV. The first Gaussian has a width of 58.6 ± 0.6 MeV and contains 172,000 J/ψ's, the second has a width of 119 ± 3 MeV and contains 88,000 J/ψ's. The measured J/ψ mass of 3070.8 $\pm$ 0.2 MeV is below the PDG value of 3097 MeV [11]. It is hypothesized that this is due to incorrect assessment of energy loss in the tracker, an inaccurate magnetic field in the reconstruction or the radial alignment of the tracker. The assessment of the energy lost by particles in the tracker may be off, if not all of the material that actually exists in the tracking volume has been accounted for. Ref. [58] shows that the magnetic field used in the reconstruction is approximately 0.24% too low, which accounts for a correction of the order of 10 MeV. We discuss the possible effect of the deviation of the mass on our measurement of the B_d^0 lifetime in section 5.6.5.

We have also applied our J/ψ selection, as outlined in this section, to a $B \to J/\psi X$ MC sample. The technical details of this sample are described in section A.2. The resulting dimuon

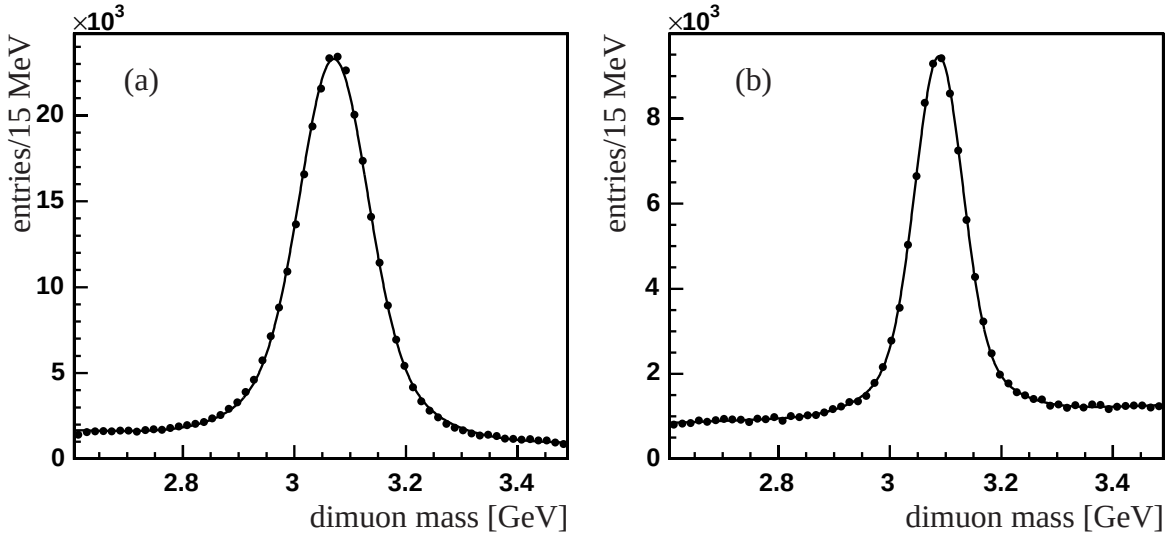


Figure 4.6: Mass spectrum of two-track vertices passing J/ψ identification cuts, using (a) our data sample and (b) a $B \to J/\psi X$ Monte Carlo sample. The fit to the real data yields a total of about 260,000 J/ψ's at a mass of 3070.8 ± 0.2 MeV.

mass spectrum is shown in Fig. 4.6 (b). The fit shape is again a double Gaussian over a linear background. We fit a J/ψ mass of 3088.0 ± 0.2 MeV. Even in the Monte Carlo sample, the mass is fit below the PDG mass of 3097 MeV (which corresponds to the J/ψ mass in our Monte Carlo program). This can happen if the peak is not well described by a double Gaussian. The widths of the Gaussians are 41.4 ± 0.2 MeV and 93 ± 1 MeV, with 84% of the events in the first Gaussian. This is somewhat more narrow than the width observed in data, and can be attributed to inaccuracies in the description of the magnetic field and the material in the detector.

4.4 K_S^0 selection

We will now make a selection of K_S^0 from the data, that we optimize for the significance of the number of K_S^0's. We will later tune our selection to be optimal for the selection of B_d^0's. We will use the sample defined in section 4.1.

The signature of a K_S^0 as seen in the detector, is a vertex of two tracks, displaced from the primary vertex, forming the K^0 mass of 498 MeV [11]. The vertex fitting takes relatively long, so to limit processing time, a track selection is made first. We will perform the vertex fit only on track pairs of opposite charge, where both tracks have a good p_T measurement ($\sigma(1/p_T)\ /\ (1/p_T) < 0.5$).

Since K_S^0's have relatively long lifetime and can decay at large radius between detector layers, we should allow for "misses" on tracks before the K_S^0 decay vertex, and we expect hits on all detector elements beyond the decay vertex. Therefore we require the number of missing hits on

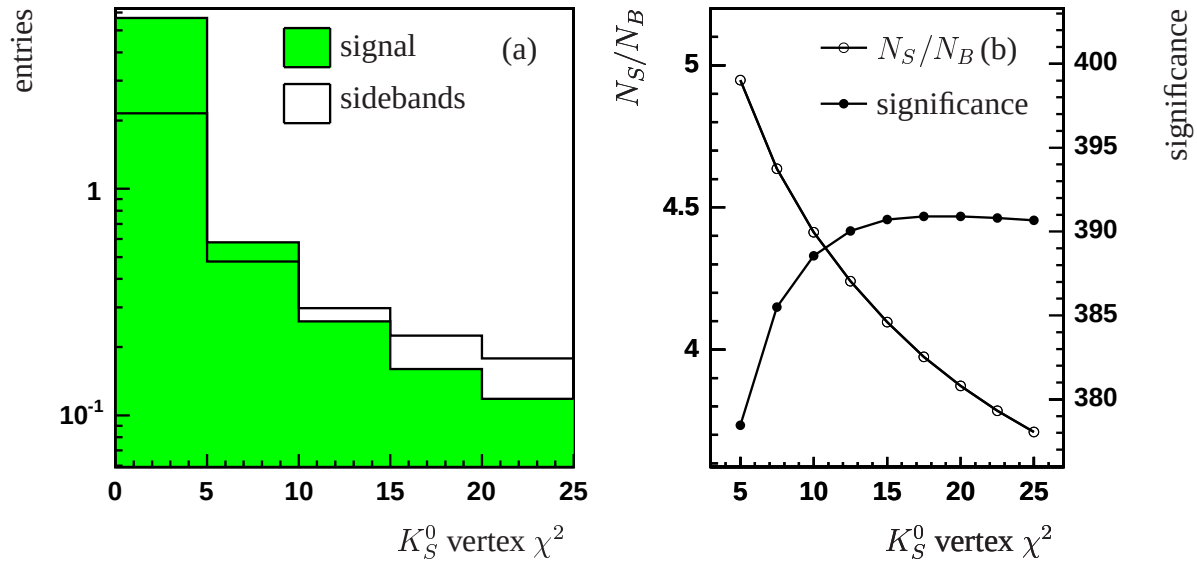


Figure 4.7: (a) The vertex χ^2 distribution of the entries in the K^0_S signal mass window and sidebands (background). (b) N_S/N_B and the signal significance as a function of the vertex χ^2 cut value.

each track beyond the vertex to be smaller than 2, and the sum of the number of hits on both tracks within the vertex radius to be smaller than 3.

Because of the long K^0_S lifetime, we can require both tracks to be displaced from the primary vertex: the χ^2 of each track with respect to the selected primary vertex, χ^2_{PV}, should be larger than 9 and at least one of the tracks should have $\chi^2_{PV} > 16$.

After the tracks have been fit into a vertex, selection criteria on the vertex quality are applied. We begin by requiring the vertex χ^2 to be smaller than 20. The vertex χ^2 distribution for K^0_S's in the K^0_S signal and sideband regions are shown in Fig. 4.7 (a). We defined the K^0_S signal region as the mass region from 470 to 520 MeV and the sidebands as the regions from 400 to 460 MeV and 530 to 600 MeV. Figure 4.7 (b) shows the signal significance as a function of the cut on the vertex χ^2, justifying the cut that has been chosen.

To further improve the signal significance, we require a minimum p_T for the reconstructed K^0_S. The p_T distributions and signal significance as a function of the cut on the p_T of the K^0_S, are shown in Fig. 4.8. A cut on the K^0_S p_T of 0.15 GeV has been chosen.

Finally, the K^0_S can be required to approximately point back to its production vertex. Most of the K^0_S's in the data have been produced at the primary vertex; only a small fraction of K^0_S's come from displaced B-meson decay vertices. A large reduction of the background in our K^0_S selection can therefore be achieved by requiring the K^0_S to point back to the primary vertex. We will later modify this requirement when developing an efficient selection for B^0_d decays. The requirement is captured in a cut on the vertex collinearity. The vertex collinearity is the cosine of the angle φ between the reconstructed K^0_S momentum and the line between the primary vertex and the reconstructed K^0_S decay vertex, all in the transverse plane (see Fig. 4.9). The vertex

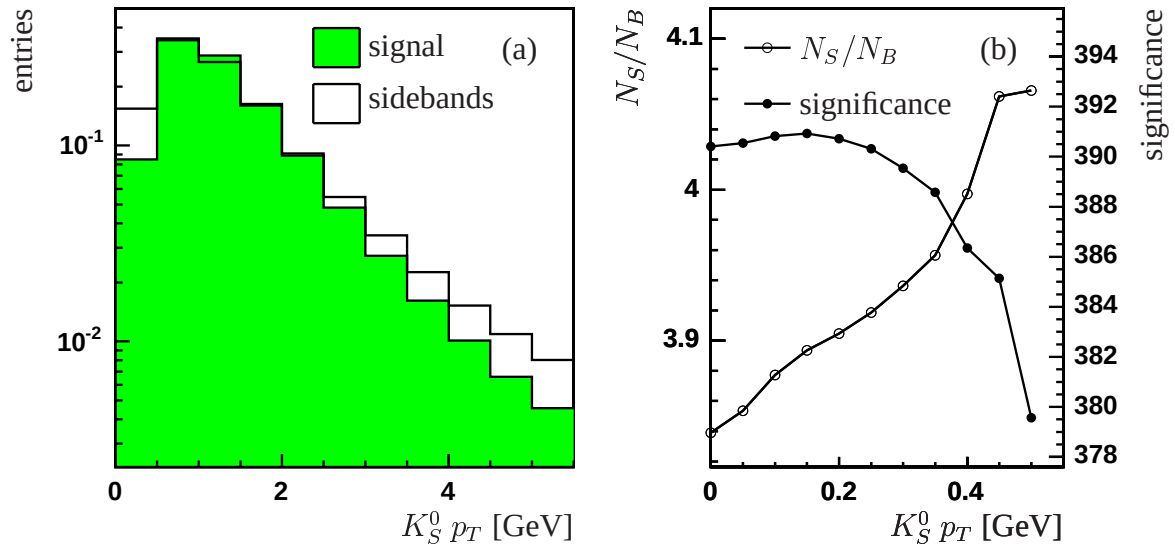


Figure 4.8: (a) The p_T distribution of K_S^0 candidates corresponding to entries in the K_S^0 signal mass window and sidebands (background). (b) N_S/N_B and the signal significance as a function of the cut value.

collinearity distributions and signal significance as a function of cut value have been plotted in Fig. 4.10. A cut on the vertex collinearity of 0.999 is used, corresponding to a maximum angle of around 2.5˚.

In summary, the selection criteria are:

- track selection: both tracks required to have $\sigma(1/p_T)/(1/p_T) < 0.5$ and $\chi^2_{PV} > 9$, and at least one of the tracks $\chi^2_{PV} > 16$, the number of misses on each track beyond vertex < 2, the total number of hits before vertex < 3;
- vertex selection: $\chi^2 < 20$, collinearity > 0.999;
- p_T of the $K_S^0 > 0.15$ GeV.

The mass spectrum of the vertices passing the above criteria is shown in Fig. 4.11 (a). The fit yields around 216,000 K_S^0's at a mass of 492.68 ± 0.03 MeV. This means that the measured K_S^0 mass is, as the J/ψ mass, lower than the PDG value of 498 MeV. This can have an effect on our B_d^0 lifetime measurement, which will be discussed in section 5.6.5. The mass spectrum is fit to the sum of two Gaussians with the same mass, but different widths, on top of a linear background. Two different Gaussians are used, because tracks with different p_T resolutions contribute to the mass spectrum. The p_T resolution is better for more central tracks, due to, for example, a more homogeneous magnetic field, and more CFT hits. For tracks pointing back to $z = 0$, full CFT coverage is only available out to a pseudo-rapidity of approximately 1.6. This is illustrated by the plots in Fig. 4.12. The widths of the Gaussians in the fit to the mass spectrum are 7.25 ± 0.06

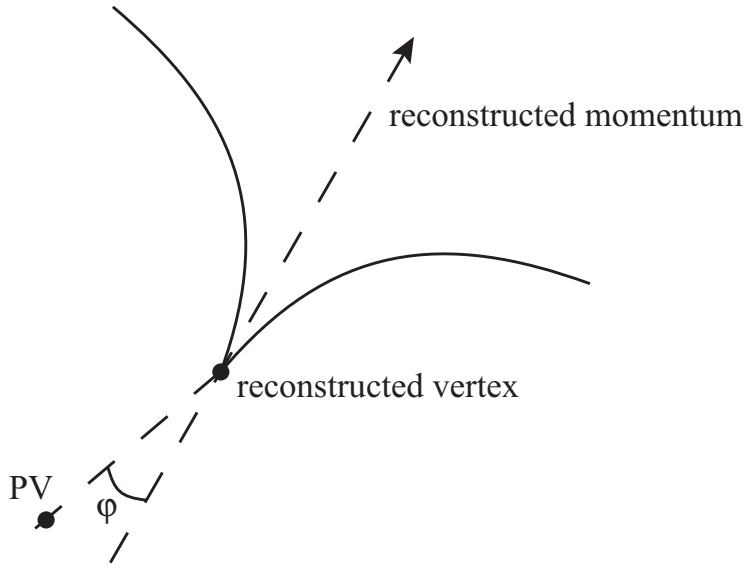


Figure 4.9: Vertex collinearity. The collinearity is the cosine of the angle φ between the reconstructed momentum and the line from the primary vertex (PV) to the reconstructed decay vertex.

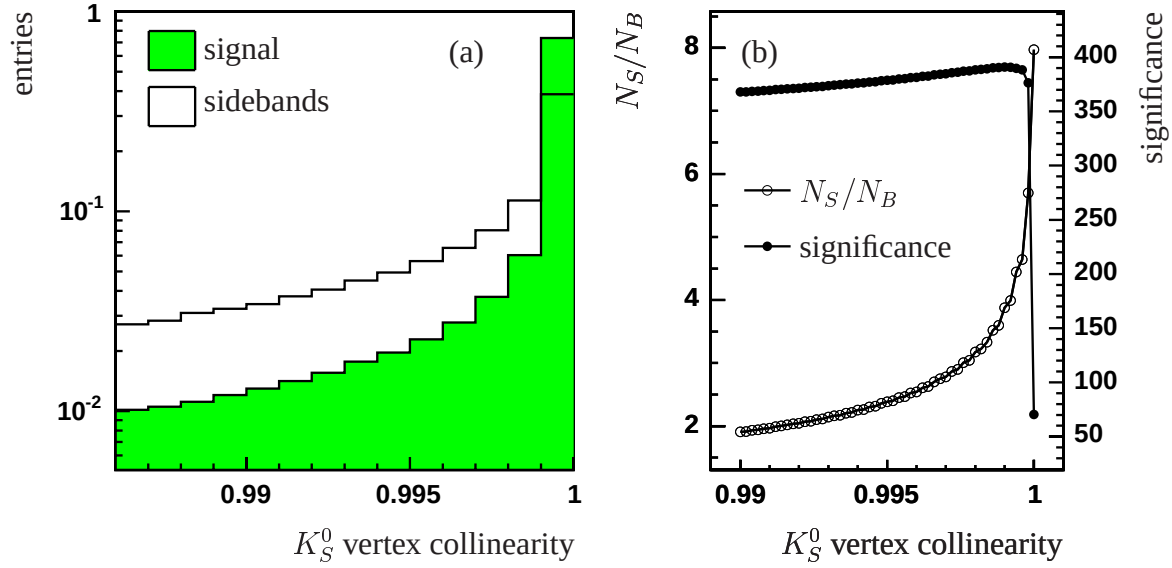


Figure 4.10: (a) The collinearity distribution of K^0_S candidates in the K^0_S signal mass window and sidebands (background). (b) N_S/N_B and the signal significance as a function of collinearity cut value.

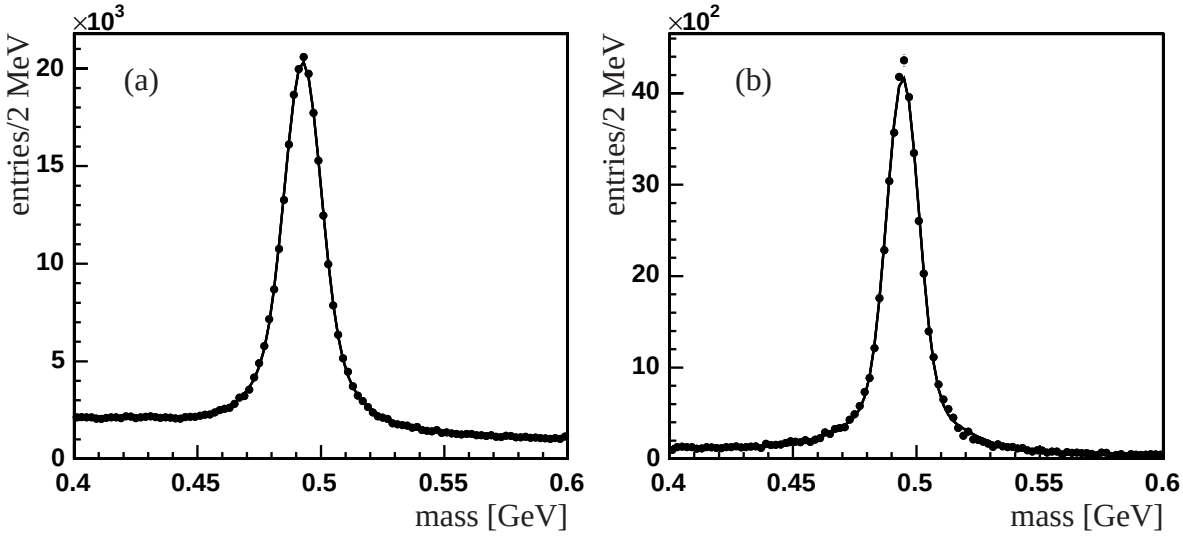


Figure 4.11: The two-track mass spectrum after K_S^0 selection cuts, using (a) our data sample and (b) a $B \to J/\psi X$ Monte Carlo sample. The fit to the data yields 216,000 K_S^0's at a mass of 492.68 ± 0.03 MeV.

MeV and 17.7 ± 0.3 MeV, where the first number may be interpreted as resulting directly from the p_T resolution of the tracker for low p_T tracks. 81% of the events are described by the narrow Gaussian.

Applying the K_S^0 selection to our $B \to J/\psi X$ Monte Carlo sample yields the mass distribution of Fig. 4.11 (b). The fit shape is identical to the shape used in Fig. 4.11 (a), a double Gaussian over a linear background. The fitted mass is 494.5 ± 0.05 MeV. Again we see a discrepancy with the K_S^0 mass according to the PDG. The widths of the Gaussians are 6.4 ± 0.07 MeV and 20.0 ± 0.5 MeV, with 87% of the events following the narrow Gaussian. This is consistent with our observations with respect to the J/ψ signal.

4.5 B_d^0 selection

To find B_d^0 candidates, we combine the J/ψ and K_S^0 candidates in each event and apply additional selection cuts. In section 4.5.1, we will begin our B_d^0 selection using the clean J/ψ and K_S^0 samples defined in the previous two sections. The J/ψ and K_S^0 selection cuts defined previously have been optimized without taking into account their efficiency for the B_d^0 signal or background. Therefore, as described in section 4.5.2, we will re-tune them to enhance the B_d^0 signal. The K_S^0 collinearity was defined with respect to the primary vertex. We will change this as described in section 4.5.2. After this, in section 4.5.3, we will form a set of selection cuts that is optimal for the B_d^0 lifetime measurement.

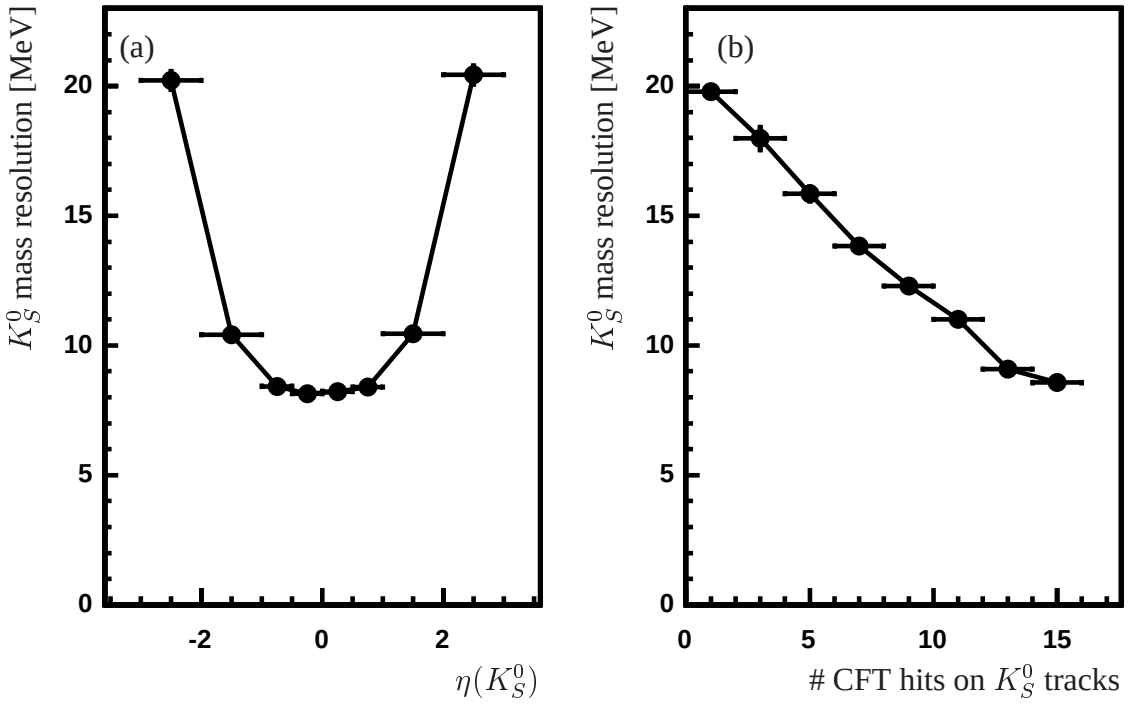


Figure 4.12: The standard deviation of a Gaussian fit to the K_S^0 mass spectrum in data, as a function of η (a), and as a function of the number of CFT hits on each track (b). At higher η the field is less well-known, and the average number of CFT hits on the track decreases, due to the CFT acceptance. Tracks with fewer CFT hits have poorer p_T resolution, resulting in a poorer K_S^0 mass resolution.

4.5.1 Selection using predefined J/ψ and K_S^0 cuts

To produce the initial set of B_d^0 candidates, we vertex the two muons from the J/ψ's, and the K_S^0 that we have found in the previous sections into a three-track vertex. The vertexing procedure is identical to the vertex fitting algorithm described in section 3.3, only the track index a in equation 3.12 runs over three tracks. The inputs to the B_d^0 mass calculation are the J/ψ and K_S^0 four-momentum vectors. To first order these are obtained by summing the four-momentum vectors of the decay products, the muons or pions respectively. To improve the B_d^0 mass and vertex resolution, a mass-constrained fit is applied to the J/ψ and K_S^0 (see section 3.3.1).

Before forming the candidate B_d^0 vertices, we apply a mass window cut on the J/ψ and K_S^0 candidates. To define the mass window, we will use the same strategy as for other cuts: We find the mass windows that maximize the B_d^0 signal significance. To determine, for instance, the lower boundary of the J/ψ mass window, we fix the upper boundary of the J/ψ mass window and measure the B_d^0 signal significance as a function of the lower bound of the J/ψ mass window. The result for the J/ψ lower mass window boundary is shown in Fig. 4.13 (a). The results for the upper J/ψ mass window bound and the K_S^0 mass window cuts are shown in Fig. 4.13 (b), (c) and (d) respectively. Based on these results, we set a J/ψ mass window from 2.9 to 3.5 GeV and a K_S^0 mass window from 0.44 to 0.53 GeV.

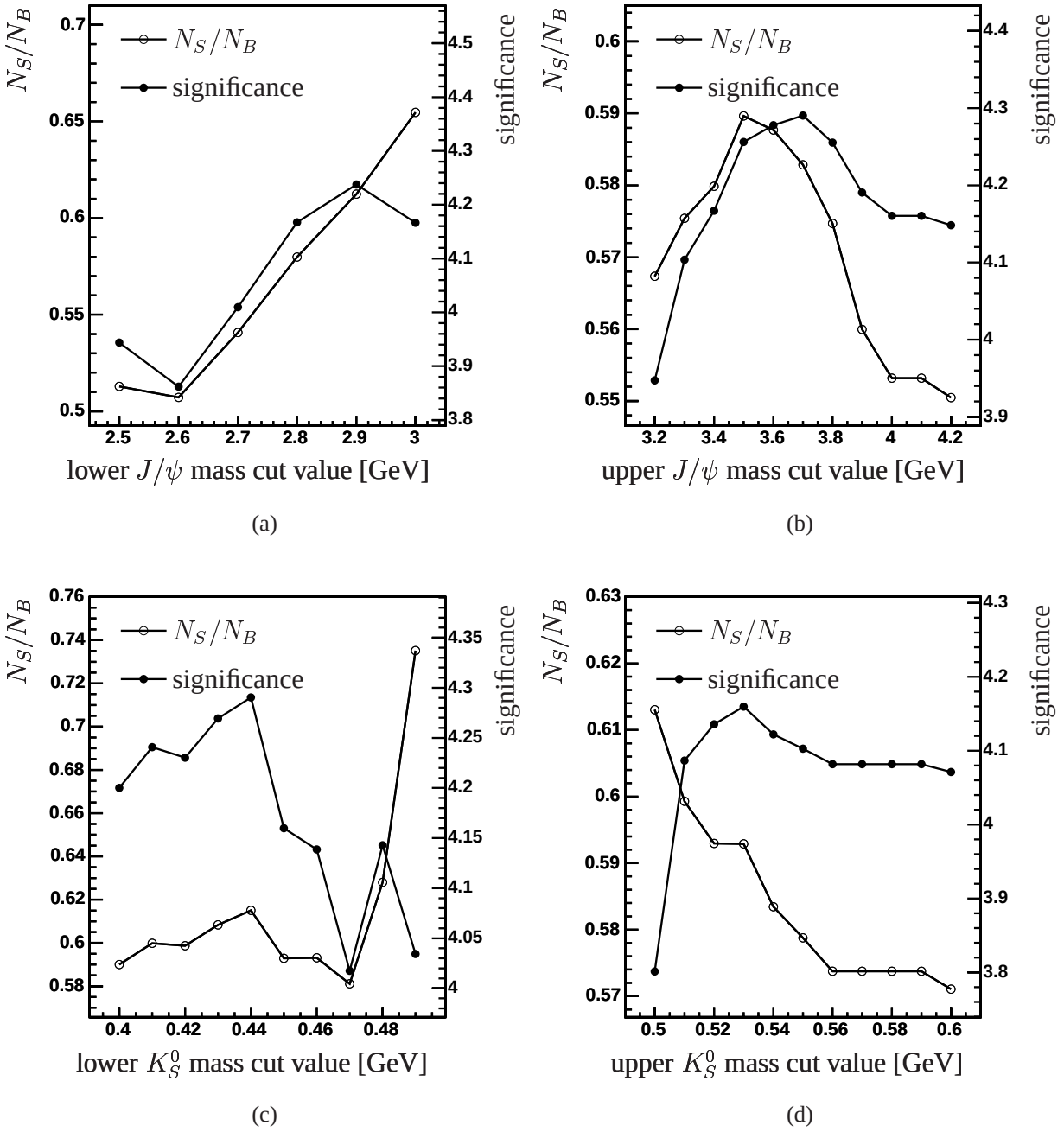
To select the B_d^0 mesons we begin by placing a vertex χ^2 cut on the $(J/\psi K_S^0)$ vertex. We define the B_d^0 signal mass window to be from 5.16 to 5.4 GeV, and the B_d^0 sidebands from 4.1 GeV to 5.1 GeV and from 5.6 GeV to 6.6 GeV. The vertex χ^2 distributions of signal and sidebands in Fig. 4.14 (a) show that the signal distribution has a similar shape as the background. The signal significance plot in Fig. 4.14 (b) indicates that we should primarily be careful not to cut too low. We have chosen to cut at a vertex χ^2 of 25. To extract the B_d^0 signal we additionally employ a series of other cuts:

- χ^2 of the B_d^0 "track" w.r.t. the primary vertex < 30;
- collinearity of the $B_d^0 > 0.95$;
- p_T $(B_d^0) > 8$ GeV.

Distributions of the above mentioned parameters, as well as the signal significance as a result of each cut, are shown in Figs. 4.15, 4.16 and 4.17.

The $(J/\psi K_S^0)$ mass spectrum resulting from this selection is shown in Fig. 4.18. The fit function is a Gaussian on top of a linear background. The fit finds the B_d^0 mass at 5.29 ± 0.01 GeV, which matches the PDG value. Since the mass of the J/ψ and K_S^0 have been constrained to their PDG value, this shows that there is no bias in the B_d^0 selection, and neither is a bias introduced due to the vertex fitting algorithm. The width of the Gaussian is 42 ± 13 MeV, and it contains around 48 ± 5 B_d^0's. The number of background events in the B_d^0 signal mass range from 5.16 to 5.4 GeV in this plot is around 67.

The $(J/\psi K_S^0)$ mass spectrum is composed of the B_d^0 signal and different background contributions. One background contribution consists of combinations of J/ψ's and K_S^0's of which either the J/ψ or the K_S^0, or both, is fake. Fake, in the case of a J/ψ, means that the J/ψ candidate passes the J/ψ selection criteria, but the tracks identified as J/ψ decay products didn't



(a) (b)

(c) (d)

Figure 4.13: Determining mass window cuts for J/ψ and K_S^0 candidates.

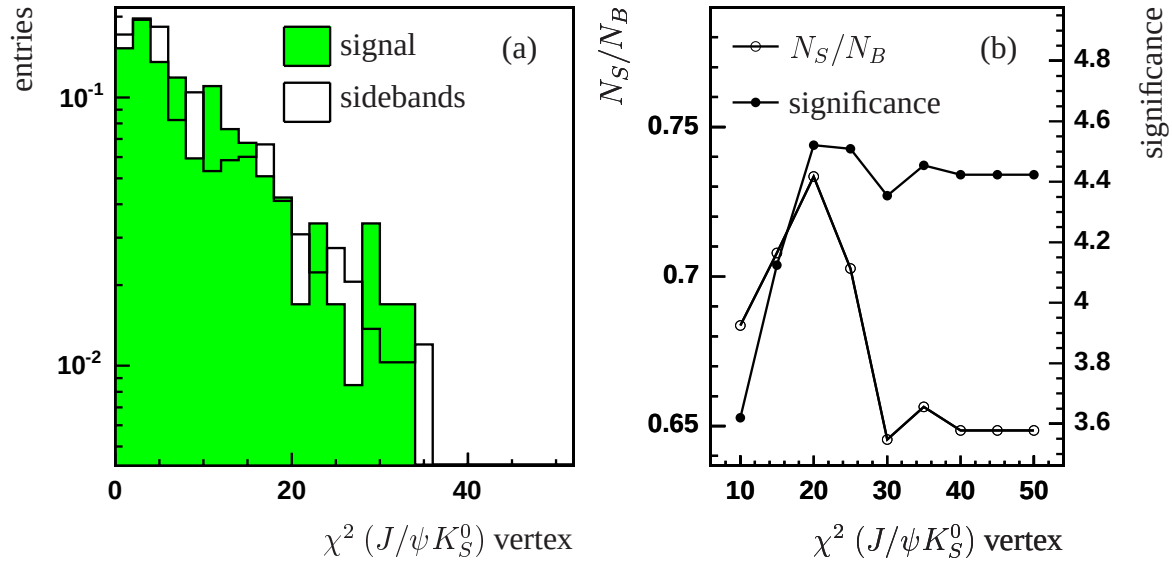


Figure 4.14: (a) The vertex χ^2 distribution of the B_d^0 candidates, in the B_d^0 signal mass window and in the sidebands (background). (b) N_S/N_B and the signal significance as a function of the cut on the vertex χ^2.

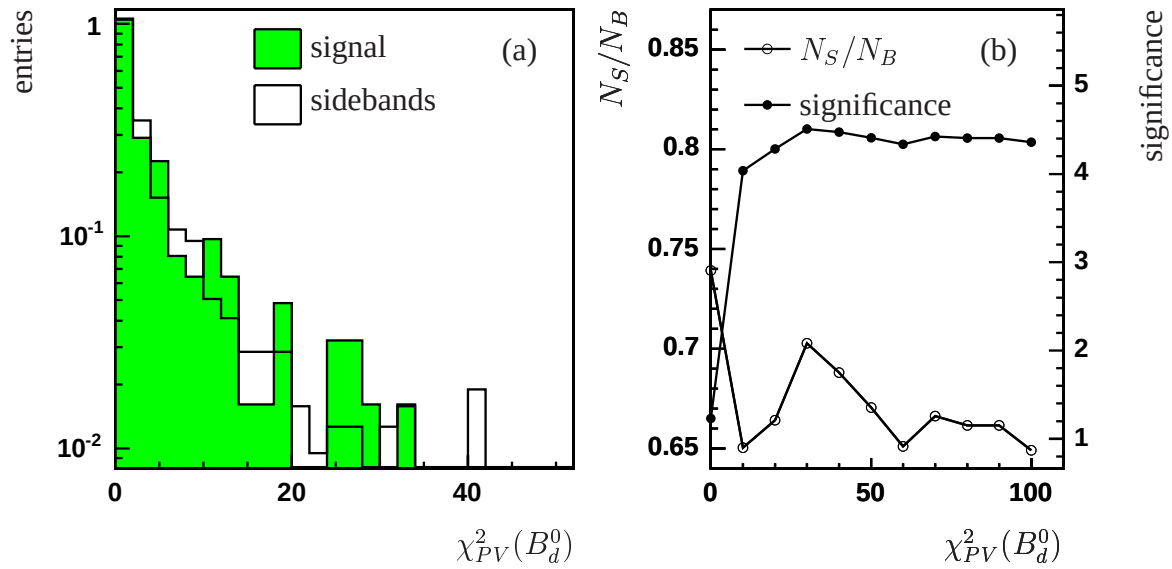


Figure 4.15: (a) The χ^2 distribution of the B_d^0 candidates w.r.t. the primary vertex, in the B_d^0 signal mass window and in the sidebands (background). (b) N_S/N_B and the signal significance for the cut on the χ^2_{PV}.

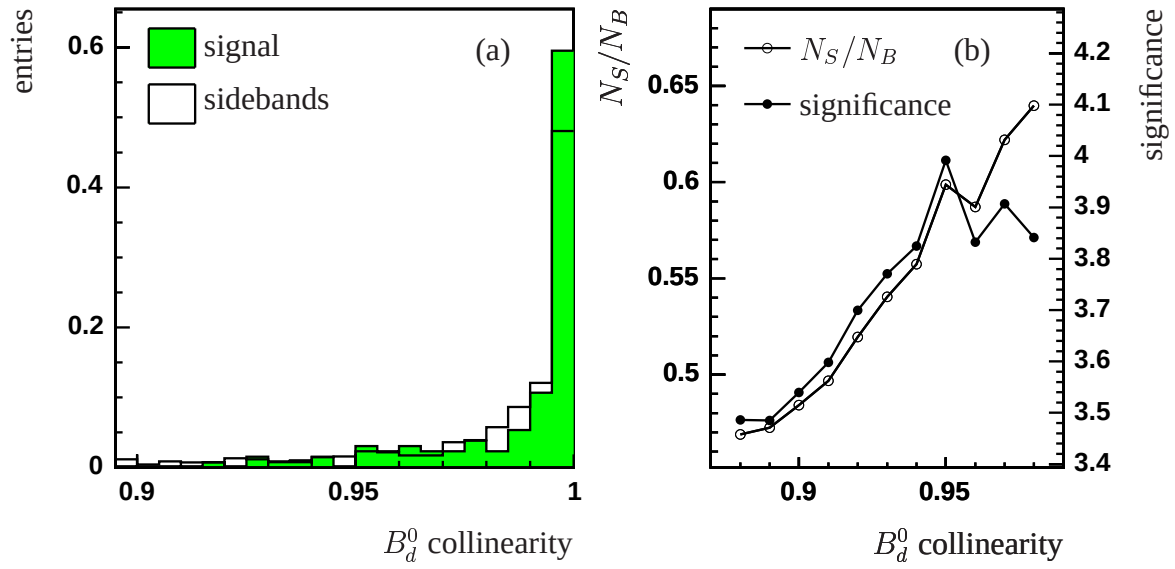


Figure 4.16: (a) The collinearity distribution of the B_d^0 candidates, in the B_d^0 signal mass window and in the sidebands (background). (b) N_S/N_B and the signal significance for the cut on the collinearity.

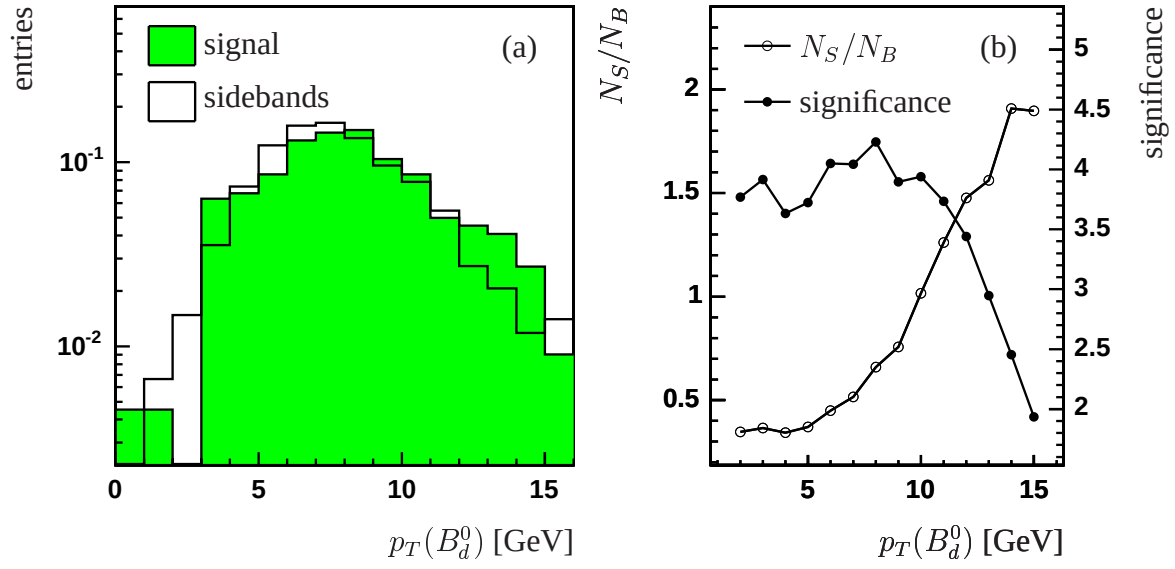


Figure 4.17: (a) The p_T spectrum of the B_d^0 candidates, in the B_d^0 signal mass window and in the sidebands (background). (b) N_S/N_B and the signal significance for the cut on the p_T of the B_d^0 candidates.

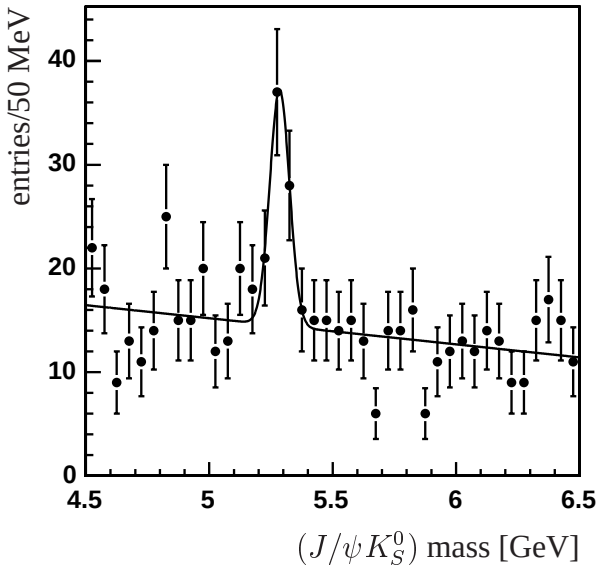


Figure 4.18: The B_d^0 mass spectrum, using the pre-defined J/ψ and K_S^0 selection cuts. The fit yields 48 ± 5 B_d^0's at a mass of 5.29 ± 0.01 GeV, over a background of 67 events.

actually originate from a J/ψ. The other type of background consists of combinations of true J/ψ's and true K_S^0's, that were not produced by the decay of a B_d^0 meson. As can be seen in Figs. 4.6 (a) and 4.11 (a), respectively, the level of the background in the dimuon mass spectrum showing the J/ψ peak, as well as the background in the two-track mass spectrum with the K_S^0 peak, is very low, compared to the level of background under the B_d^0 mass peak. The background under the B_d^0 mass peak does not consist primarily of fake J/ψ's and/or K_S^0's; it primarily consists of true J/ψ's and true K_S^0's that did not originate from a $B_d^0 \to J/\psi K_S^0$ decay.

4.5.2 Selection with optimal cuts for B_d^0 significance

In the previous sections, we have defined a B_d^0 selection, that was tuned in three steps. First, the J/ψ selection was tuned to optimize the J/ψ significance. Secondly, the K_S^0 selection was optimized for the K_S^0 significance. And finally the selection on the B_d^0 cuts was optimized for the significance of the number of B_d^0's. Taking these selection cuts as a starting point, we can go back to the J/ψ and K_S^0 selections, and tune them for high B_d^0 significance. We expect some cuts to change, because for instance J/ψ's from B's have higher p_T, on average, than J/ψ's from prompt production processes. Also, if we use a J/ψ and K_S^0 selection that is better suited for J/ψ's and K_S^0's from B's, maybe the B_d^0 selection can be relaxed.

We tune all cuts described in the previous sections for optimal B_d^0 significance. Most cuts remain the same, so we will only discuss the cuts that will be modified. In the K_S^0 selection, the collinearity was defined as the cosine of the angle between the K_S^0 momentum and the line connecting the K_S^0 decay vertex and the primary vertex, because most K_S^0's are produced at the primary vertex. Since we now tune our selection to be efficient for K_S^0's from B-decays, we replace the primary vertex in the definition of the K_S^0 collinearity with the $(J/\psi K_S^0)$ vertex. The

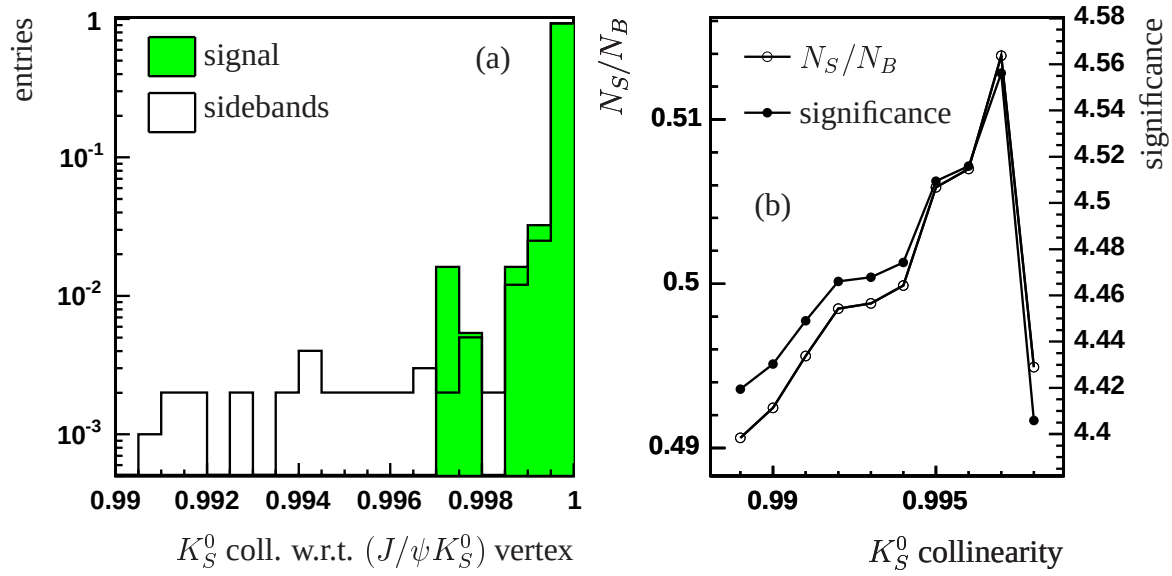


Figure 4.19: (a) The K_S^0 collinearity with respect to the J/ψ vertex, and (b) the B_d^0 significance as a function of the cut on the K_S^0 collinearity.

plot of the significance is shown in Fig. 4.19 (b). We place the cut at 0.996, which corresponds to approximately 5˚. As an additional track quality cut, we can require a minimum number of CFT hits on each track. The distribution of the number of CFT hits for K_S^0's in the B_d^0 signal and sideband regions ranges from 0 to 16, and is shown in Fig. 4.20 (a). We decided to require only a minimum of 1 CFT hit, because the statistical fluctuations in the significance graph in Fig. 4.20 (b) are relatively large, and the gain of raising the cut appears small.

In the selection of J/ψ's, as expected, the B_d^0 significance was improved if we raise the p_T cut on the muons to 2.5 GeV and the cut on the $p_T(J/\psi)$ to 3 GeV. The plots are shown in Fig. 4.21 and 4.22. Also, we found a slight improvement after widening the J/ψ mass window to 2.8 – 3.6 GeV. Finally, when we search for the optimal value of the B_d^0 collinearity cut, we find that it has shifted to a slightly lower value, as shown in Fig. 4.23. We placed our cut at 0.92, or 23˚.

The mass spectrum resulting from this selection is shown in Fig. 4.24. The number of B_d^0's in the peak is 67 ± 7, over a background of 114 events in a mass window from 5.16 to 5.4 GeV. This corresponds to a signal to noise ratio of 0.59, which is poorer than the selection described in section 4.5.1, yielding 0.72. However, the signal significance improved from 4.48 to 4.98.

We can compare our sample of 67 reconstructed B_d^0's to the event yield as predicted by Ref. [7]. Scaling for the difference in integrated luminosity, the main factor for the difference is the trigger efficiency. The prediction uses the assumption that a muon A-stub trigger would be available. Triggering on A-stubs in stead of scintillator coincidences lowers the effective p_T threshold from 3 GeV to 1.5 GeV, because muons are not required to penetrate the toroid. The muon A-stub trigger has only become available after our data-taking period. Therefore we have used a muon scintillator trigger. The lower p_T threshold was predicted to increase the event yield by a factor 10. If we remove this factor from the calculation, the observed event yield is of the

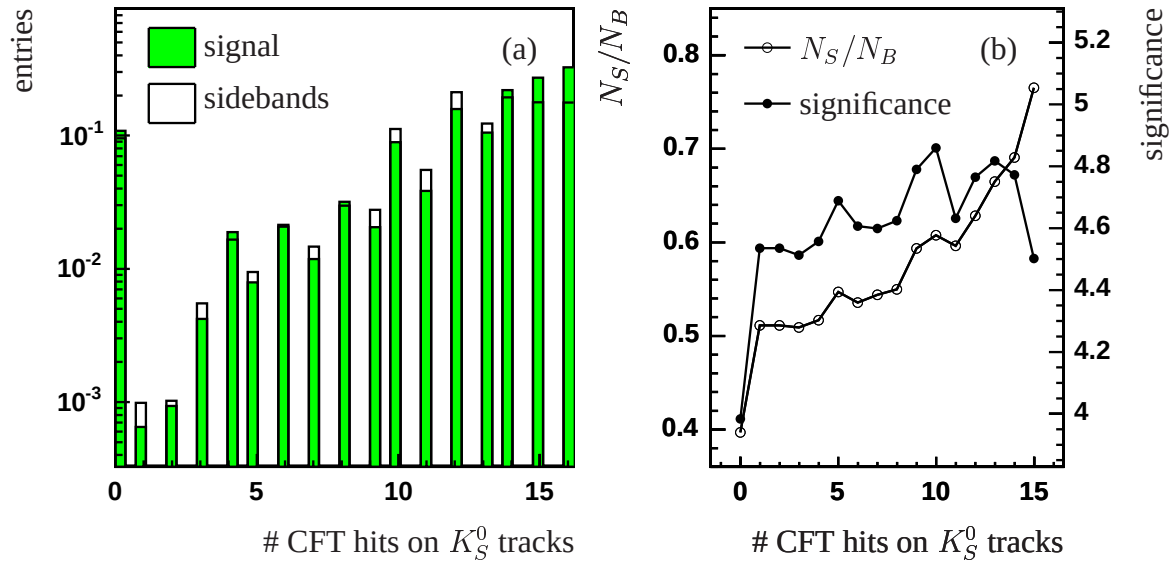


Figure 4.20: (a) The number of CFT hits on each track on the K_S^0 vertex, and (b) the B_d^0 significance as a function of the cut on the number of CFT hits.

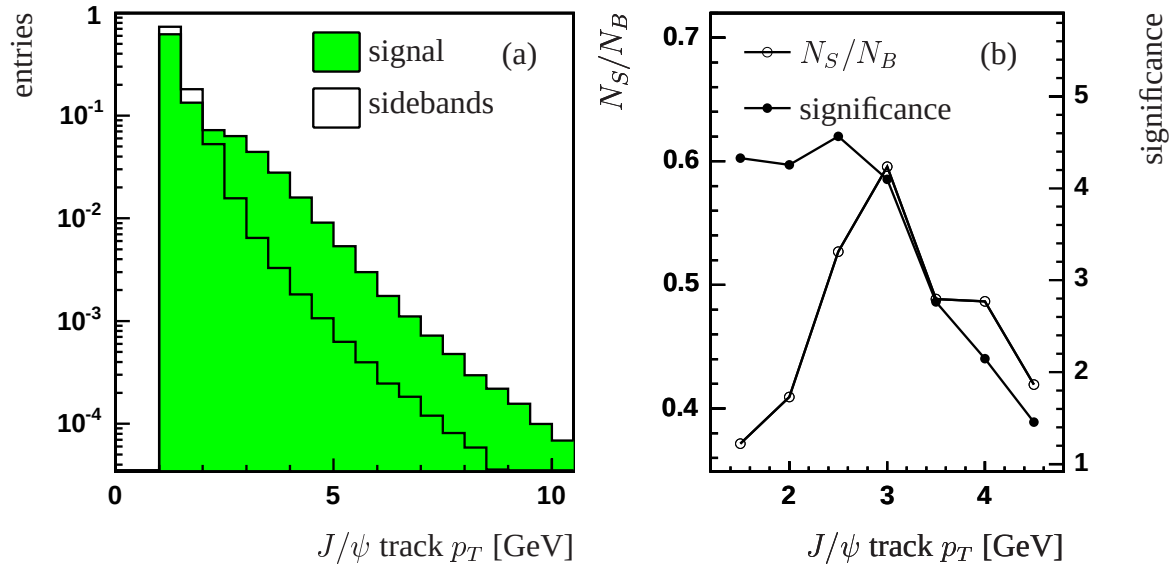


Figure 4.21: (a) The p_T distribution of the lowest-p_T muon on a J/ψ, and (b) the B_d^0 significance as a function of the cut on this p_T.

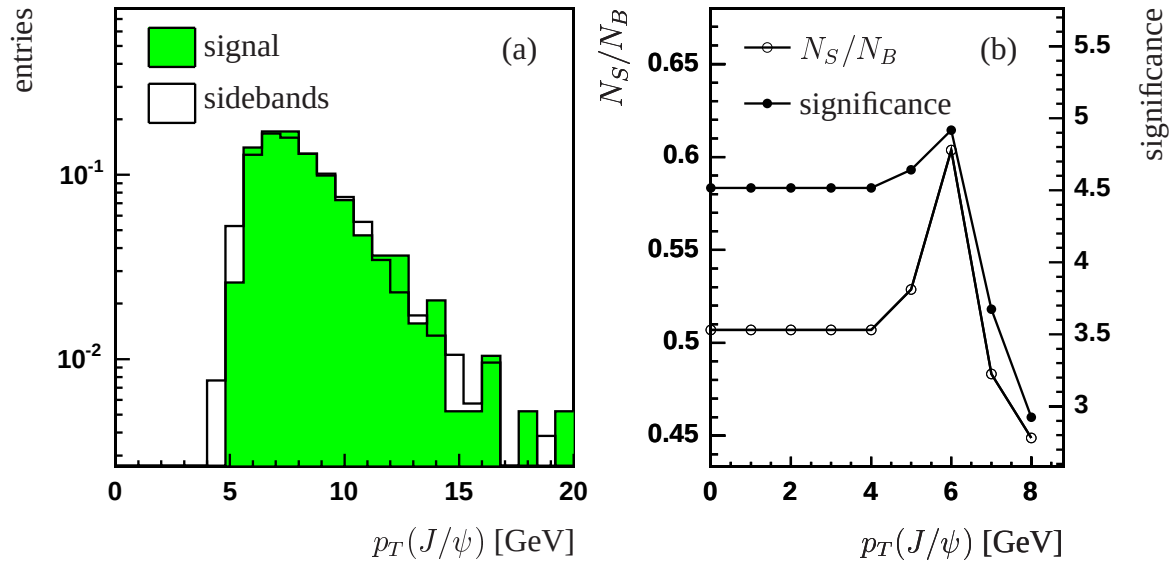


Figure 4.22: (a) The p_T distribution of the J/ψ's, and (b) the B_d^0 significance as a function of the cut on the $p_T(J/\psi)$.

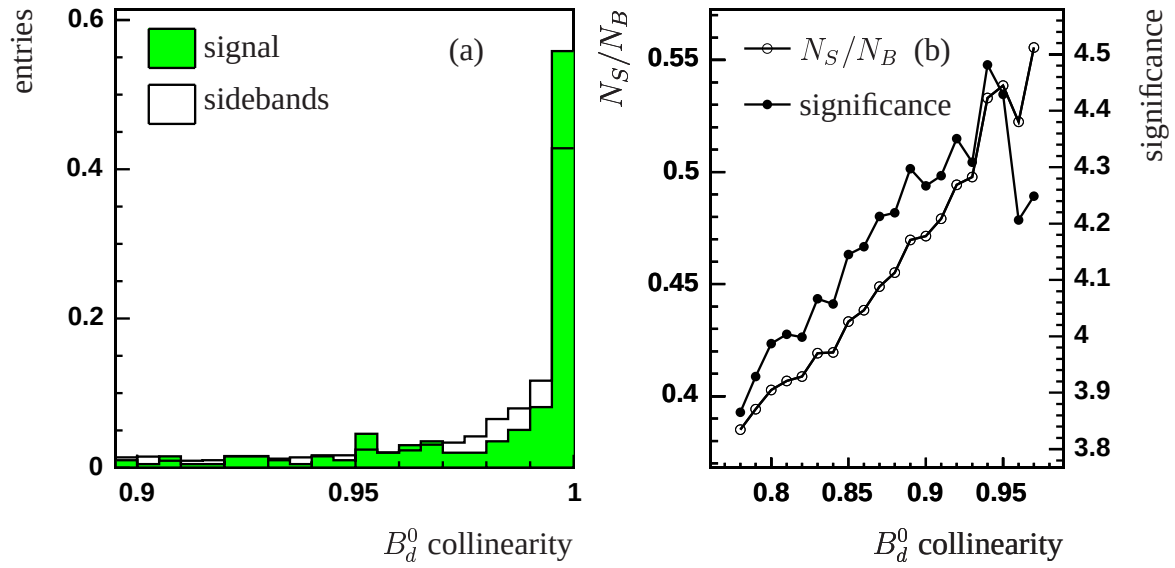


Figure 4.23: (a) The collinearity distribution of the B_d^0's, and (b) the B_d^0 significance as a function of the cut on the collinearity.

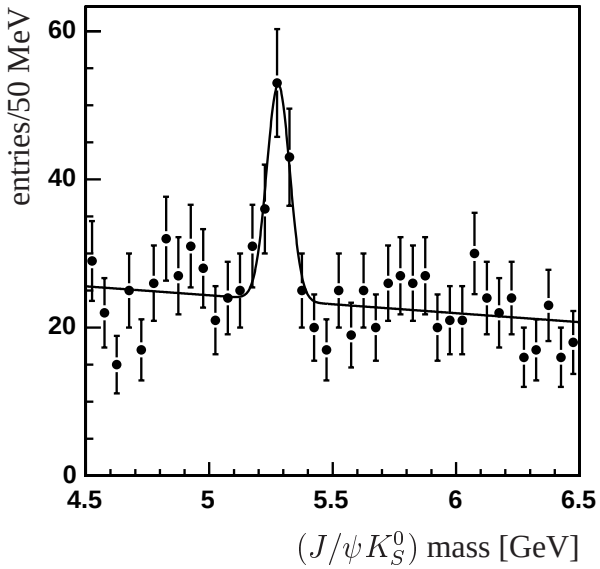


Figure 4.24: The $(J/\psi K_S^0)$ mass spectrum after tuning also the J/ψ and K_S^0 selections to optimize the B_d^0 significance. The fit yields 67 ± 7 B_d^0's at a mass of 5.29 ± 0.01 GeV, over a background of 114 events.

same order of magnitude as the prediction.

4.5.3 Selection for lifetime measurement

It is of crucial importance in the selection of B_d^0 candidates not to impose any cut that could bias the proper time distribution in any way. The B_d^0 collinearity cut, which has served us well in establishing a clean, signal enhanced sample, does affect the proper time distribution, because it has a lower efficiency for lower decay lengths. This cut will therefore have to be relinquished for a lifetime measurement. Since removing this cut will significantly increase the background, a new optimization will have to be performed for the remaining cuts to obtain the best signal significance.

The re-tuning resulted in the modification of three cuts. First, 13 CFT hits are required on each track of the K_S^0 (see Fig. 4.25). Secondly, the cut on the χ^2 of the J/ψ vertex was tightened to 6, as shown in Fig. 4.26. Finally, the cut on the χ^2 of the B_d^0 "track" w.r.t. the primary vertex was relaxed to 50 (see Fig. 4.27). The final selection cuts thus are:

- J/ψ:
  - confirmation of both tracks by the muon system;
  - track $p_T > 2.5$ GeV;
  - at least 3 SMT hits on both tracks;
  - vertex $\chi^2 < 6$;

  - $p_T(J/\psi) > 3$ GeV;
  - mass window of 2.8 – 3.6 GeV;

- K^0_S:
  - good momentum measurement for both tracks: $\sigma(1/p_T)/(1/p_T) < 0.5$;
  - tracks displaced w.r.t. the primary vertex: $\chi^2_{PV} > 9$ for both tracks and $\chi^2_{PV} > 16$ for at least one track;
  - less than 2 misses on each track beyond the vertex;
  - no more than 2 hits on both tracks combined before the vertex;
  - vertex $\chi^2 < 30$;
  - $p_T(K^0_S) > 0.4$ GeV;
  - mass window of 0.44 – 0.53 GeV;
  - collinearity w.r.t. the $(J/\psi K^0_S)$ vertex > 0.997;

- B^0_d:
  - χ^2 of B^0_d "track" w.r.t. the primary vertex < 50;
  - vertex $\chi^2 < 25$;
  - $p_T(B^0_d) > 7$ GeV;

The mass spectrum that results from these selections, and that will be used for the lifetime measurement, is shown in Fig. 4.28 (a). We used a Gaussian plus a straight line to fit the spectrum. The fitted B^0_d mass is 5.29 ± 0.01 GeV, and the width of the Gaussian is 31 ± 9 MeV. The number of B^0_d's is 56 ± 6 over a background of 133 events in a mass window of 5.16 to 5.4 GeV.

Figure 4.28 (b) shows the $(J/\psi K^0_S)$ mass spectrum from the $B^0_d \rightarrow J/\psi K^0_S$ signal MC sample containing 20,000 events. The mass peak contains 496 B^0_d candidates. The reconstruction efficiency strongly depends on the generated p_T spectrum of the b-quarks, but it is clear that of the B^0_d's that are not reconstructed, most are not reconstructed because the K^0_S has been missed. The low K^0_S reconstruction efficiency is due in part to the long lifetime: The average decay length of K^0_S's at DØ is 5.8 cm. In addition the energy release in the decay to two pions is rather low, which results in many pions falling below the p_T threshold in the track reconstruction.

To properly fit the B^0_d mass peak in the signal MC sample, a double Gaussian is needed. The first Gaussian holds 57 ± 9% of the events, and has a width of 18 ± 2 MeV. The second Gaussian has a width of 46 ± 5 MeV. The weighted average of the Gaussian widths is 33 MeV.

4.6 Background samples

There are two important processes contributing to background of the selection of the B^0_d in the decay mode to $(J/\psi K^0_S)$. One is the decay of B-hadrons to J/ψ plus other decay products. The

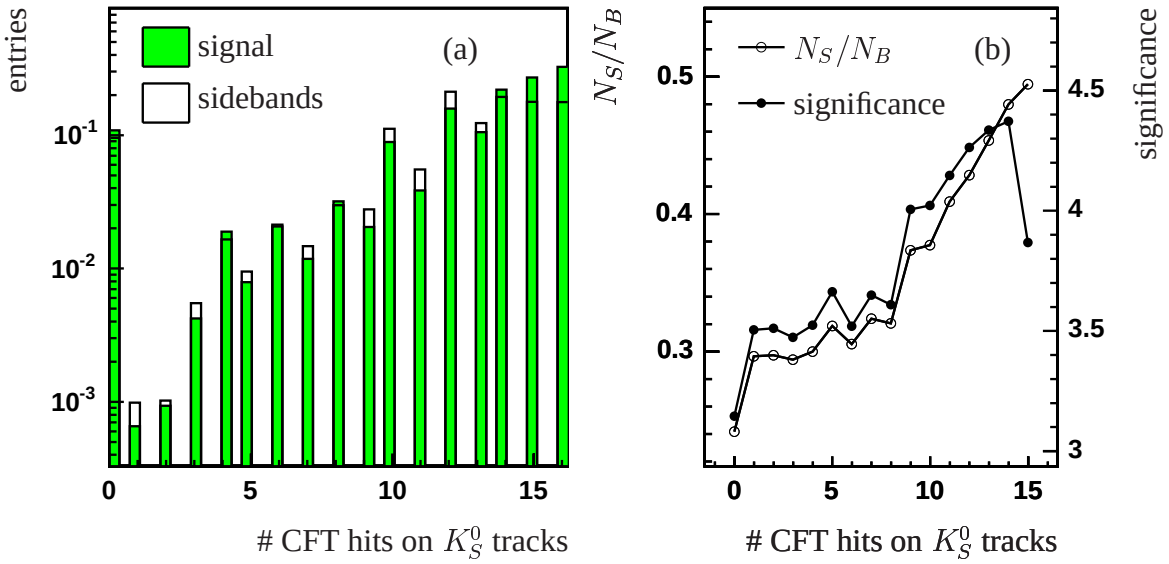


Figure 4.25: (a) The distribution of the number of CFT hits on each track of the K_S^0, and (b) the B_d^0 significance as a function of the cut on this quantity.

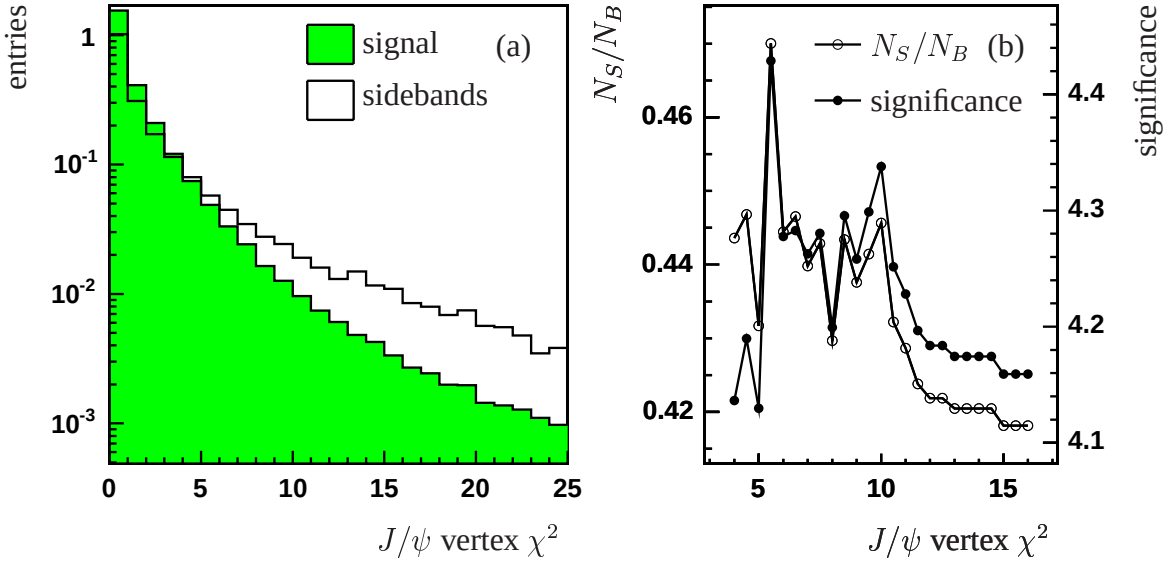


Figure 4.26: (a) J/ψ vertex χ^2 distribution in B_d^0 signal and sideband regions, and (b) B_d^0 significance as a function of this cut.

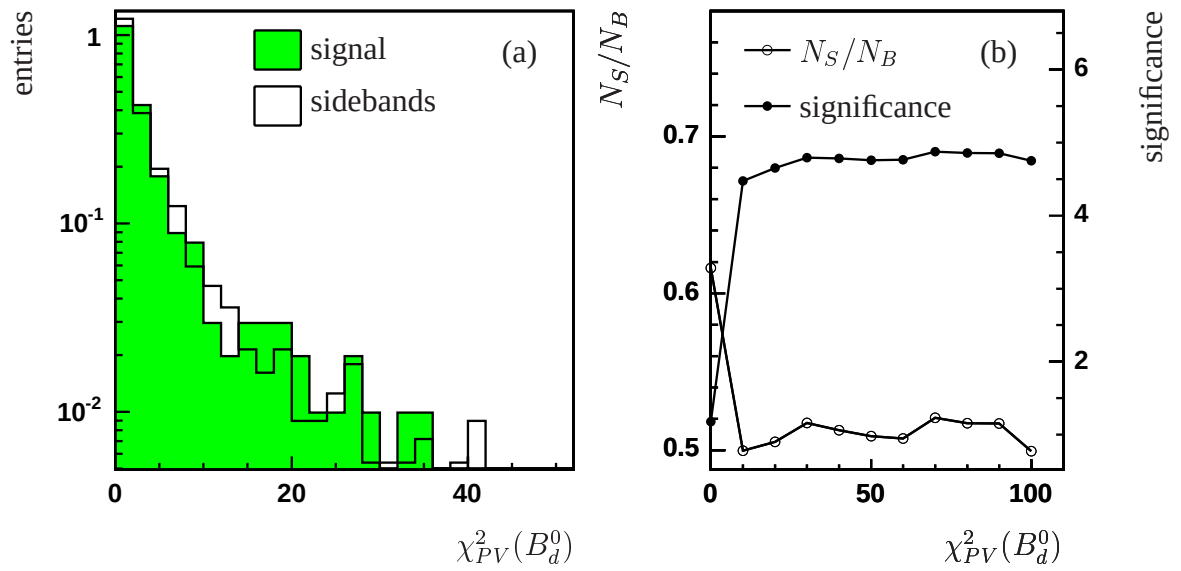


Figure 4.27: (a) Distribution of the χ^2 of B^0_d "track" w.r.t. the primary vertex in B^0_d signal and sideband regions, and (b) B^0_d significance as a function of the cut on this quantity.

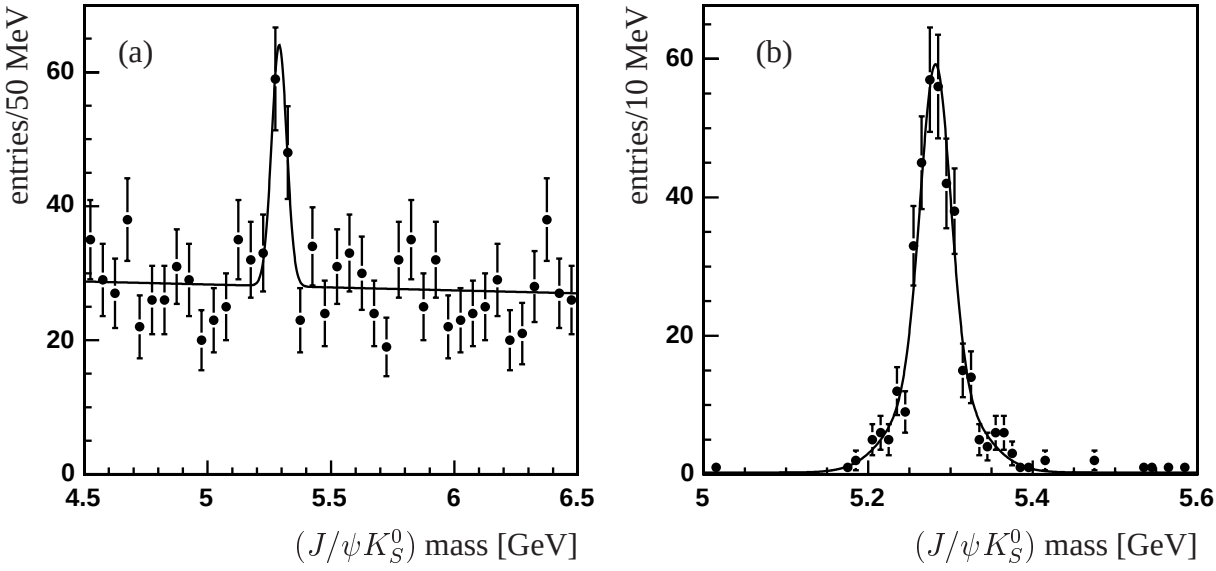


Figure 4.28: $(J/\psi K^0_S)$ invariant mass distribution, resulting from the selection that has been optimized for the B^0_d lifetime measurement, using (a) our data sample and (b) a $B^0_d \to J/\psi K^0_S$ Monte Carlo sample. The fit to our data sample yields 56 ± 6 B^0_d's over a background of 133 events, at a mass of 5.29 ± 0.01 GeV. The width of the Gaussian B^0_d mass peak in data is 31 ± 9 MeV. The fit function used for the Monte Carlo sample is a double Gaussian and has a standard deviation of 33 MeV.

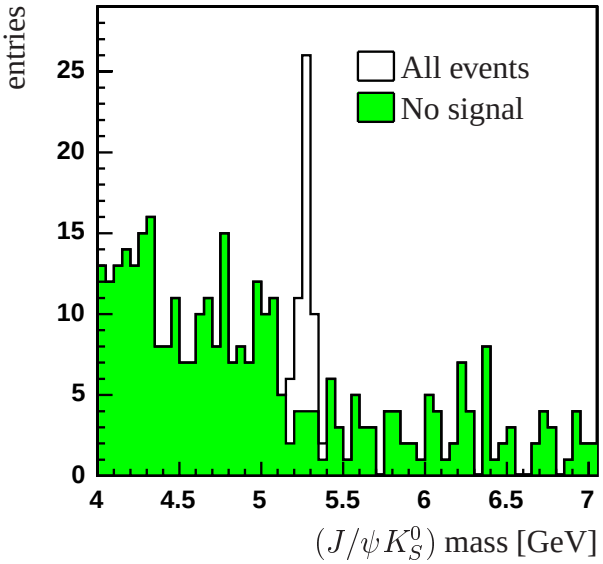


Figure 4.29: The $B \to J/\psi X$ mass spectrum, and the same spectrum with the $B_d^0 \to J/\psi K_S^0$ signal removed, obtained from our $B \to J/\psi X$ MC sample.

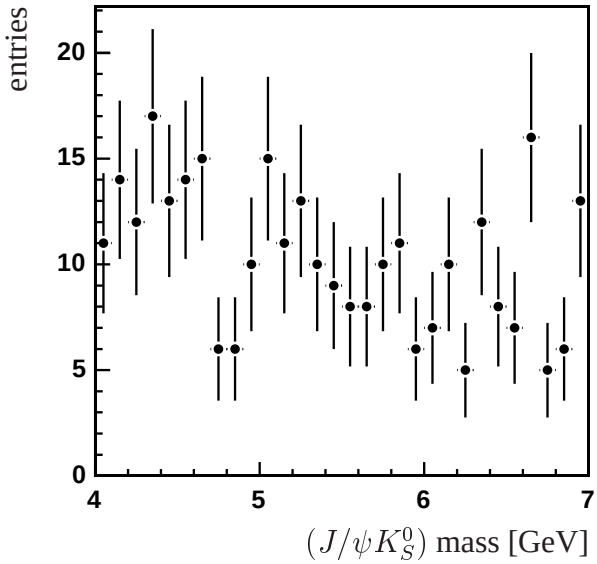


Figure 4.30: The $(J/\psi K_S^0)$ mass spectrum in prompt J/ψ decays, obtained from our prompt J/ψ MC sample.

B_d^0 candidate is formed by combining this J/ψ with a K_S^0 either from the primary vertex or from the B-decay. We denote this background as $B \to J/\psi X$, or long-lived background. The other background is due to prompt J/ψ production.

We have run our selection on a Monte Carlo sample of $B \to J/\psi X$ events and on a Monte Carlo sample of events with prompt J/ψ's. The technical details of these MC samples are described in appendix A. In this section we show the results of this exercise, to prove that we are not drawing fake B_d^0 candidates into the mass peak and that the mass peak in data consists of real B_d^0's.

Our Monte Carlo sample of $B \to J/\psi X$ events holds 100,000 events, including genuine $B_d^0 \to J/\psi K_S^0$ signal candidates. We can remove these from the sample by comparing our reconstructed B_d^0 candidates to the generator information. Figure 4.29 shows the $B \to J/\psi X$ mass spectrum before and after removing the signal candidates. After removal of the signal candidates, there is no peak left in the mass spectrum, indicating that no $B \to J/\psi X$ decays are drawn into the B_d^0 mass peak by our reconstruction or selection.

Our sample of prompt J/ψ events does not have any signal events, so we can simply run our selection on these events and check the mass spectrum. Figure 4.30 shows the mass spectrum of B_d^0 candidates. The sample consists of around 100,000 events, but the selection rate is much lower in this sample than in the $B \to J/\psi X$ sample, because the number of K_S^0's is much smaller in these events. It is clear that the mass spectrum from our prompt J/ψ sample is also not biased and does not reconstruct events preferentially around the B_d^0 mass.

Chapter 5

Measurement of the B_d^0 lifetime

In this chapter the measurement of the lifetime of B_d^0 mesons will be described. The measurement strategy is to reconstruct the transverse decay length of the B_d^0, L_{xy}, and its transverse momentum p_T. The proper time $c\tau(B_d^0)$ will be calculated using the relation:

$$c\tau(B_d^0) = \frac{L_{xy}}{p_T} \cdot m_{B_d^0}. \tag{5.1}$$

Note that we refer to $c\tau$ as the proper time, although it has units of distance. This is just a matter of definition. The $c\tau$ distribution follows an exponential distribution, and the B_d^0 lifetime is the time constant of this distribution. For the mass of the B_d^0 we will use the PDG value of 5.2794 GeV [11]. Equation 5.1 follows directly from the standard relation:

$$c\tau(B_d^0) = \frac{L_{xyz}}{|\vec{p}|} \cdot m_{B_d^0}. \tag{5.2}$$

L_{xyz} is the decay length (in three dimensions) and $|\vec{p}|$ is the momentum of the B_d^0. Equation 5.1 can be found by multiplying both the numerator and the denominator of Equation 5.2 with $\sin(\theta)$, where θ is the angle of the B_d^0 flight direction with the beam axis. The B_d^0 lifetime $\lambda(B_d^0)$ will be determined from a fit to the proper time distribution.

In addition to the signal, the measured $c\tau$ distribution will contain background from a number of different sources. We have studied the proper time distribution of the most important backgrounds in MC. We compare this to the distributions we see in data in the sideband mass regions to the low and high side of the mass peak. We use the results from these studies to determine our background in the signal region.

Finally, a maximum likelihood fit is performed to extract $\lambda(B_d^0)$ from the data. We will describe our fit method and fit results. The chapter is concluded with a discussion of the systematic uncertainty on the result.

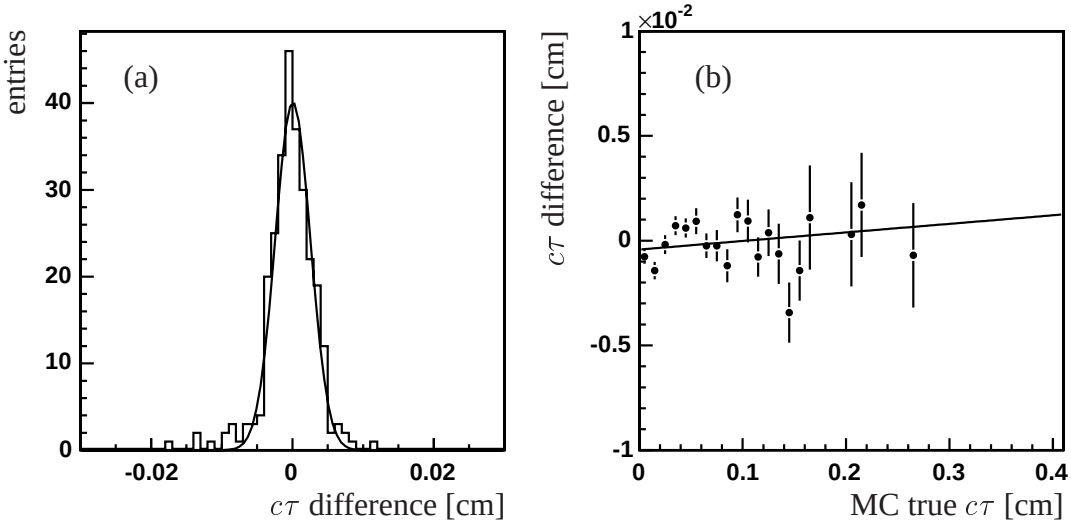


Figure 5.1: The difference between reconstructed $c\tau$ and MC true $c\tau$, for reconstructed B_d^0 signal candidates in a $B_d^0 \to J/\psi K_S^0$ Monte Carlo sample. The fit in (b) is a straight line, with a slope of $(4.1 \pm 3.6) \cdot 10^{-3}$ and has a χ^2 per degree of freedom of 1.88.

5.1 Proper time reconstruction

The proper time of each candidate is calculated using the relation 5.1. We have shown in sections 3.4 and 3.5, that our reconstruction algorithm does not introduce a bias in p_T and L_{xy} for MC data. The MC being a realistic description of the data, there is no reason to expect biases to be introduced in these variables by our algorithm in real data. However, in principle it is still possible that the resulting proper time distribution is biased. To check this, we plot the distribution of the difference between the reconstructed $c\tau$ and MC true $c\tau$ of reconstructed B_d^0 candidates in our signal MC sample. The result is shown in Fig. 5.1 (a). The width of this distribution is 24 ± 1 μm, and the mean is 0.6 ± 2 μm. We can also plot the difference as a function of MC true $c\tau$. This distribution is shown in Fig. 5.1 (b). The points in this plot are the average in each $c\tau$ bin. The error bars indicate the uncertainty on the average, calculated as $\sigma/\sqrt{N_i}$, with σ the standard deviation of all residuals (24 μm) and N_i the number of entries in each bin. The fit is a straight line fit, and has a slope of $(4.1 \pm 3.6) \cdot 10^{-3}$, which is consistent with zero.

Finally, it is necessary to verify that the B_d^0 reconstruction efficiency as a function of proper time is constant, because a varying reconstructing efficiency will modify the expected shape of the $c\tau$ distribution. In Monte Carlo data we can determine the reconstruction efficiency from the generator information. We have used our signal MC sample for this study. If we match all reconstructed B_d^0's to their partners in the generator information, we can plot the reconstructed B_d^0's as a function of MC true $c\tau$. This is shown in Fig. 5.2 (b). The efficiency can be found by simply dividing this distribution by the $c\tau$ distribution of all B_d^0's in the generator information (see Fig. 5.2 (a)). The result is shown in Fig. 5.3. We fit a straight line to the data-points, and find that the slope of the line is -0.001 ± 0.02, i.e. consistent with zero.

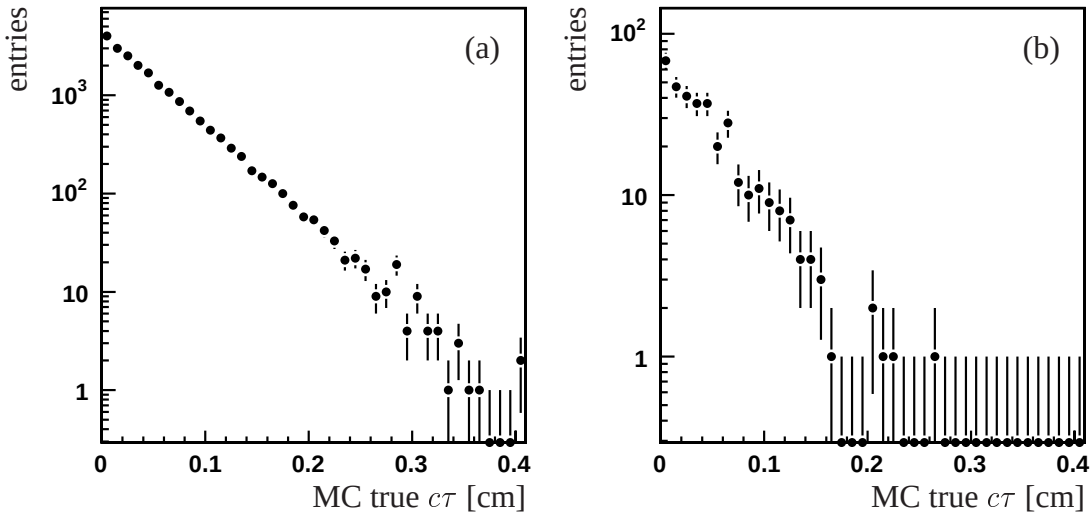


Figure 5.2: Monte Carlo true proper time distribution of (a) all generated B_d^0's, and (b) reconstructed B_d^0 candidates, from a $B_d^0 \to J/\psi K_S^0$ MC sample.

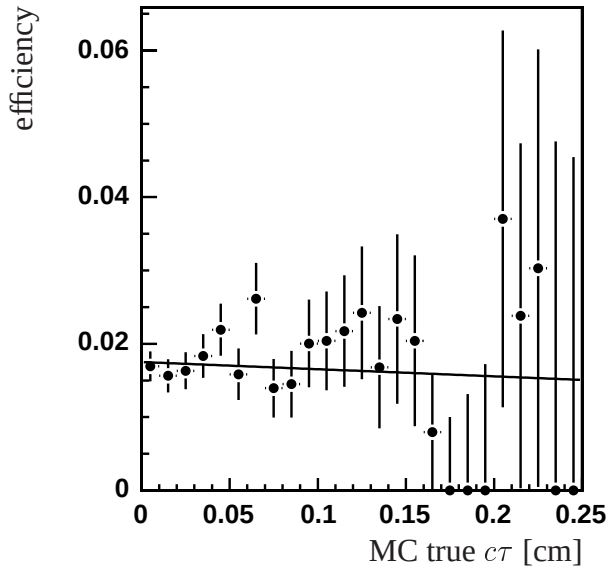


Figure 5.3: Reconstruction efficiency as a function of MC true $c\tau$, determined from a $B_d^0 \to J/\psi K_S^0$ MC sample. The line is a straight-line fit to the data-points. The slope of the line is compatible with zero. The fit has a χ^2 per degree of freedom of 0.73.

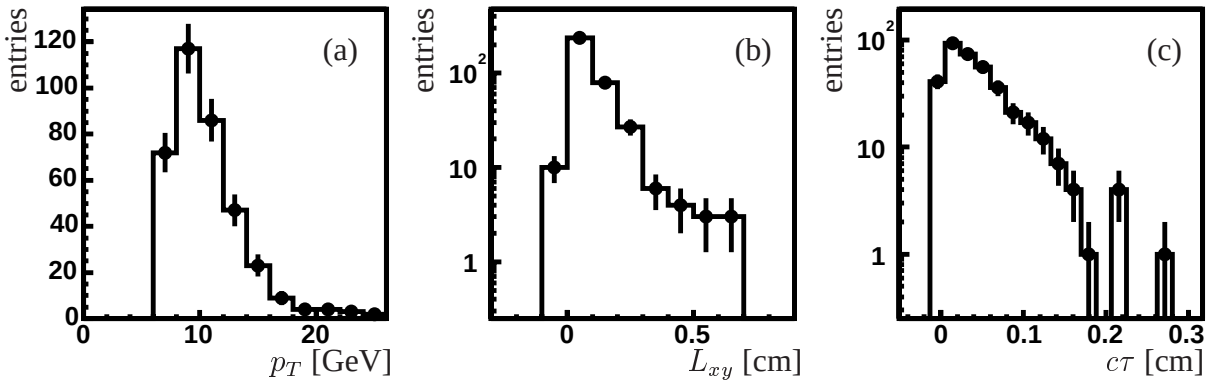


Figure 5.4: The number of selected events from a $B_d^0 \to J/\psi K_S^0$ MC sample, as a function of (a) p_T, (b) decay length and (c) proper time.

5.2 Proper time distributions

The proper time distribution of the events in the signal mass region has a contribution from signal events, as well as contributions from different kinds of backgrounds. We can begin by dividing the background into a class of events that contains a fake J/ψ, and a class of events that has a true reconstructed J/ψ, but does not contain a $B_d^0 \to J/\psi K_S^0$ decay. By comparing the background under the J/ψ peak in the dimuon mass spectrum in Fig. 4.6 to the background under the B_d^0 peak in the (J/ψ K_S^0) mass spectrum (see Fig. 4.28 (a)), we can tell that most of the background falls in the second category. Thus, the background events generally have a true J/ψ. The main sources of J/ψ production at the Tevatron are prompt J/ψ production and J/ψ's from B-decays (which we denote as $B \to J/\psi X$).

In the following sections we will study the proper time distribution in a signal MC sample, as well as in two background MC samples, followed by a discussion of proper time distributions in data. Technical details of the generation of the MC samples are given in appendix A.

5.2.1 $B_d^0 \to J/\psi K_S^0$ decays in Monte Carlo

Figure 5.4 shows the decay length, p_T and proper time distributions for a MC sample of $B_d^0 \to J/\psi K_S^0$ decays. The decay length and proper time distributions both display the expected positive tail.

5.2.2 Prompt J/ψ decays in Monte Carlo

We reconstruct and select events from a prompt J/ψ MC sample of 96,328 events, using the same selection cuts we use to select B_d^0 candidates. The decay length, p_T and proper time distributions of the events passing this selection are shown in Fig. 5.5.

From Fig. 5.5 we conclude, that the $c\tau$ spectrum measured from a prompt J/ψ MC sample is compatible with a Gaussian centered at zero. No long positive or negative tails are observed.

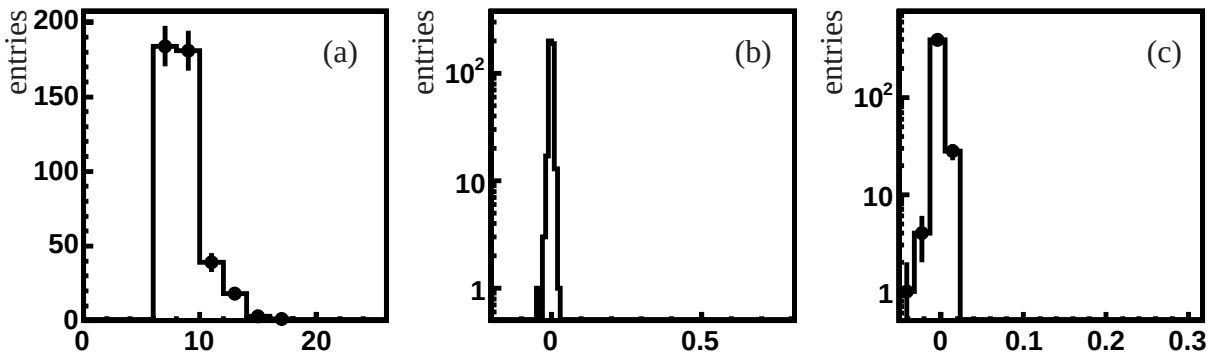


Figure 5.5: The number of selected events from a prompt J/ψ MC sample, as a function of (a) p_T, (b) decay length and (c) proper time. A very loose mass window cut has been applied to the selected candidates: 4.5 to 7.0 GeV.

5.2.3 $B \to J/\psi X$ decays in Monte Carlo

In the previous subsection it was concluded that prompt J/ψ's do not have a measurable lifetime. B-hadrons decaying to J/ψ on the other hand, do have a measurable lifetime. To within approximately 20%, the lifetime of all B-hadrons, but the B_c, (i.e. the B_d^0, $B^{\pm}$, B_s, and Λ_b) have the same lifetime. The B_c has a much shorter lifetime, because also the c-quark can decay weakly. The B_c doesn't have a large effect on the average B-hadron lifetime though, because it is produced at a rate which is only a small fraction of the total B-hadron production rate. This means that the expected lifetime distribution on a sample of $B \to J/\psi X$ is expected to be similar to the lifetime distribution of our signal events. In this sample 50% of the events contain a $B_d^0 \to J/\psi K_S^0 X$ decay, and of these events, 20% contain a $B_d^0 \to J/\psi K_S^0$ decay, meaning that 10% of the entire sample are $B_d^0 \to J/\psi K_S^0$ events.

We reconstruct and select $B_d^0 \to J/\psi K_S^0$ events from a large Monte Carlo sample of $B \to J/\psi X$ decays in exactly the same way as we do in data. The mass spectrum formed by the candidates from the $B \to J/\psi X$ sample has been shown in Fig. 4.29. Distributions of p_T, decay length and $c\tau$ of these candidates are shown in Fig. 5.6. Comparing these distributions to the distributions from direct J/ψ production in Fig. 5.5, shows that the p_T distributions are similar, but the decay length and thus the proper time distribution have long positive tails.

5.2.4 Sideband regions in real data

We will now move from MC to real data. In the $(J/\psi K_S^0)$ mass spectrum shown in Fig. 4.28 (a) we defined a signal mass region from 5.16 GeV to 5.4 GeV, containing a good fraction of signal events compared to the number of background events. We also define two sideband regions containing only background events: a low-mass sideband from 4.55 GeV to 5.10 GeV, and a high-mass sideband from 5.6 GeV to 6.6 GeV.

The proper time distributions for the low-mass and the high-mass sidebands are shown in Fig. 5.7 and Fig. 5.8, respectively. It is clear that there are many more long-lived decays in

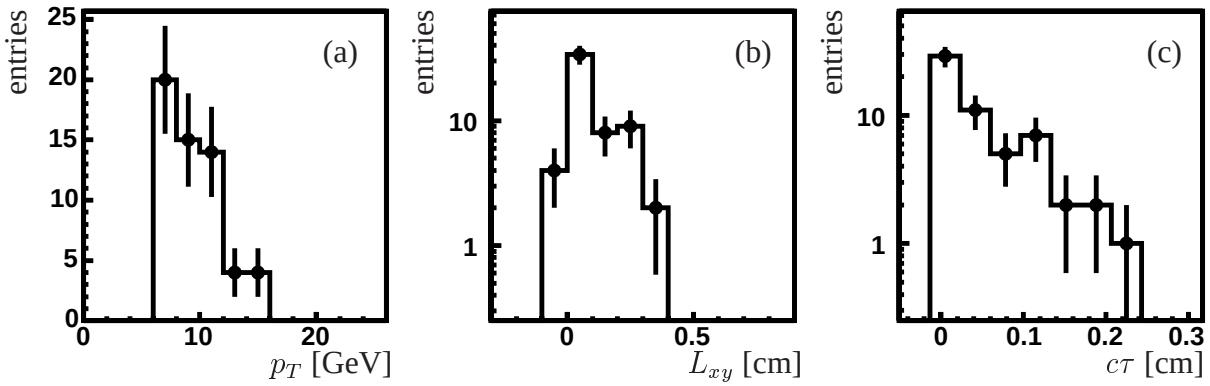


Figure 5.6: The number of selected events from a $B \to J/\psi X$ MC sample, as a function of (a) p_T, (b) decay length and (c) proper time. The selected events are from the B_d^0 signal mass window (5.16 to 5.40 GeV).

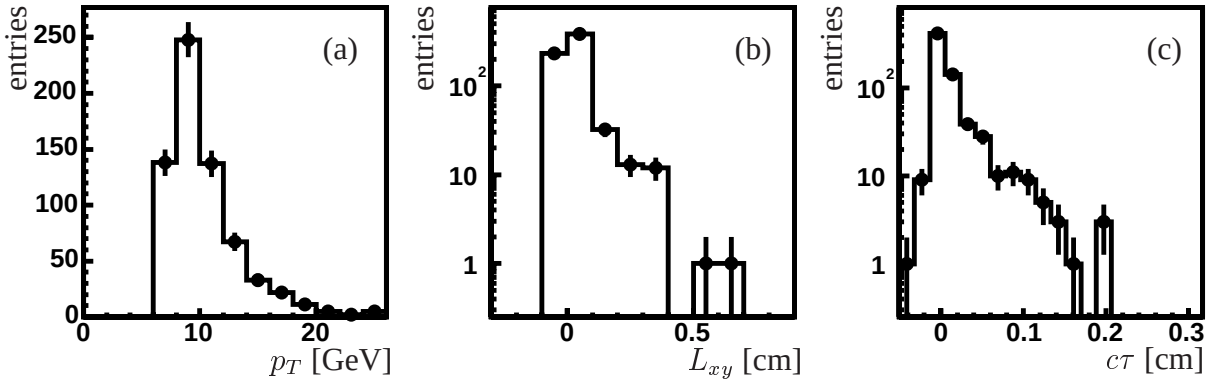


Figure 5.7: The number of selected events from the low-mass sideband region (real data), as a function of (a) p_T, (b) decay length and (c) proper time.

the low-mass sideband than there are in the high-mass sideband. Also, if we compare the mass spectra of the two most important backgrounds in Figs. 4.29 and 4.30, we see that the rise in background below the B-mass is due to the increase in $B \to J/\psi X$ background toward lower mass. We calculate the mass of the B-meson by combining the J/ψ and the K_S^0 decay products, neglecting the rest of the possible decay products. This will cause some B-mesons to be reconstructed too low in mass, and thus to fall in the low-mass sideband region.

5.2.5 Signal region in real data

The decay length, p_T and proper time distributions of events from the signal region (5.16 – 5.4 GeV) are shown in Fig. 5.9. The shape of these distributions mimics what is found in the low-mass sideband, which obviously complicates the measurement of the lifetime of the signal. A solution to this problem will be presented later in this chapter.

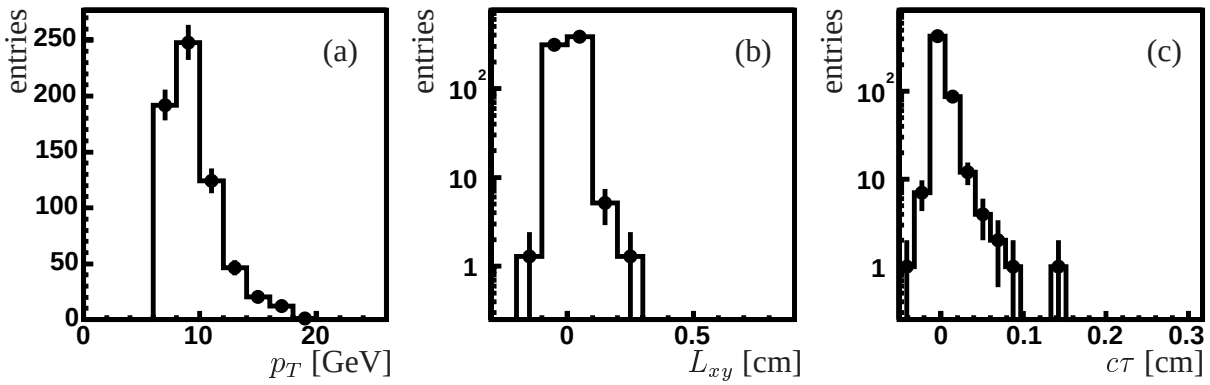


Figure 5.8: The number of selected events from the high-mass sideband region (real data), as a function of (a) p_T, (b) decay length and (c) proper time.

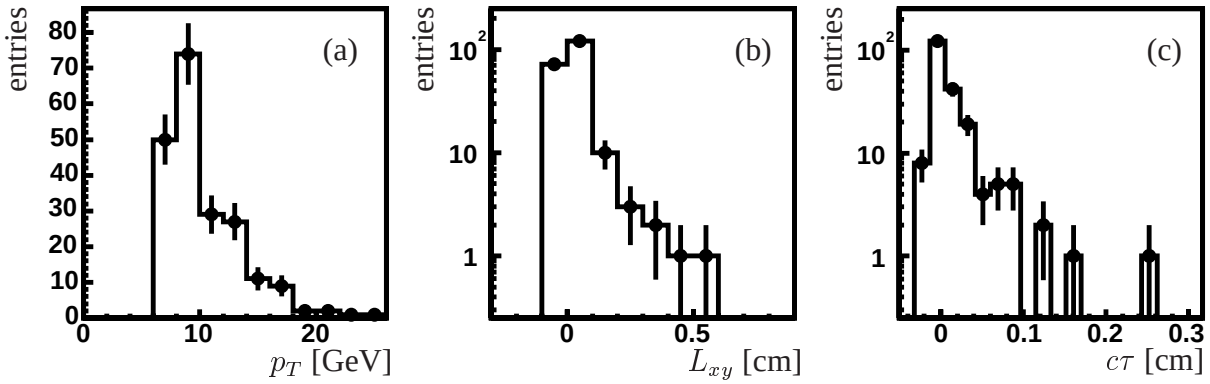


Figure 5.9: The number of selected events from the signal mass region (real data), as a function of (a) p_T, (b) decay length and (c) proper time.

5.3 B_d^0 lifetime measurement using a one-dimensional fit

In a one-dimensional fit we fit the proper time distribution of the sum of various contributions to the data in the signal region. The implementation of the one-dimensional fit is an unbinned maximum likelihood fit. The likelihood function describes the form of the distribution the events are expected to follow. For each set of parameters to be determined the likelihood function takes on a different shape. The set which maximizes the likelihood function is the fit result. The likelihood function is given by:

$$\mathcal{L} = \prod_{i=1}^{N} \Big[\alpha f^i_{sig}\big((c\tau)_i, \sigma(c\tau)_i\big) + (1-\alpha) f^i_{bkg}\big((c\tau)_i, \sigma(c\tau)_i\big) \Big], \tag{5.3}$$

where α is the fraction of signal events in the candidate sample. The index i runs from 1 to the number of B_d^0 candidates N. f^i_{sig} and f^i_{bkg} are probability density functions (PDF's) for signal and background respectively, as a function of the reconstructed proper time $(c\tau)_i$, with error $\sigma(c\tau)_i$, of each candidate.

The background shape, f_{bkg}, is the sum of a contribution from prompt J/ψ production, f_{prompt} and a long-lived contribution f_{LL}. The expected shape of the proper decay length distribution for the prompt J/ψ contribution is a delta peak at zero (representing the negligible lifetime of the J/ψ), smeared with the $c\tau$ resolution function. The shape in $c\tau$ for the prompt J/ψ contribution is a double Gaussian:

$$\begin{aligned} f_{prompt} \;=\; & \frac{1-\alpha_2}{F_{c\tau,1}\cdot\sigma(c\tau)_i\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{(c\tau)_i}{F_{c\tau,1}\cdot\sigma(c\tau)_i}\right)^2\right) + \\ & \frac{\alpha_2}{F_{c\tau,2}\cdot\sigma(c\tau)_i\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{(c\tau)_i}{F_{c\tau,2}\cdot\sigma(c\tau)_i}\right)^2\right). \end{aligned} \tag{5.4}$$

In this function we introduced two scale factors, $F_{c\tau,1}$ and $F_{c\tau,2}$, and the relative weight of $F_{c\tau,2}$, α_2.

The shape of the long-lived background is an exponentially decaying function with decay constant $(c\tau)_{LL}$, convoluted with a Gaussian with width σ_w:

$$f_{LL} = C_{LL}\cdot\exp\left(-\frac{(c\tau)_i}{(c\tau)_{LL}}\right) \otimes \exp\left(-\frac{1}{2}\left(\frac{(c\tau)_i}{\sigma_w}\right)^2\right). \tag{5.5}$$

C_{LL} is a constant normalizing f_{LL} to 1. The width of the Gaussian σ_w is the weighted average of the widths of the double Gaussian from the prompt J/ψ's:

$$\sigma_w = \sqrt{(1-\alpha_2)\cdot\big(F_{c\tau,1}\cdot\sigma(c\tau)_i\big)^2 + \alpha_2\cdot\big(F_{c\tau,2}\cdot\sigma(c\tau)_i\big)^2}. \tag{5.6}$$

The function describing the signal distribution, f_{sig}, is also an exponentially decaying function convoluted with a Gaussian, similar to f_{LL}:

$$f_{sig} = C_{sig} \cdot \exp\left(-\frac{(c\tau)_i}{\lambda(B_d^0)}\right) \otimes \exp\left(-\frac{1}{2}\left(\frac{(c\tau)_i}{\sigma_w}\right)^2\right), \tag{5.7}$$

where C_{sig} is again the normalization constant, and $\lambda(B_d^0)$ is the B_d^0 lifetime.

5.3.1 Fit to low-mass sideband data

In the low-mass sideband, the long-lived contribution can easily be distinguished from the contribution of prompt J/ψ. The $c\tau$ distribution of the low-mass sideband has already been shown in Fig. 5.7 (c). In Fig. 5.10 we show this distribution again, this time with a fit. The fitted function is the sum of f_{prompt} and f_{LL}, as given in equation 5.4 and 5.5. The fit results are listed in Table 5.1.

We define a χ^2 per degree of freedom of the fit, by comparing the deviation of each data-point N_i with error σ_i with the fit function f:

$$\chi^2_{fit} = \frac{1}{N_{bins} - N_{par}} \sum_{i=1}^{N_{bins}} \left(\frac{N_i - f_i}{\sigma_i}\right)^2, \tag{5.8}$$

where the index i runs from 1 to the number of bins N_{bins}, and f_i is the value of the fit function in bin i. N_{par} is the number of free fit parameters. For this fit χ^2_{fit} is 0.47.

5.3.2 Fit to high-mass sideband data

We can also apply this fit to events from the high-mass sideband. In this mass region, the long-lived background contribution is less prominent and falls off more quickly than it does in the low-mass sideband. Therefore, it cannot be distinguished as easily from the prompt J/ψ production and we have chosen to fix the parameters of the double Gaussian to the values determined from the fit to the low-mass sideband. The remaining free parameters are the fraction of events that are from prompt J/ψ's, and the slope of the exponential fitting the long-lived background. The fit is shown in Fig. 5.11. It has $\chi^2_{fit} = 0.74$, and the results are listed in Table 5.2. It is clear that the lifetime of the long-lived background measured in the low-mass sideband is incompatible with the value measured in the high-mass sideband.

5.3.3 Extraction of the lifetime from signal region data

The fit to the events in the signal region will involve the sum of f_{prompt} and f_{LL} to accommodate the background, plus f_{sig} for the signal. With the current statistics, it is impossible to resolve two exponentials, f_{sig} and f_{LL}, from a single fit, if the slopes of the exponentials are of the same order of magnitude. We will therefore have to fix the slope $(c\tau)_{LL}$ and the relative normalization $(1 - \alpha_{prompt})$ of the long-lived background in our final fit. In the previous section, it has become clear that the two sidebands yield different values for $(c\tau)_{LL}$ and α_{prompt}. This issue is difficult to resolve using a one-dimensional fit, because this fit does not use information from the mass

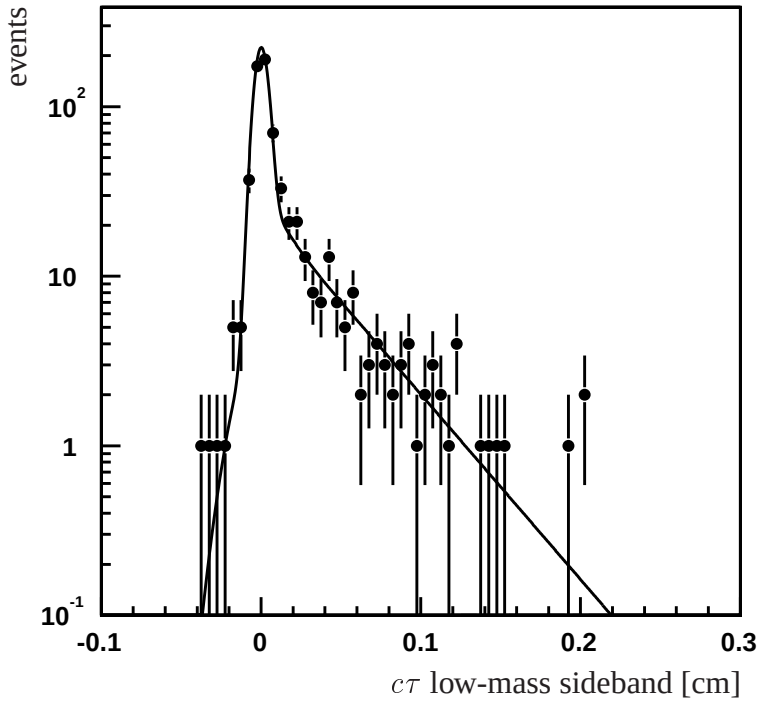


Figure 5.10: A fit to the proper time distribution of data events from the low-mass sideband (mass range 4.1 – 5.1 GeV). The fit shape is a double Gaussian to fit the prompt J/ψ contribution and an exponential convoluted with a Gaussian to accommodate the long-lived background. The fit χ^2 per degree of freedom is 0.47.

Variable	value
lifetime $(c\tau)_{LL}$	483^{+46}_{-41} μm
prompt fraction α_{prompt}	0.66 ± 0.04
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.65 ± 0.150
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	5.49 ± 2.731
weight second scale α_2	0.05 ± 0.06

Table 5.1: Results from fitting a double Gaussian and an exponential convoluted with a Gaussian to events from the low-mass sideband.

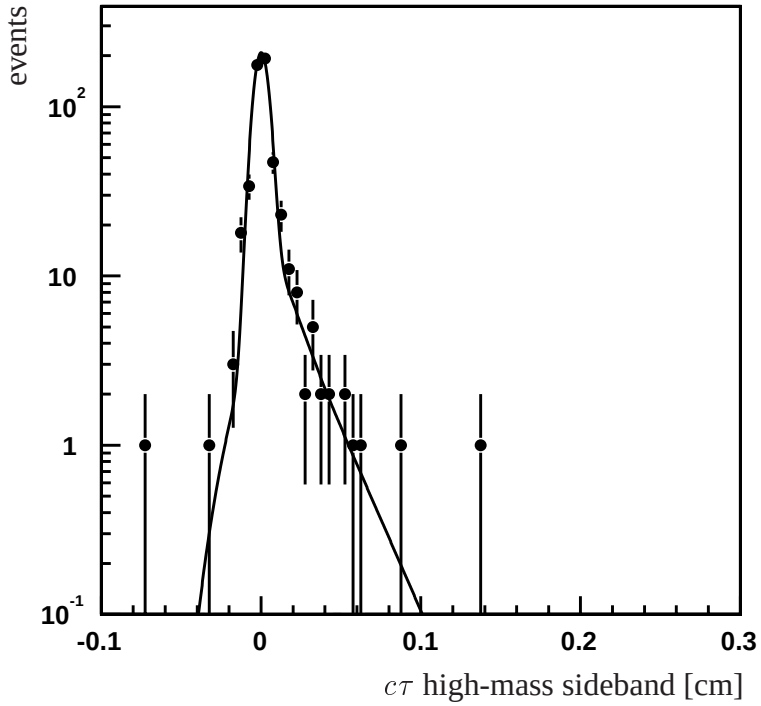


Figure 5.11: A fit to the proper time distribution of data events from the high-mass sideband (mass range 5.6 – 6.6 GeV). The fit shape is a double Gaussian to fit the prompt J/ψ contribution and an exponential convoluted with a Gaussian to accommodate the long-lived background in this sample. The fit χ^2 per degree of freedom is 0.74.

Variable	value
lifetime $(c\tau)_{LL}$	169^{+34}_{-27} μm
prompt fraction α_{prompt}	0.86 ± 0.04
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.65 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	5.49 (fixed)
weight second scale α_2	0.05 (fixed)

Table 5.2: Results from fitting a double Gaussian and an exponential convoluted with a Gaussian to events from the high-mass sideband. The parameters of the prompt J/ψ PDF have been fixed to the values obtained from the low-mass sideband (see Table 5.1).

spectrum. For instance, it is impossible to introduce a mass dependence of $(c\tau)_{LL}$. This forces us to use a two-dimensional fit, which utilizes the reconstructed mass of the events in addition to the proper time, for our final result.

5.4 B_d^0 lifetime measurement using a two-dimensional fit

The extraction of the B_d^0 lifetime using a one-dimensional method is problematic, due to the limited discrimination between signal and long-lived background, and the varying lifetime of the long-lived background as a function of mass. To counter these problems, we have introduced a second dimension, the reconstructed mass, in the fit. This allows us to assign a much higher probability to signal events, when they are compatible with the mass peak. This greatly improves the discrimination between signal and background. Moreover, it enables us to implement a mass dependence in the long-lived background contribution.

5.4.1 Two-dimensional fit method

The two-dimensional fit is implemented, like the one-dimensional fit, as an unbinned maximum likelihood fit. The likelihood function for the two-dimensional fit is given by:

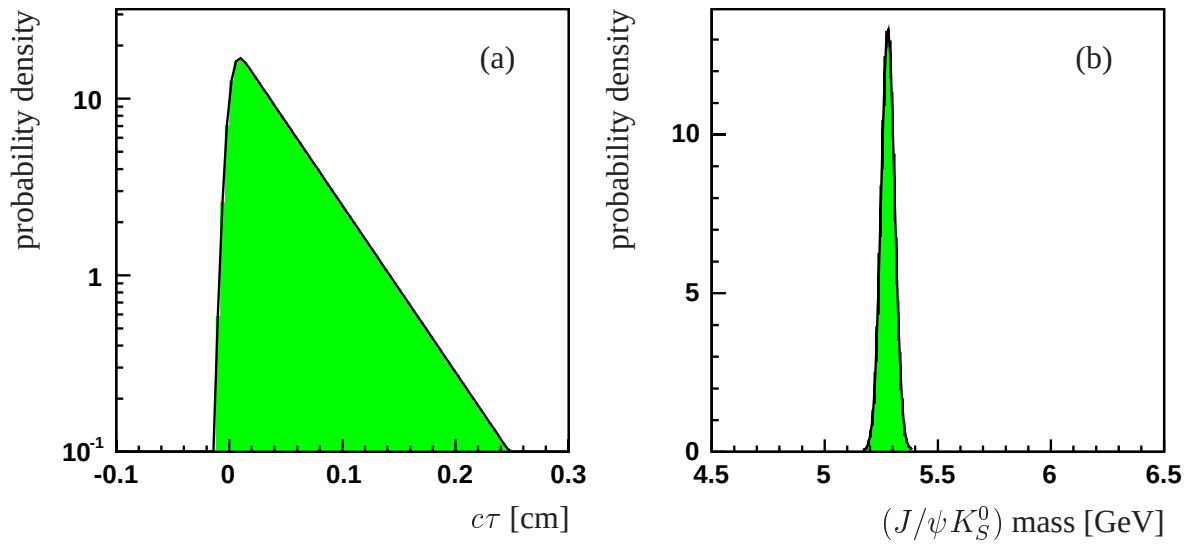
$$\mathcal{L} = \prod_{i=1}^{N} \Big[\alpha f_{sig}^{i}\big(m_i, (c\tau)_i, \sigma(c\tau)_i\big) + (1-\alpha) f_{bkg}^{i}\big(m_i, (c\tau)_i, \sigma(c\tau)_i\big) \Big]. \tag{5.9}$$

Note that the PDF's for the signal and background defined here also use the reconstructed mass m_i of each candidate. The PDF for signal, f_{sig}, is the product of the signal PDF's in mass and time: $f_{sig} = f_{sig}^{mass} \cdot f_{sig}^{time}$. The PDF in mass, f_{sig}^{mass}, is a double Gaussian with mean $m_{B_d^0}$ and standard deviations $\sigma_{B,1}$ and $\sigma_{B,2}$. The fraction of events in the second Gaussian we label $\alpha_{B,2}$. The PDF in time, f_{sig}^{time}, has already been defined in equation 5.7. The functional shapes describing the signal are shown in Fig. 5.12 (a) and (b).

The PDF for background is the sum of a prompt J/ψ and a long-lived contribution. In the two-dimensional fit, however, the PDF of each contribution is the product of a PDF in proper time and a PDF in mass, so that we can write the PDF of the prompt J/ψ contribution f_{prompt} as $f_{prompt}^{mass} \cdot f_{prompt}^{time}$. Fig. 4.30 shows the mass distribution of prompt J/ψ decays in Monte Carlo events. For f_{prompt}^{mass} we will use a uniform distribution. The proper time distribution of prompt J/ψ decays, f_{prompt}^{time}, is a double Gaussian, as defined in equation 5.4. We show the shapes describing the prompt J/ψ mass and proper time distribution in Fig. 5.12 (c) and (d). Finally, the implementation of the long-lived background in the two-dimensional fit is described in the next section.

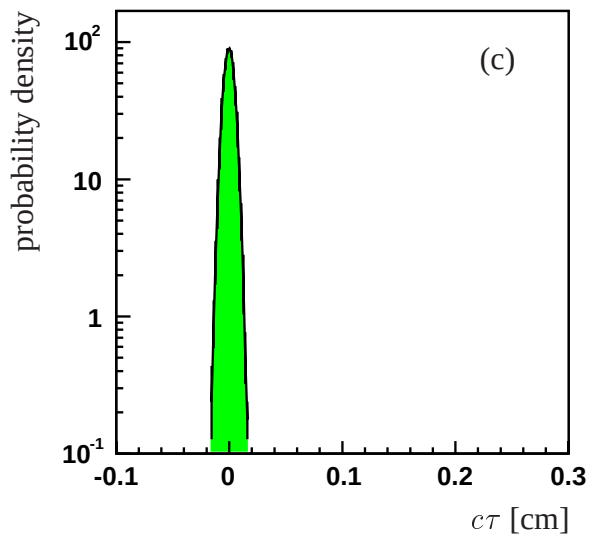
5.4.2 Long-lived background description

It was first noted in section 5.2.4 that the proper time distribution of the long-lived background below and above the B_d^0 mass differs significantly. In this section we will describe the source of this effect and its incorporation into the two-dimensional fit.

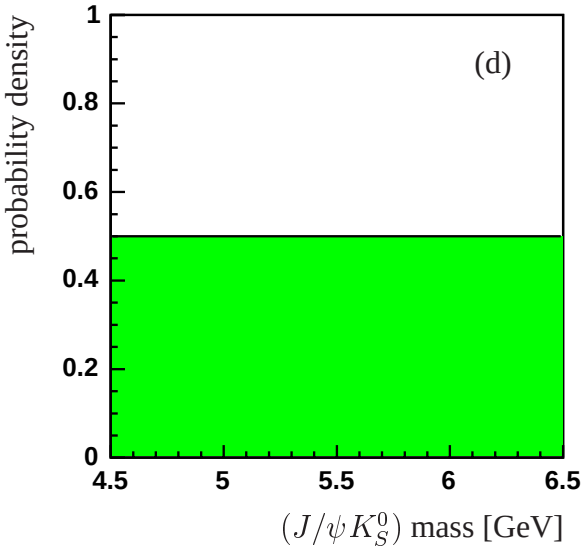


(a) PDF in time for signal.

(b) PDF in mass for signal.



(c) PDF in time for prompt J/ψ contribution.



(d) PDF in mass for prompt J/ψ contribution.

Figure 5.12: Shape of PDF in mass and proper time used for signal and the background from prompt J/ψ production.

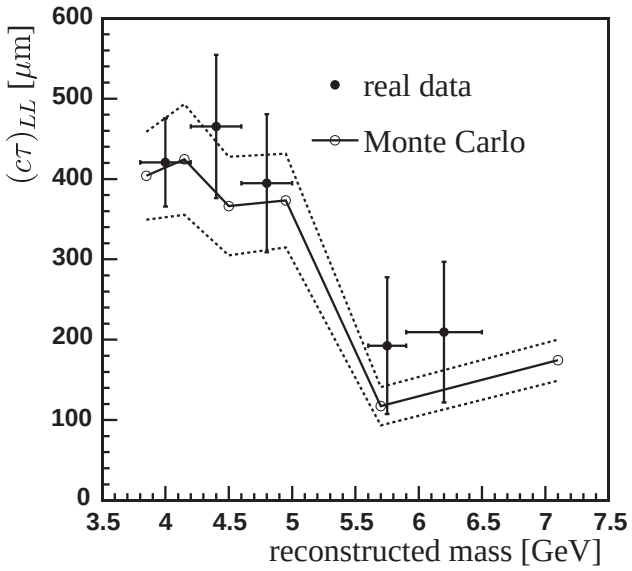


Figure 5.13: The evolution of $(c\tau)_{LL}$ as a function of reconstructed mass, in real data and in a $B \to J/\psi X$ MC sample. The dashed lines indicate the statistical uncertainty on the MC points.

To develop a clearer picture of the effect, we have divided our real data sample in a number of mass bins, and fit for $(c\tau)_{LL}$. The evolution of $(c\tau)_{LL}$ as a function of reconstructed mass in data is shown as points with error bars in Fig. 5.13. We have studied the effect in our $B \to J/\psi X$ MC sample also, after removing the $B \to J/\psi K_S^0$ signal candidates from the sample. The result from this MC sample is shown as the solid line in Fig. 5.13, and is compatible with the shape found in data. The dashed lines indicate the statistical error on the MC value.

The event signature required is a reconstructed J/ψ and K_S^0. Because the effect being studied occurs in long-lived background, and given the stringent selection criteria on the J/ψ reconstruction, it is safe to assume that the J/ψ's originate from B-decays. The pattern observed in the proper time $(c\tau)_{LL}$ then leads us to identify two classes of events. The first class of events we call "unrelated K_S^0". The K_S^0 that joins the J/ψ in the $(J/\psi K_S^0)$ vertex is either from the same B as the J/ψ, or it is unrelated to it. If the K_S^0 is unrelated to the B-decay, it generally means that it originated from the primary vertex. Note that also fake K_S^0's fall into this category, since what has been mistaken for a K_S^0, is completely unrelated to the B-decay. The effective lifetime of events in this class is labeled $(c\tau)_{unrel.K}$.

If the K_S^0 was from the B-decay, but the $(J/\psi K_S^0)$ mass did not reconstruct to the B_d^0 mass within resolution, the B-hadron must have had additional decay products that were not assigned to the $(J/\psi K_S^0)$ vertex. This class of events is called "missing decay products" and the effective lifetime of these events is labeled $(c\tau)_{miss}$. Figure 5.14 shows the mass spectrum and the value of $(c\tau)_{miss}$ for events from this category. The reconstructed mass is almost always lower than the B-hadron mass, which is compatible with the fact that in these events not all of the decay products of the B-hadron are included in the event reconstruction. There are a few events at a mass higher than the B-hadron mass, due to an occasional inaccurate p_T measurement of the J/ψ and/or K_S^0. The decay length, however, is properly reconstructed. Since the events in this

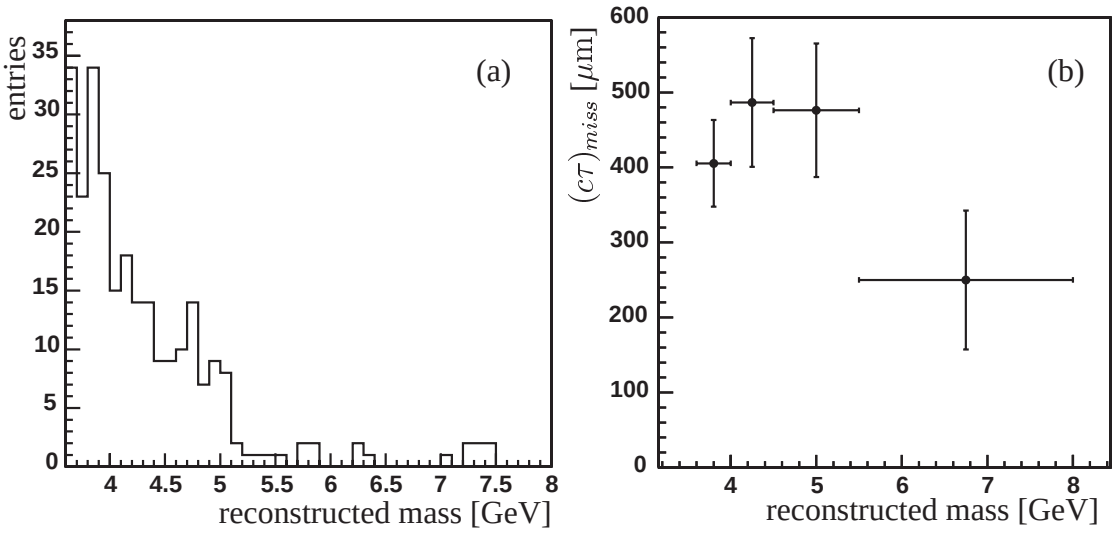


Figure 5.14: (a) Mass spectrum and (b) effective lifetime $(c\tau)_{miss}$ for events in the "missing decay products" category.

category are genuine B-decays we expect to measure a lifetime of the same magnitude as the lifetime of the B-hadron, as is observed (see Fig. 5.14 (b)).

The lifetime measured for events in the "unrelated K_S^0" category, as shown in Fig. 5.15 (b), is smaller than the measured lifetime for events in the "missing decay products" category. The K_S^0 in events of this type generally originate from the primary vertex. Combining the K_S^0 and the J/ψ in one vertex will pull the J/ψ vertex toward the primary vertex. Moreover, for this type of events, the efficiency will be lower for larger decay lengths. The separation between the J/ψ vertex and the K_S^0 "track" is proportional to the decay length of the B-hadron, resulting in a larger vertex χ^2 for large decay lengths.

When comparing the mass spectra of the two event categories, we see that only the "unrelated K_S^0" events contribute to the mass range above the B mass. This results in the low $(c\tau)_{LL}$ seen in this range in Fig. 5.13. At masses below the B mass, the "missing decay products" category dominates.

To incorporate this understanding in the two-dimensional fit, we introduce two contributions to the two-dimensional fit. For each we define a PDF in proper time and a PDF in mass. The shape of the PDF in proper time for both event categories will be identical to the original PDF f_{LL} as defined in equation 5.5, only the time constants will be different. We take the PDF in mass of the unrelated K_S^0 contribution to be uniform. In Fig. 5.14 (a) we see that the mass distribution of the missing decay products events has a more distinct shape. If we fit a polynomial to this distribution, we will find that the largest discrepancy between the fit function and the histogram will be around the B mass. This is unsatisfactory, because this is the region where the shape of the polynomial has the biggest effect on our measurement. Therefore we chose to use the mass spectrum as the PDF directly, without fitting. Unfortunately, as is often the case with background MC samples, our signal selection leaves only very few events from our MC sample.

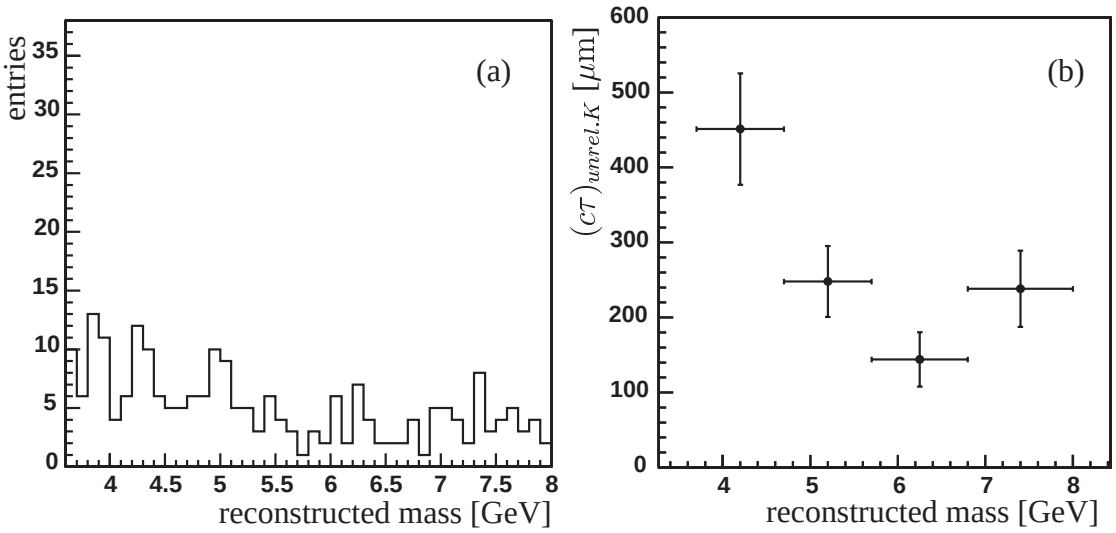


Figure 5.15: (a) Mass spectrum and (b) effective lifetime $(c\tau)_{unrel.K}$ for events in the "unrelated K_S^0" category.

To improve this situation, we reproduce the shape of the mass histogram directly from generator level information. The generator level histogram will have to be smeared with the appropriate resolutions. These resolutions can be derived from the widths of the J/ψ and K_S^0 mass peak, as observed in data, as follows.

We make the assumption that the resolution of the B mass can be completely attributed to the p_T resolution of the J/ψ and the K_S^0. We will now describe the calculation of the p_T resolution of the J/ψ, followed by a description of its relation to the $(J/\psi K_S^0)$ invariant mass resolution. The calculation of the p_T resolution of the K_S^0 is analogous.

If $\vec{p}_1$ and $\vec{p}_2$ are the three-momenta of the muons to which the J/ψ decays, then, ignoring the mass of the muons, the mass of the J/ψ is $m_{J/\psi} = \sqrt{2|p_1||p_2|(1-\cos\omega)}$, with ω the opening angle between the two muons. We want to express the width of the J/ψ peak in the dimuon mass spectrum, $\sigma_{m_{J/\psi}}$, in terms of the muon momentum resolution. The error on the muon momentum, however, is not constant, but is Gaussian distributed in q/p_T (with q the reconstructed charge). We therefore define $k_1 = q_1/p_{T,1}$ and $k_2 = q_2/p_{T,2}$ as Gaussian distributed measures of the p_T of muon 1 and 2. We know that p_T is related to the momentum $|\vec{p}|$ through the angle with the beam axis θ: $|\vec{p}| = p_T/\sin(\theta)$, and therefore $|\vec{p}| = |q/(k\sin(\theta))|$. If we insert this relation in our expression for the J/ψ mass $m_{J/\psi}$, we obtain:

$$m_{J/\psi} = \sqrt{\frac{2(1-\cos\omega)}{k_1\sin(\theta_1)k_2\sin(\theta_2)}}. \tag{5.10}$$

We dropped the charge q, because it is overridden by the absolute-value signs. We can then drop the absolute-value signs if we assume the convention that θ is positive and smaller than $\pi/2$. The mass resolution can be found using the standard relation:

$$\sigma^2_{m_{J/\psi}} = \left(\frac{\partial m_{J/\psi}}{\partial k_1}\sigma_{k_1}\right)^2 + \left(\frac{\partial m_{J/\psi}}{\partial k_2}\sigma_{k_2}\right)^2. \tag{5.11}$$

Note that σ_k is independent of k (to first order), and therefore $\sigma_{k_1} = \sigma_{k_2}$. Inserting equation 5.10 in equation 5.11 yields:

$$\sigma^2_{m_{J/\psi}} = \mathcal{A}^2 \left(\frac{1}{k_1^2} + \frac{1}{k_2^2}\right)\sigma_k^2, \tag{5.12}$$

where we defined $\mathcal{A}$ as:

$$\mathcal{A} = \frac{-1+\cos(\omega)}{m_{J/\psi}\sin(\theta_1)\sin(\theta_2)k_1k_2}. \tag{5.13}$$

The relation between the J/ψ momentum resolution $\sigma_{p_{J/\psi}}$ and the muon momentum resolution follows from the relation $|\vec{p}_{J/\psi}| = \sqrt{|\vec{p}_1|^2 + |\vec{p}_2|^2 + 2|\vec{p}_1||\vec{p}_2|\cos\omega}$. Again we use $k = q/p_T$ to find:

$$|\vec{p}_{J/\psi}| = \sqrt{\frac{1}{k_1^2\sin^2(\theta_1)} + \frac{1}{k_2^2\sin^2(\theta_2)} + \frac{2\cos(\omega)}{k_1\sin(\theta_1)k_2\sin(\theta_2)}}. \tag{5.14}$$

We can use the relation analogous to equation 5.11 to find the J/ψ momentum resolution:

$$\sigma^2_{p_{J/\psi}} = \frac{1}{p^2_{J/\psi}}(\mathcal{B}_1^2 + \mathcal{B}_2^2)\sigma_k^2. \tag{5.15}$$

Here we defined:

$$\begin{aligned} \mathcal{B}_1 &= \frac{1}{k_1^3\sin(\theta_1)} + \frac{\cos(\omega)}{k_1^2\sin(\theta_1)k_2\sin(\theta_2)}, \\ \mathcal{B}_2 &= \frac{1}{k_2^3\sin(\theta_2)} + \frac{\cos(\omega)}{k_1\sin(\theta_1)k_2^2\sin(\theta_2)}. \end{aligned} \tag{5.16}$$

Combining equations 5.12 and 5.15, we find the J/ψ momentum resolution from the width of the mass peak:

$$\sigma^2_{p_{J/\psi}} = \frac{(\mathcal{B}_1^2 + \mathcal{B}_2^2)}{p^2_{J/\psi}\mathcal{A}^2\left(\frac{1}{k_1^2} + \frac{1}{k_2^2}\right)}\sigma^2_{m_{J/\psi}}. \tag{5.17}$$

The Monte Carlo true invariant mass spectrum, using J/ψ's and K_S^0's that come from a single B-decay, is shown in Fig. 5.16 (a). Again, the $B_d^0 \to J/\psi K_S^0$ signal has been removed here. If we smear this spectrum using the procedure outlined above, we find the distribution in Fig. 5.16 (b). This is the distribution that we will use as the mass PDF for the missing decay products event category.

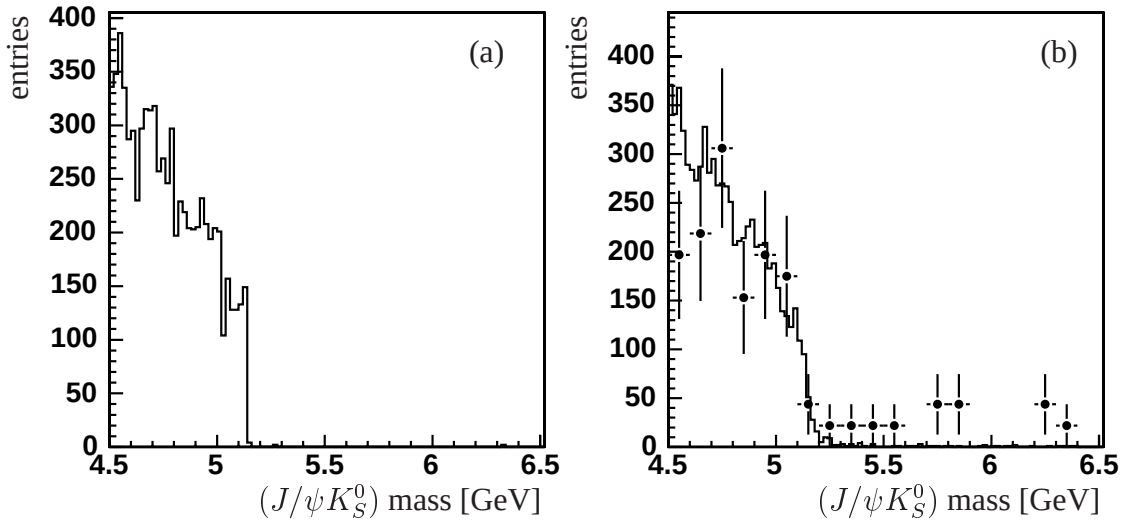


Figure 5.16: The $(J/\psi K_S^0)$ mass spectrum using generator information, obtained from a $B \to J/\psi X$ Monte Carlo sample. Figure (b) shows as a histogram the generator spectrum smeared using the J/ψ and K_S^0 mass resolution from real data. We have overlaid the smeared generator spectrum with the reconstructed spectrum from this sample from Fig. 5.14 (a), shown as points with error bars.

To verify our smearing procedure, we have applied it to our signal MC sample. Our signal MC sample also contains B-decays that are not $B_d^0 \to J/\psi K_S^0$, mostly from the decay of the other-side b-quark. We have removed these for this study, so that before smearing, we expect to find an extremely narrow B_d^0 mass peak, precisely at the B_d^0 mass. If our smearing procedure is correct, applying it to the MC true B_d^0 mass peak will reproduce the shape of the reconstructed B_d^0 mass peak from this sample, as shown in Fig. 4.28 (b). The smeared peak is shown as the points in Fig. 5.17. We have fit a double Gaussian, indicated as the line in Fig. 5.17. The fit results in the first Gaussian having $\sigma = 19 \pm 0.4$ MeV and containing 85% of the events. The second Gaussian has $\sigma = 72 \pm 2$ MeV. The weighted average of the widths of these Gaussians is $\sigma_w = 33$ MeV. The fit to the original reconstructed B_d^0 mass peak, shown in Fig. 4.28 (b), resulted in $\sigma_w = 33$ MeV as well. We therefore conclude that our smearing procedure is adequate and that the smeared generator spectrum from Fig. 5.16 (b) is a sufficiently accurate representation of the mass spectrum of the "missing decay products" events. The PDF's for the "unrelated K_S^0" and "missing decay products" categories, in mass and in proper time, are then all known, and shown in Fig. 5.18.

5.4.3 Reducing the number of free fit parameters

At this point in the analysis the data set is represented by events from four different event categories. First and foremost we have the signal events. The background has been separated into three different classes: Prompt J/ψ production with a very short lifetime; the "unrelated K_S^0"

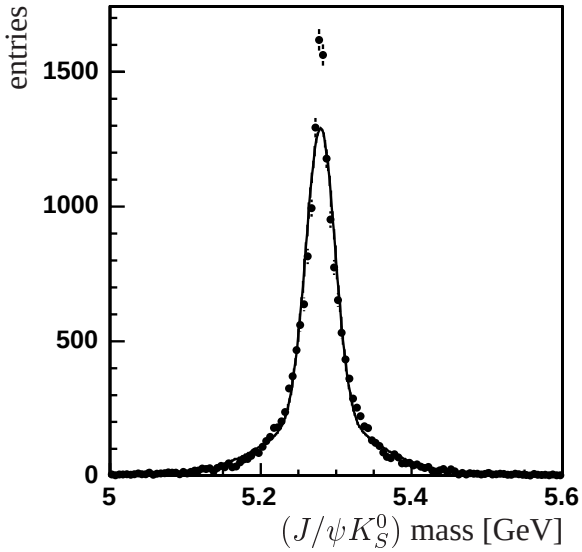
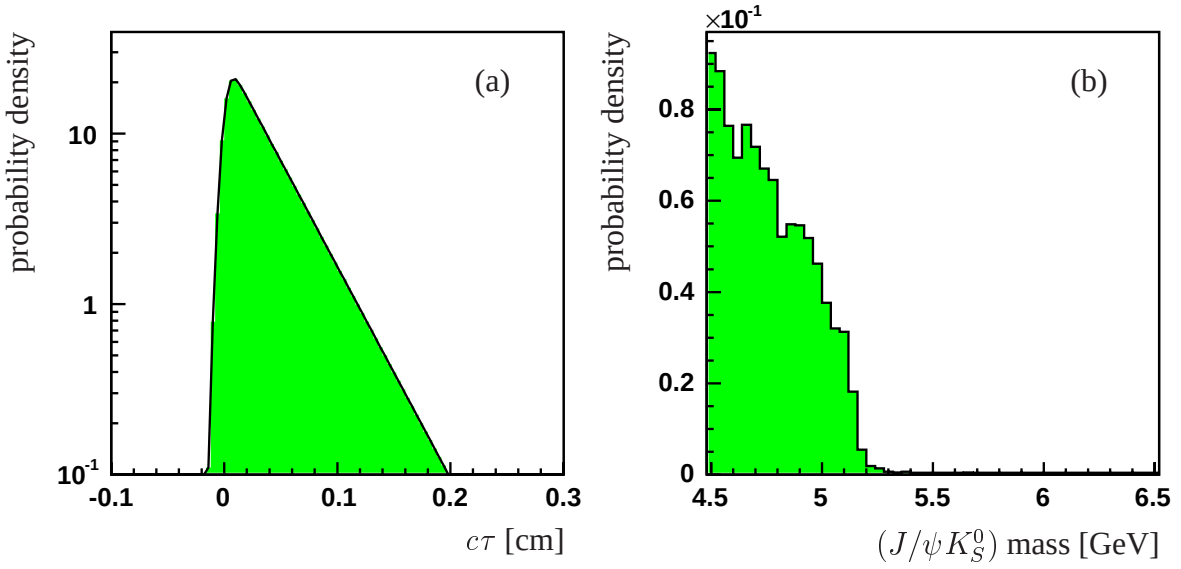


Figure 5.17: The $(J/\psi K_S^0)$ invariant mass spectrum of $B_d^0 \to J/\psi K_S^0$ decays in Monte Carlo, shown as points with errors. We also show a double Gaussian fit to the spectrum. The spectrum was obtained by smearing the invariant mass spectrum of the generator level J/ψ and K_S^0 four-vectors. The smearing procedure was devised to reproduce the reconstruction resolution. The spectrum shown here can therefore be compared to the mass peak shown in Fig. 4.28 (b).

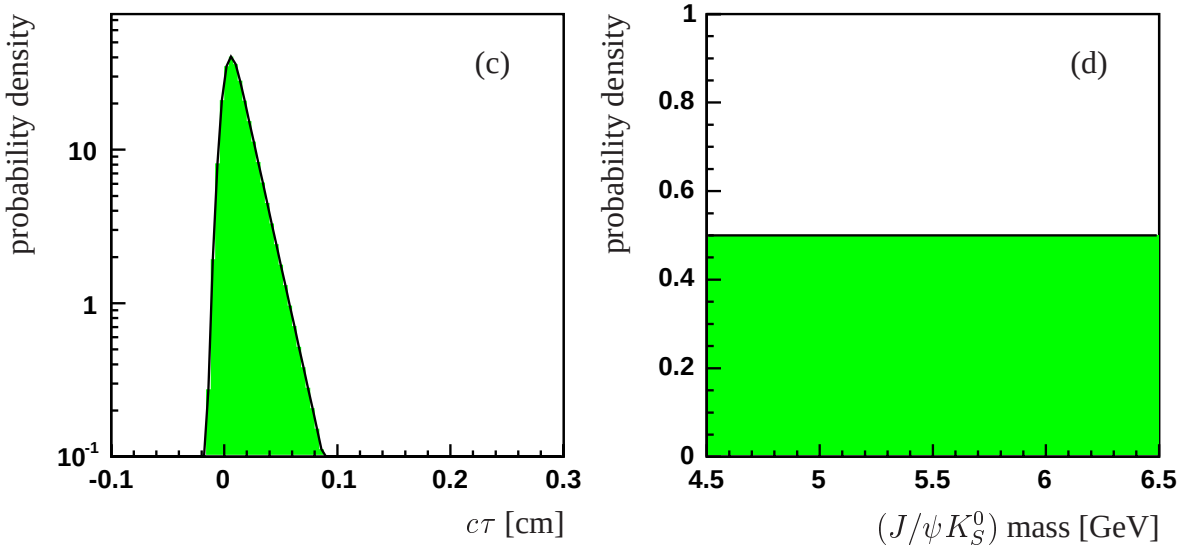
category characterized by a short lifetime and a uniform distribution in mass; and the "missing decay products" category, characterized by a long lifetime, and a mass below the B mass. For all event classes we defined a PDF in mass and in proper time, resulting in thirteen free parameters:

- six describing the signal:
  - the signal fraction α,
  - the B_d^0 mass, $m_{B_d^0}$, widths $\sigma_{B,1}$ and $\sigma_{B,2}$, and weight of the second width $\alpha_{B,2}$,
  - the B_d^0 lifetime $\lambda(B_d^0)$;
- four describing the prompt J/ψ production:
  - the fraction of all background, α_{prompt}, that is prompt J/ψ,
  - the $c\tau$ resolution scales $F_{c\tau,1}$ and $F_{c\tau,2}$, and α_2 which is the fraction of events following $F_{c\tau,2}$;
- two describing the "unrelated K_S^0" event category:
  - the fraction of long-lived background that is "unrelated K_S^0", $\alpha_{unrel.K}$,
  - the effective lifetime $(c\tau)_{unrel.K}$;
- one describing the "missing decay products" event category:



(a) Proper time PDF for the "missing decay products" contribution.

(b) Mass PDF for the "missing decay products" contribution.



(c) Proper time PDF for the "unrelated K_S^0" contribution.

(d) Mass PDF for the "unrelated K_S^0" contribution.

Figure 5.18: PDF's describing the long-lived background contributions to the data.

– the effective lifetime $(c\tau)_{miss}$.

Five of these parameters can be constrained as follows. We first fix the mass of the B_d^0 to the PDG value of 5.2794 GeV [11], because this has better accuracy than can be achieved in our fit. From the lifetime fit to the low-mass sideband data (see Fig. 5.10), we are able to obtain the shape of the prompt J/ψ $c\tau$ distribution. The results from the fit are listed in Table 5.1: $F_{c\tau,1}$ and $F_{c\tau,2}$ we set to 1.65 and 5.49, and α_2 to 0.05. In the high-mass sideband, the only long-lived contribution is from "unrelated K_S^0". We can therefore fix $(c\tau)_{unrel.K}$ to the slope of the exponential tail determined in a fit to the high-mass sideband data, 169 μm (see Table 5.2). This leaves eight free fit parameters.

5.4.4 Result of the two-dimensional fit

In order to include a reasonable number of events from all event categories, we apply a mass window from 4.5 GeV to 6.5 GeV. A tighter window would reduce the number of events in either the "missing decay products" event category, or the number of events in the "unrelated K_S^0" event class. Too few events in an event class leads to an inaccurate determination of the free fit parameters describing the event class ($(c\tau)_{miss}$ or $\alpha_{unrel.K}$), which adversely affects the accuracy on the B_d^0 lifetime measurement. On the other hand, we cannot widen our window too much, because our description of the different event categories is only valid over a limited mass range. For instance, sliding the lower limit of our mass range to a value smaller than the sum of the mass of the J/ψ and K_S^0 would invalidate the uniform mass distribution of the prompt J/ψ's and unrelated K_S^0's. No proper time window cut has been applied.

The fit result is shown in Fig. 5.19 and Table 5.3. The measured value of the B_d^0 lifetime is 412 μm. The uncertainty on this measurement is determined from the variation of the likelihood as a function of $\lambda(B_d^0)$, as shown in Fig. 5.20. We define the statistical uncertainty as the deviation from the optimal value of $\lambda(B_d^0)$, that causes the natural logarithm of the likelihood, $\log \mathcal{L}$, to increase by 0.5. The statistical uncertainty on our result is 72 μm in the positive direction, and 59 μm in the negative direction.

Figure 5.19 (b) shows that the distribution in proper time of the signal and the background contribution from events of the "missing decay products" category is very similar. It is precisely the strength of the two-dimensional fit that we developed that allows us to disentangle these two contributions, because each is determined by a different mass region. This can also be seen from the correlation matrix of the fit, shown in Table 5.4. The statement that we have disentangled $\lambda(B_d^0)$ and $(c\tau)_{miss}$ is confirmed by their small correlation coefficient of 0.018. The correlation between the signal fraction α and $\lambda(B_d^0)$ is stronger, with a correlation coefficient of -0.33. This can be understood as follows. The most important background contribution in the B_d^0 mass region is from prompt J/ψ production, which has no significant lifetime. If $\lambda(B_d^0)$ is increased, the total number of signal events, corresponding to α, must be reduced, to match the number of events observed at high lifetime.

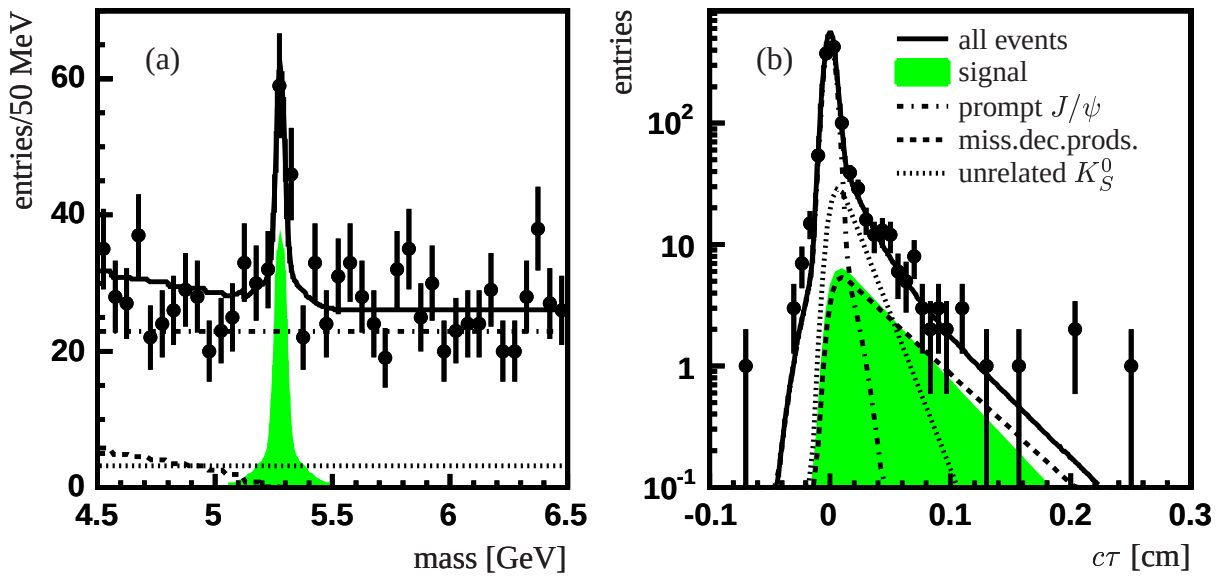


Figure 5.19: Result from the two-dimensional fit, using mass and proper time information.

Variable	value
signal fraction α	0.044 ± 0.0122
$m(B_d^0)$	5.2794 GeV (fixed)
B_d^0 width 1	18.6 ± 8.30 MeV
B_d^0 width 2	79.8 ± 75.03 MeV
$\alpha_{B,2}$	0.43 ± 0.255 MeV
$\lambda(B_d^0)$	412^{+72}_{-59} μm
lifetime $(c\tau)_{miss}$	478 ± 114 μm
fraction unrelated K_S^0's $\alpha_{unrel.K^0}$	0.71 ± 0.09
lifetime $(c\tau)_{unrel.K^0}$	169 μm (fixed)
prompt fraction α_{prompt}	0.84 ± 0.02
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.65 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	5.49 (fixed)
weight second scale α_2	0.05 (fixed)

Table 5.3: Result from the two-dimensional fit, using mass and proper time information.

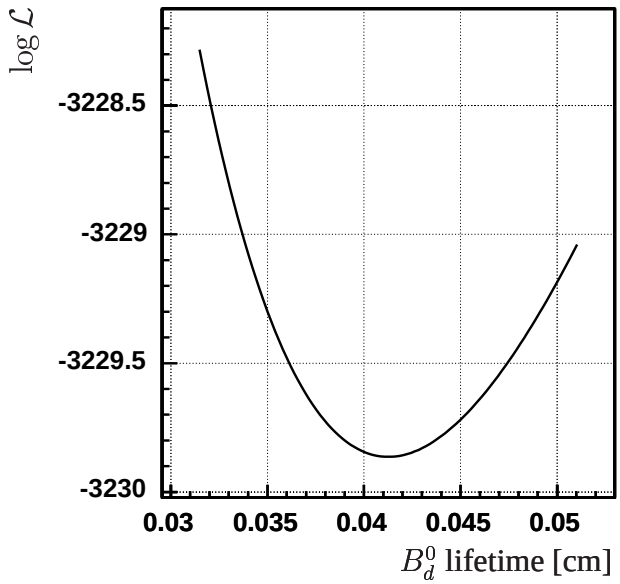


Figure 5.20: The logarithm of the likelihood $\mathcal{L}$ as a function of the B_d^0 lifetime, $\lambda(B_d^0)$. A 1σ statistical deviation is defined as the deviation in $\lambda(B_d^0)$ at which the likelihood is 0.5 above its minimum.

	α	$\sigma_{B,1}$	$\sigma_{B,2}$	$\alpha_{B,2}$	$\lambda(B_d^0)$	$(c\tau)_{miss}$	$\alpha_{unrel.K}$	α_{prompt}
α	1	0.074	0.19	0.15	-0.33	-0.045	-0.17	0.19
$\sigma_{B,1}$	0.074	1	0.36	-0.56	-0.054	-0.0064	-0.029	0.036
$\sigma_{B,2}$	0.19	0.36	1	-0.44	-0.048	-0.0082	-0.054	0.11
$\alpha_{B,2}$	0.15	-0.56	-0.44	1	-0.071	-0.019	-0.073	0.079
$\lambda(B_d^0)$	-0.33	-0.054	-0.048	-0.071	1	0.018	0.085	-0.077
$c\tau_{miss}$	-0.045	-0.0064	-0.0082	-0.019	0.018	1	0.37	-0.016
$\alpha_{unrel.K}$	-0.17	-0.029	-0.054	-0.073	0.085	0.37	1	-0.34
α_{prompt}	0.19	0.036	0.11	0.079	-0.077	-0.016	-0.34	1

Table 5.4: The correlation matrix of the two-dimensional fit shown in Fig. 5.19 and Table 5.3.

5.5 Monte Carlo verification

As a consistency check, we have tested our lifetime fit on a Monte Carlo sample of similar composition as the data. We created this sample by merging our $B \to J/\psi X$ and prompt J/ψ MC samples. We removed the $B_d^0 \to J/\psi K_S^0$ decays from the $B \to J/\psi X$ sample, because due to technical issues with the generation, the proper time distribution of these decays does not follow an exponential shape, as explained in section A.1. We introduced $B_d^0 \to J/\psi K_S^0$ signal by adding 100 reconstructed events from our signal MC sample, which has been generated such that the proper time distribution of the B_d^0's does follow an exponential. Note that this is more than the ~ 50 reconstructed B_d^0's that we have in data, so we expect a $\lambda(B_d^0)$ result from a two-dimensional fit to this MC sample that is more accurate than our nominal result.

The proper time resolution and hence the resolution scale factors $F_{c\tau,1}$ and $F_{c\tau,2}$ can differ between data and Monte Carlo, so the fits determining these parameters, as done for data in Figs. 5.10 and 5.11, have to be redone for this sample. The fits for this Monte Carlo sample are shown in Figs. 5.21 and 5.22. The fit to the low-mass sideband in Fig. 5.21 determines $F_{c\tau,1}$ to be 1.01 ± 0.186 and $F_{c\tau,2}$ to be 2.15 ± 0.757, with α_2 at 0.24 ± 0.23 (see Table 5.5). The value of α_{prompt} results from the relative size of the $B \to J/\psi X$ MC sample and the prompt J/ψ sample. These parameters have been fixed in the fit to the high-mass sideband data from this sample, to enable the determination of $(c\tau)_{unrel.K}$. As indicated in Table 5.6, $(c\tau)_{unrel.K}$ is determined in the fit in Fig. 5.22 to be 274^{+73}_{-52} μm.

We use these parameters in the two-dimensional fit to this sample, the result of which is shown in Fig. 5.23 and Table 5.7. The fitted value of $\lambda(B_d^0)$ is 462^{+44}_{-39} μm, which is consistent with the value of 464 μm, that has been used in the generation. The errors are, as expected, smaller than in our nominal result of 412^{+72}_{-59} μm, where fewer reconstructed $B_d^0 \to J/\psi K_S^0$ decays were available.

5.6 Systematic uncertainties

In this section we will study different effects that potentially affect the lifetime measurement. These effects can be introduced by the event reconstruction algorithm, by the event selection, by our lifetime measurement method, or by detector effects.

We have proven in sections 3.4 and 3.5 that no significant bias is introduced in the L_{xy} and p_T measurement by the reconstruction algorithm on Monte Carlo data. In the event selection, we have taken care not to introduce any cuts affecting the shape of the proper time spectrum. As a result, no significant change in efficiency is observed as a function of proper time when looking at Monte Carlo data. However, we can only exclude efficiency effects to a certain level, and this introduces some uncertainty on our result, as described in section 5.6.1. Moreover, we have to verify that cuts on the vertex quality do not introduce a bias, since these cuts are sensitive to possible differences between Monte Carlo and data, such as alignment. We have studied this possibility in section 5.6.2.

Some uncertainty on our result is introduced because a number of parameters have been fixed in the two-dimensional fit, as they have been determined with a certain accuracy. We have

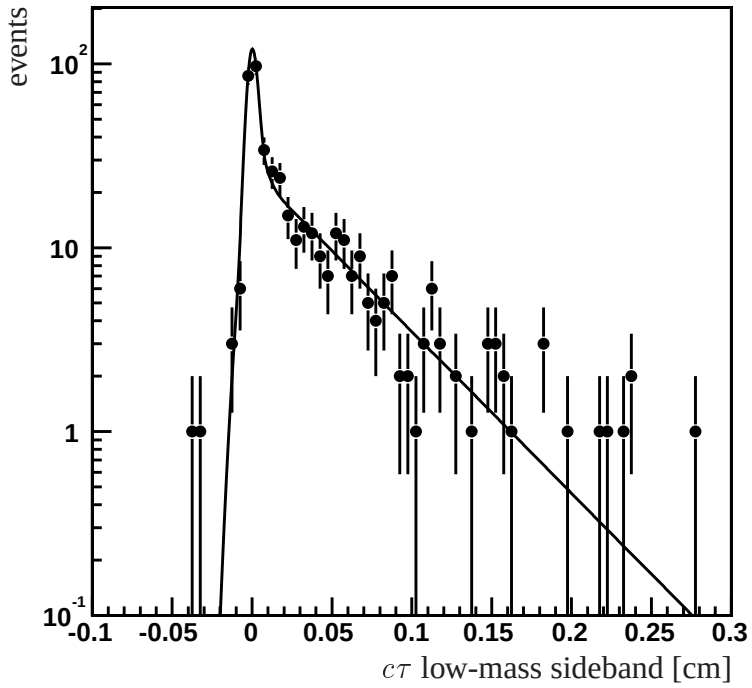


Figure 5.21: A fit with a double Gaussian and exponential to the proper time distribution in a Monte Carlo sample, in the low-mass sideband (mass range 4.55 – 5.1 GeV). The χ^2_{fit} per degree of freedom is 0.7.

Variable	value
lifetime $(c\tau)_{LL}$	494^{+33}_{-30} μm
prompt fraction α_{prompt}	0.41 ± 0.04
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.01 ± 0.186
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	2.15 ± 0.757
weight second scale α_2	0.24 ± 0.23

Table 5.5: Results from fitting a double Gaussian and an exponential convoluted with a Gaussian to events from the low-mass sideband in a Monte Carlo sample containing $B \to J/\psi X$ and prompt J/ψ decays.

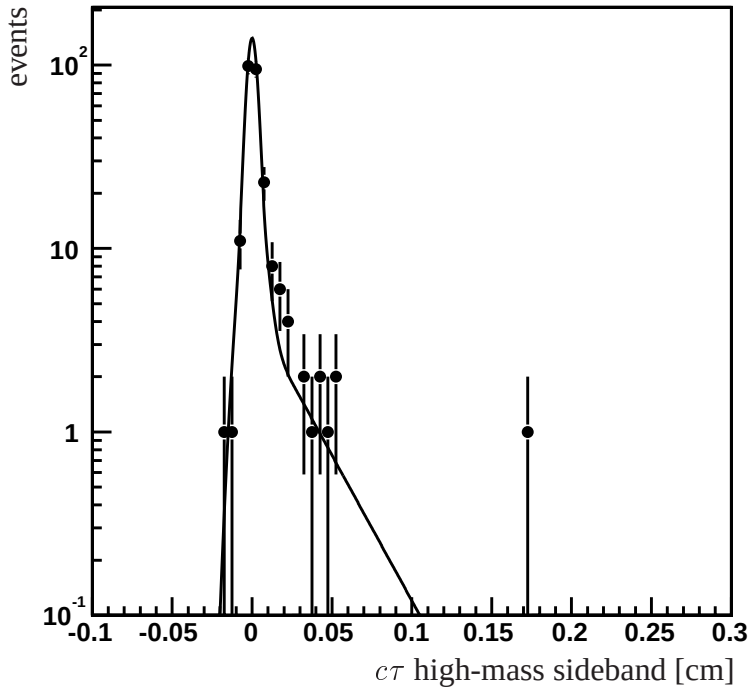


Figure 5.22: A fit with a double Gaussian and exponential to the proper time distribution in a Monte Carlo sample, in the high-mass sideband (mass range 5.6 – 6.6 GeV). The χ^2_{fit} per degree of freedom is 0.2.

Variable	value
lifetime $(c\tau)_{unrel.K}$	274^{+73}_{-52} μm
prompt fraction α_{prompt}	0.90 ± 0.04
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.01 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	2.15 (fixed)
weight second scale α_2	0.24 (fixed)

Table 5.6: Results from fitting a double Gaussian and an exponential convoluted with a Gaussian to events from the high-mass sideband in a Monte Carlo sample containing $B \to J/\psi X$ and prompt J/ψ decays. The parameters of the prompt J/ψ PDF have been fixed to the values obtained from the low-mass sideband (see Table 5.5).

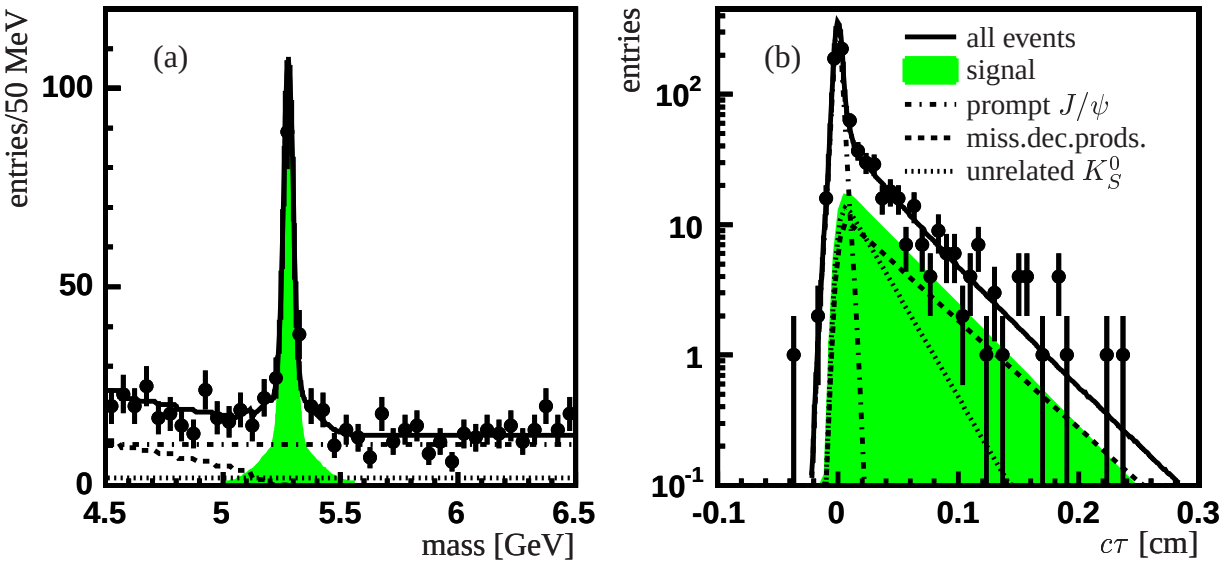


Figure 5.23: Result from a two-dimensional fit, using a Monte Carlo sample composed of prompt J/ψ, $B \to J/\psi X$ and $B^0_d \to J/\psi K^0_S$ decays.

Variable	value
signal fraction α	0.195 ± 0.0249
$m(B^0_d)$	5.2794 GeV (fixed)
B^0_d width 1	20.4 ± 4.42 MeV
B^0_d width 2	100.0 ± 89.24 MeV
$\alpha_{B,2}$	0.39 ± 0.121 MeV
$\lambda(B^0_d)$	462^{+44}_{-39} μm
lifetime $(c\tau)_{miss}$	525 ± 87 μm
fraction unrelated K^0_S's $\alpha_{unrel.K^0}$	0.43 ± 0.10
lifetime $(c\tau)_{unrel.K^0}$	274 μm (fixed)
prompt fraction α_{prompt}	0.71 ± 0.03
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.01 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	2.15 (fixed)
weight second scale α_2	0.24 (fixed)

Table 5.7: Result from a two-dimensional fit, using a Monte Carlo sample composed of prompt J/ψ, $B \to J/\psi X$ and $B^0_d \to J/\psi K^0_S$ decays.

estimated this uncertainty in section 5.6.3. Finally, we consider the possibility of inaccuracies in our background model in section 5.6.4.

The remaining effects are detector effects. The only conceivable relevant effect on our measurement comes from alignment. Misalignment of tracking detectors degrades the p_T resolution. This is covered by the two resolution scale factors, $F_{c\tau,1}$ and $F_{c\tau,2}$, which are determined from data. The other effect from misalignment occurs if the radius of detector layers is off. We have studied this effect in section 5.6.5.

We will conclude with an estimate of the combined contribution of these effects to the uncertainty on our result.

5.6.1 Selection efficiency

It has already been noted in section 5.1 that a B_d^0 selection efficiency that changes with proper time has a strong effect on the B_d^0 lifetime measurement. The efficiency was found to be constant within errors. The uncertainty on the dependence of the efficiency as a function of proper time, however, needs to be propagated. The straight-line fit to the efficiency as a function of MC true $c\tau$, in Fig. 5.1 (b), gave a slope of $(-0.98 \pm 2.16) \cdot 10^{-2}$. We introduce the effect of a varying selection efficiency by multiplying our PDF in proper time for signal from equation 5.7 with the straight line efficiency function f_{eff}:

$$f_{sig}^{eff} = C_{sig}^{eff} \cdot \left[\exp\left(-\frac{(c\tau)_i}{\lambda(B_d^0)} \right) \otimes \exp\left(-\frac{1}{2} \left(\frac{(c\tau)_i}{\sigma_w} \right)^2 \right) \right] \cdot f_{eff}, \tag{5.18}$$

where we must be careful to appropriately adjust the constant C_{sig}^{eff} to keep the PDF normalized.

Our nominal result for the B_d^0 lifetime is 412^{+72}_{-59} μm. We will now use f_{sig}^{eff} as the PDF in time for signal in our fit. We first choose the slope of f_{sig}^{eff} to be $(-0.98 + 2.16) \cdot 10^{-2}$, and perform the two-dimensional fit, yielding a B_d^0 lifetime of 424 μm, 12 μm higher than our nominal result. We then choose the slope to be $(-0.98 - 2.16) \cdot 10^{-2}$, which yields a B_d^0 lifetime of 387 μm, 25 μm below our nominal result. We therefore assign a 12 μm positive systematic error to our result, and a negative error of 25 μm.

5.6.2 Vertex quality cuts

To study how the vertex quality cuts that we make on the B_d^0 candidates affect the lifetime measurement, we have relaxed those cuts and fit the B_d^0 lifetime using the resulting candidate sample. The cuts that affect vertex quality are the B_d^0 vertex χ^2, and also the number of SMT hits required on each track from the J/ψ. The optimal values of these cuts are 25 for the B_d^0 vertex χ^2 and a minimum of 3 SMT hits. We have relaxed these to 100 and 1, respectively. The number of events increases from 1139 to 1224. The result of the two-dimensional lifetime fit is shown in Fig. 5.24 and Table 5.8. The B_d^0 lifetime measured using these modified cuts is 401^{+67}_{-55} μm.

At least part of the difference with the nominal result (-11 μm) can be attributed to statistics, because there are more events in this fit. We quantify the statistical effect as follows. The number

of events used in the nominal fit is 1139. Relaxing the vertexing cuts increases this to 1224. We now use the 1224 events and randomly select 1139 events from this sample. The resulting event sample, containing 1139 events selected with looser vertexing cuts, will be used for a two-dimensional fit. If we repeat this procedure 10 times, we get 10 different answers. The spread on the answers is an indication of the statistical effect of the expansion of the event sample from 1139 to 1224 events. In our case, the RMS of the answers is 19 μm. The difference between the nominal result and the result with relaxed vertexing cuts was 11 μm, so we conclude that this difference is consistent with a statistical effect.

5.6.3 Parameters fixed in the two-dimensional fit

A number of parameters has been determined with a certain accuracy, and later fixed in the two-dimensional fit. The uncertainty on these parameters has to be propagated to the final result separately. We do this by varying each parameter by $\pm 1\sigma$ and combining the variations in $\lambda(B_d^0)$ as independent effects.

The parameters that have been fixed in the two-dimensional fit are the B_d^0 mass, $(c\tau)_{unrel.K}$, and the parameters of the prompt J/ψ proper time distribution, $F_{c\tau,1}$, $F_{c\tau,2}$ and α_2. The uncertainty on the B_d^0 mass is the uncertainty on the PDG value, 0.5 MeV. A variation of the B_d^0 mass by this amount has no measurable effect on $\lambda(B_d^0)$. We determined the parameters of the prompt J/ψ proper time distribution in a fit to the low-mass sideband (see Table 5.1). The parameter $(c\tau)_{unrel.K}$ was found using a fit to the high-mass sideband data (see Table 5.2). We have tabulated the measured values again in Table 5.9. We have also indicated the deviation in the measured value of $\lambda(B_d^0)$, when each parameter is decreased or increased by one standard deviation, keeping the other parameters at their nominal values.

We collect all negative variations of $\lambda(B_d^0)$, and find the combined negative uncertainty by squared addition: 6.0 μm. Using the same method we find the positive uncertainty to be 4.5 μm. For simplicity, we assign a symmetric uncertainty of 6 μm to this effect.

5.6.4 Background model

An incorrect model of the background may affect the lifetime measurement. A possible modification to our background model is the introduction of a slope in the mass distribution of the "unrelated K_S^0" contribution. One may expect this mass distribution to fall toward higher masses, because relatively higher-p_T K_S^0's are needed, and the K_S^0 p_T spectrum is a falling distribution. On the other hand it is expected that, if there is a slope, the slope is small (see Fig. 5.15 (a)). To allow a slope C in the mass distribution, we use the following normalized probability density function:

$$f_{mass} = \frac{1 + C \cdot (m_i - m_{min})}{\int_{m_{min}}^{m_{max}} \left[1 + C \cdot (m_i - m_{min})\right] dm_i} \tag{5.19}$$

Here m_{min} and m_{max} are the boundaries of the mass window, and m_i is the mass of the B_d^0 candidate. For the slope of the "unrelated K_S^0" mass distribution we find $C = -0.07 \pm 0.42$ GeV^{-1},

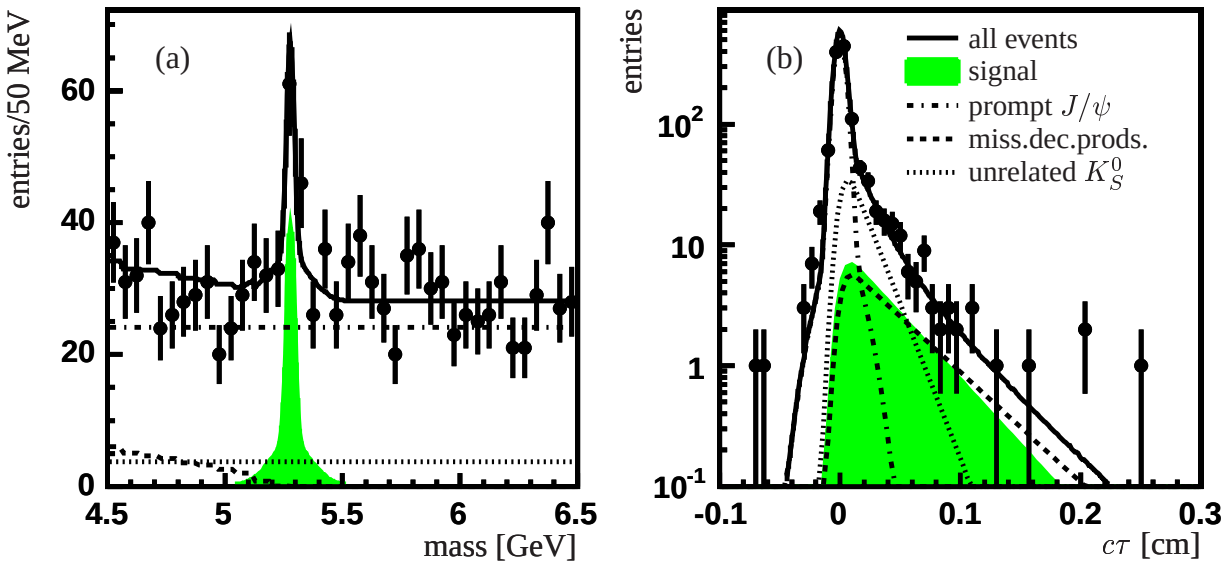


Figure 5.24: Result from a two-dimensional fit, where we relaxed the cuts on the B_d^0 vertex χ^2 and the minimum number of SMT hits on the tracks from the J/ψ.

Variable	value
signal fraction α	0.045 ± 0.0124
$m(B_d^0)$	5.2794 GeV (fixed)
B_d^0 width 1	17.0 ± 7.37 MeV
B_d^0 width 2	83.7 ± 75.08 MeV
$\alpha_{B,2}$	0.49 ± 0.226 MeV
$\lambda(B_d^0)$	401^{+67}_{-55} μm
lifetime $(c\tau)_{miss}$	474 ± 112 μm
fraction unrelated K_S^0's $\alpha_{unrel.K^0}$	0.74 ± 0.08
lifetime $(c\tau)_{unrel.K^0}$	169 μm (fixed)
prompt fraction α_{prompt}	0.83 ± 0.02
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.65 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	5.49 (fixed)
weight second scale α_2	0.05 (fixed)

Table 5.8: Result from a two-dimensional fit, using mass and proper time information. The cut on the B_d^0 vertex χ^2 has been relaxed to 100, and only a minimum of 1 SMT hit was required on each track from the J/ψ.

parameter	measured value	variation of $\lambda(B_d^0)$ if decreased by 1σ [μm]	variation of $\lambda(B_d^0)$ if increased by 1σ [μm]
$(c\tau)_{unrel.K}$	169^{+34}_{-27} μm	+0	-3
$F_{c\tau,1}$	1.65 ± 0.15	-3	+2
$F_{c\tau,2}$	5.49 ± 2.731	+3	-4
α_2	0.05 ± 0.06	+3	-2

Table 5.9: Variation of $\lambda(B_d^0)$, as a result of $\pm 1\sigma$ variations in parameters that were fixed in the two-dimensional fit (see Table 5.3). The nominal value of $\lambda(B_d^0)$ is 412^{+72}_{-59} μm.

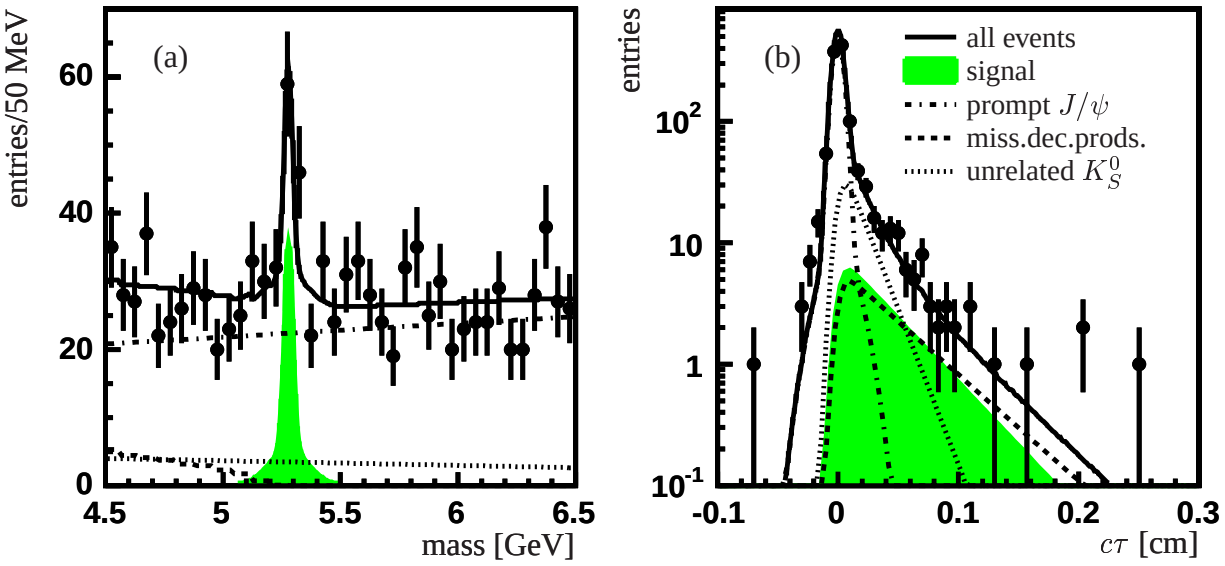


Figure 5.25: Result from a two-dimensional fit, where we introduced a slope in the mass distributions of the prompt J/ψ and "unrelated K_S^0" contributions.

which is consistent with zero. The B_d^0 lifetime measured this way is 414^{+74}_{-60} μm. Similarly, we can allow a slope in the prompt J/ψ mass distribution. The fit determines this slope to be 0.09 ± 0.10 GeV^{-1} and the B_d^0 lifetime at 409^{+72}_{-58} μm. We can also allow both slopes simultaneously to be different from zero. This yields a slope of 0.10 ± 0.11 GeV^{-1} and -0.17 ± 0.32 GeV^{-1} for the prompt J/ψ and unrelated K_S^0 mass distribution, respectively. The B_d^0 lifetime result is 414^{+74}_{-58} μm. This result is shown in Fig. 5.25.

We now adjust the mass window for the two-dimensional fit to 5.1 – 6.5 GeV, from 4.5 – 6.5 GeV. According to the $(J/\psi K_S^0)$ mass spectrum of events from the missing decay products category, shown in Fig. 5.16 (b), only a very small contribution from this event class will be left in our mass window. We therefore exclude this contribution from the fit by setting $\alpha_{unrel.K}$ to 1. The result, shown in Fig. 5.26 and Table 5.10, is a B_d^0 lifetime of 416^{+72}_{-59} μm, which is consistent with our nominal result of 412^{+72}_{-59} μm.

Finally, we consider the possibility of an additional background contribution that we failed

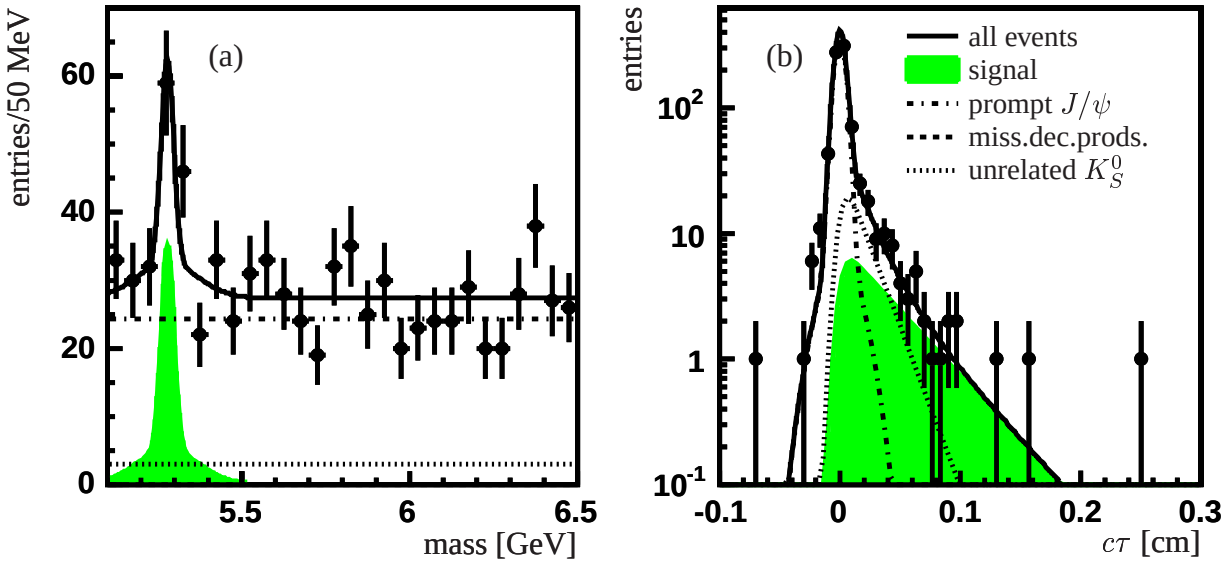


Figure 5.26: Result from a two-dimensional fit, where we adjusted the mass window to 5.1 – 6.5 GeV and removed the "missing decay products" contribution from the fit.

Variable	value
signal fraction α	0.062 ± 0.0150
$m(B_d^0)$	5.2974 GeV (fixed)
B_d^0 width 1	25.4 ± 11.27 MeV
B_d^0 width 2	73.5 ± 27.88 MeV
$\alpha_{B,2}$	0.51 ± 0.235 MeV
$\lambda(B_d^0)$	416^{+72}_{-59} μm
fraction unrelated K_S^0's $\alpha_{unrel.K^0}$	1.00 (fixed)
lifetime $(c\tau)_{unrel.K^0}$	169 μm (fixed)
prompt fraction α_{prompt}	0.89 ± 0.03
$c\tau$ resol. scale factor 1 $F_{c\tau,1}$	1.65 (fixed)
$c\tau$ resol. scale factor 2 $F_{c\tau,2}$	5.49 (fixed)
weight second scale α_2	0.05 (fixed)

Table 5.10: Result from a two-dimensional fit, where we adjusted the mass window to 5.1 – 6.5 GeV and removed the "missing decay products" contribution from the fit.

to identify. It's difficult to conceive of a background that is significant, but does not show up in any of the fits, but we can simply introduce it in the two-dimensional fit and see how our result is affected.

We have implemented a background X, which we gave a normalization α_X defined as a fraction of the "missing decay products" contribution. As a mass distribution, we use a uniform distribution, just as was used for the unrelated K^0_S's. As the proper time shape, we can use an exponential convoluted with a Gaussian. If we do this, the unidentified background has exactly the same characteristics as the unrelated K^0_S's, and the fit will give X a lifetime and normalization similar to the unrelated K^0_S's and a large correlation with the parameters of the unrelated K^0_S contribution. We can also assume that our background is completely random in lifetime, i.e. that it has a uniform lifetime distribution. If we do this, we fit a normalization α_X of 0.04 ± 0.06 or 2 ± 3 events. This is consistent with zero events. The B^0_d lifetime is marginally affected: The result is 407^{+73}_{-60} μm.

We now have six results, obtained with different background models. We will assign the root mean square (RMS) of these results as the systematic error due to uncertainty in background model. The six results are:

- 412 μm (nominal result);
- the three results from fits where different slopes were allowed: 414 μm, 409 μm, 414 μm;
- 416 μm, obtained by excluding the missing decay products event category;
- 407 μm, after allowing an additional background in our fit.

The RMS of these results (given by $\sqrt{< x^2 > - < x >^2}$) is 3.1 μm. We assign a systematic uncertainty of 3 μm to the possibility that our background model is incorrect.

5.6.5 Radial alignment

The hits on SMT detector layers are the most important inputs to the decay length measurement, due to the single-hit resolution in these layers, and the fact that they are nearest to the decay vertex. If the radius of each of these layers is off by a certain distance, the decay length measurement will consistently be wrong by this amount. From a detector survey, the actual position of each silicon sensor matches the design position to within 40 μm. However, to produce an overall increase in the decay length measurement, an average shift of the sensors in each layer is required. We consider two effects:

- a common scale factor: $R_i \rightarrow (1 + \epsilon_1) \cdot R_i$;
- a common offset: $R_i \rightarrow R_i + \epsilon_2$;

where R_i is the radius of each SMT layer as used in the reconstruction, and ϵ_i gives the magnitude of each effect. Both effects affect the L_{xy} measurement linearly.

A scale factor will also affect the p_T measurement of the tracks. If the detector is larger than designed, the radius of curvature of the tracks will appear smaller, leading to an underestimate of the actual p_T. This is reflected in the invariant mass of the J/ψ and K_S^0 in data. These masses were shown to be off by $\sim 1\%$ in sections 4.3 and 4.4. If we attribute this entirely to the p_T measurement, this constrains ϵ_1 to values smaller than 1%. The p_T measurement, however, is calibrated by the mass-constrained fit. As explained in section 3.3.1, we have scaled the p_T of the tracks, so that the reconstructed mass matches the nominal J/ψ or K_S^0 mass. The p_T of the tracks is used to calculate the invariant mass of the B_d^0. In the $(J/\psi K_S^0)$ mass spectrum in Fig. 4.28 (a), the B_d^0 mass peak is at the correct mass, verifying that the p_T measurement has been calibrated correctly. The scale factor will also lead to an underestimate of the L_{xy} measurement by a factor $(1+\epsilon_1)$, though. From equation 5.1 follows that a 1% error in the L_{xy} measurement leads to the same error in the $c\tau$ measurement. Therefore we assign a systematic uncertainty of 1%, 4 μm, to a possible scale factor.

A radial layer offset by ϵ_2 does not affect the p_T measurement, but it does lead to an offset in L_{xy}. The measured decay length is $L_{xy} - \epsilon_2$. To study the effect, we introduced ϵ_2 as a free fit parameter. We find a negligible offset of -0.11 ± 0.4 μm. Our lifetime measurement is not affected by this shift.

5.6.6 Summary of systematic effects

We have studied different possible effects on our B_d^0 lifetime measurement, and assigned the following uncertainties to our result:

- Selection efficiency: +12 / -25 μm
- Fixed fit parameters: $\pm$ 6 μm
- Background model: $\pm$ 3 μm
- Radial alignment: $\pm$ 4 μm

Adding these uncertainties quadratically as independent effects yields an overall systematic error toward higher values of 14 μm and toward lower values of 26 μm.

Chapter 6

Conclusions

In this thesis, we have used data from the DØ experiment, taken between October 2002 and June 2003, to make a measurement of the B_d^0 lifetime in decays to $(J/\psi K_S^0)$. We have developed an event selection, studied the relevant backgrounds, defined and implemented a fit method to extract the lifetime, and finally, investigated the systematical uncertainties. The result of this work is a measurement of the B_d^0 lifetime of:

$$\tau = 1.38\ ^{+0.26}_{-0.20}\ (\mathrm{stat.})\ ^{+0.05}_{-0.09}\ (\mathrm{syst.})\ \mathrm{ps}. \tag{6.1}$$

We have indicated both the statistical and the systematic uncertainty. This preliminary result has been presented by the DØ Collaboration at the Rencontres de Moriond 2004 conference on Electroweak Interactions and Unified Theories [60].

The current world average of the B_d^0 lifetime as reported for the Winter 2004 conferences by the Heavy Flavor Averaging Group [23] is:

$$\tau_{WA} = 1.536 \pm 0.014\ \mathrm{ps}. \tag{6.2}$$

The result presented in this thesis is consistent with the world average. The main reason for the better accuracy of the world average is that the world average is an average over multiple decay modes, and multiple experiments. Even though higher-statistics decay modes are available, we have performed a B_d^0 lifetime measurement in decays to $(J/\psi K_S^0)$, because this is an important benchmark for a $\sin(2\beta)$ measurement at DØ. A $\sin(2\beta)$ measurement using $B_d^0 \to J/\psi K_S^0$ will require the same event selection and vertexing, as were developed in this thesis. The consistency of our result with the world average is an important first step.

It is clear that the limiting factor on the accuracy of our result is the signal statistics. The availabe statistics also limits our ability to uncover systematic effects in the data to reduce the systematic error. For a measurement of $\sin(2\beta)$ it is important that the systematic uncertainties in the measurement presented in this thesis are under control. Even with our limited dataset, we have shown the systematics to be understood at the 5% level.

Another important issue for a $\sin(2\beta)$ measurement is the level and nature of the background under the mass peak. A significant long-lived background in this region will severely complicate the measurement. Figure 6.1 shows the proper time distribution of events from the B_d^0 mass peak

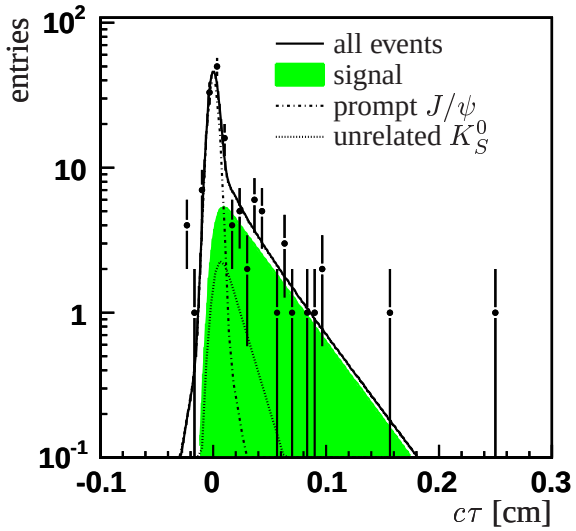


Figure 6.1: Lifetime distribution of events from the B^0_d mass peak region (5.20 – 5.36 GeV). The fit result shown corresponds to the nominal result, shown in Table 5.3.

region (5.20 – 5.36 GeV). The figure also shows the nominal fit result, with the normalization of the different contributions adjusted to their relative weight in the selected mass region. The missing decay products contribution is negligible, and therefore not drawn. The most prominent background, the prompt J/ψ's, has no significant lifetime. Furthermore, the "unrelated K^0_S" contribution becomes insignificant at proper decay lengths greater than approximately 0.5 mm. Beyond this value, the data is completely dominated by signal.

6.1 Outlook

From September 8 to November 17, 2003, the Tevatron was shut down to perform a broad range of projects, aimed at improving the instantaneous luminosity. It appears that the improvements that have been made have had a highly beneficial effect on the Tevatron performance. In the months of October to December 2002, the typical luminosity delivered to DØ was 5 pb^{-1} per week. In February 2004, the weekly delivered luminosity has typically been twice that number. This is conform the design projection, as presented to the Department of Energy during a review in July 2003 [61]. It is therefore not unrealistic to assume that also the projection for the integrated luminosity until the year 2006 will be met. According to the design projection, the integrated luminosity will exceed 2 fb^{-1} during the fiscal year (FY) 2006, and will reach 8.6 fb^{-1} by the end of FY 2009. It is currently foreseen that Run II will be terminated at that point, because of the expectation that, at that time, the Large Hadron Collider at CERN will be accumulating luminosity at a significantly higher rate.

The report "B-physics at the Tevatron – Run II and beyond" [7] presents detailed predictions of the accuracy of different B-physics measurements, assuming an integrated luminosity of 2 fb^{-1}. CDF projects a B^0_d lifetime measurement accuracy of 0.01 ps, and the expectation is to

measure the lifetime ratio $\tau(B_d^0)/\tau(B_s)$ to an accuracy of 0.5%. DØ can expect similar precision. The improvement of a factor of ~ 30 in the lifetime measurement, compared to the result presented in this thesis, is attributed to three factors: Luminosity, trigger efficiency (p_T threshold) and the decay channels considered.

Regarding the measurement of $\sin(2\beta)$, DØ expects a precision of 0.04 based on the decay mode $B_d^0 \to J/\psi K_S^0$ alone. This is comparable to the accuracy of the current world average of $\sin(2\beta)$, 0.047 [23]. To achieve this accuracy however, the flavor tagging at DØ will have to be improved. Currently, only the opposite-side muon B-flavor tag is available. A similar algorithm using electrons is foreseen, and algorithms using the charge of pions from the B fragmentation will also be developed.

The BTeV experiment, designed for B-physics measurements [62], is expected to start taking data at the Tevatron in 2009, initially at a rate of 2 fb^{-1} per year. Using the first 2 fb^{-1} of data, an accuracy on $\sin(2\beta)$ of 0.025 is foreseen. These measurements will pose more strict constraints on the Standard Model than are currently available.

In 2009, new competitors will have arrived at the scene, provided the LHC starts operation in 2007 [63]. The LHCb Collaboration has designed a dedicated B-physics experiment for the LHC. ATLAS and CMS are general-purpose experiments for the LHC, but will also make B-physics measurements. After one year of data-taking, the projected accuracies on $\sin(2\beta)$ are 0.021, 0.017 and 0.015 for LHCb, ATLAS and CMS respectively [64]. The instantaneous luminosity at the LHCb interaction point is lower than at the other interaction points, yielding a poorer accuracy on $\sin(2\beta)$ after one year of running than ATLAS and CMS. The proper time resolution of the LHCb experiment is superior, which eventually will lead to the most accurate measurement, but they will need more time to accumulate statistics.

After having listed the impending experimental advances, it appears that some of the theoretical predictions may be insufficiently accurate to fully benefit from them. Unfortunately, it is much more difficult to predict the improvement in theoretical precision, such as in HQET, or its successor. Moreover, even assuming that the existing theoretical predictions will improve, without experimental confirmation it is impossible to tell whether the improved calculations describe nature in sufficient detail. Experimental measurements will have to point the way.

Appendix A

Monte Carlo event samples

Three Monte Carlo samples are used throughout this thesis: A $B_d^0 \to J/\psi K_S^0$ (signal) sample, a $B \to J/\psi X$ sample, and a prompt J/ψ sample. In this appendix we explain the details of the generation of each sample, as well as the generation of the minimum bias events that have been added to the samples.

For all samples, PYTHIA [59] version 6.202 was used for the generation of events. For the B-decay samples, a modified version of the program called EvtGen [65] was used to simulate the B-decays. In these samples, PYTHIA has generated the full event, like it does by default, only the B-mesons were set to be stable. The decay of the B-mesons was then simulated by EvtGen.

The EvtGen program has been developed for the simulation of B-decays at colliders operating at the $\Upsilon(4S)$ resonance, such as the PEP-II collider at SLAC [20] and the KEKB collider at KEK [66]. The B-mesons at these colliders are predominantly produced by the decay of a single particle, and the two B-mesons produced by such decays are coherent quantum states. One of the effects of this coherence is that, if one of the B-mesons mixes to its anti-particle, at the same time the other B-meson mixes to its anti-particle as well. This coherence generally does not exist between B-meson pairs produced at a hadron collider, because it is disturbed by the uncorrelated quarks, that are drawn from the vacuum (or the proton) to form the mesons. The version of the EvtGen program used at DØ has therefore been modified to remove the coherence between the B-mesons.

After the event generation, the DØMess package (DØ Monte Carlo Event Selection System [67]) allows to make generator level cuts, that go beyond the selections that are offered by the generator or EvtGen. For example, DØMess allows one to require the presence of a J/ψ in the event, coming from a B_d^0 or $\bar{B}_d^0$ decay, decaying to two muons with a p_T greater than 1 GeV. These are requirements that cannot be implemented by the modification of the particle decay table of the generator or EvtGen.

The DØGSTAR program (DØ GEANT Simulation of the Total Apparatus Response) is used to simulate the effects of the material in the detector. The DØSim program simulates the digital signals in the detector, based on the DØGSTAR output. Finally, event reconstruction is performed by the DØreco program. For all DØ programs, DØGSTAR, DØSim and DØreco, version p14.02.00 was used. In DØreco, we enabled the "AA extended" tracking algorithm, the same algorithm that was used to reprocess the data.

In all samples minimum bias events were overlaid. The number of minimum bias events was randomly distributed following a Poisson distribution with an average of 0.8. We describe the details of the generation of the minimum bias events in section A.4.

A.1 $B^0_d \to J/\psi K^0_S$ signal MC sample

We used PYTHIA to generate $b\bar{b}$ quark pairs, using the selector `MSEL=5`. Ref. [59] has a detailed explanation of the generation process. Each quark was required to have a minimum p_T of 2 GeV.

EvtGen was used to simulate the decay of the B-mesons. A modification to EvtGen was necessary to generate the correct proper time distributions, in addition to the deletion of the coherence between the B-mesons described in the previous section. The low branching fraction of the B^0_d to $(J/\psi K^0_S)$ of 4.4×10^{-4} makes the CPU time required to generate a large number of events containing a $B^0_d \to J/\psi K^0_S$ decay excessive. The standard solution is to force all B^0_d's to decay to $(J/\psi K^0_S)$, by setting the branching fraction for this decay equal to 1, and leaving the $\bar{B}^0_d$'s to decay freely. This way we obtain an event sample, containing many decays of B^0_d to $(J/\psi K^0_S)$ and almost no decays of $\bar{B}^0_d$ to $(J/\psi K^0_S)$. Due to the CP violation in B^0_d mixing, the proper time distribution of the B^0_d's in decays to CP eigenstates like $(J/\psi K^0_S)$, is an exponential, modulated with a sine:

$$N(t) = N_0\, e^{-t/\lambda(B^0_d)} \cdot \big[1 - \sin(2\beta)\sin(\Delta m t)\big]. \tag{A.1}$$

Here $N(t)$ is the number of B^0_d's decaying to $(J/\psi K^0_S)$ when time t has passed since their production, and N_0 is a scale factor proportional to the number of B^0_d's produced at $t = 0$. β is the angle in the unitarity triangle shown in Fig. 1.1, and Δm is the mass difference between the B^0_d and $\bar{B}^0_d$ mass states. The proper time distribution of the $\bar{B}^0_d$ in the decay to $(J/\psi K^0_S)$ is modulated with the opposite sign:

$$\bar{N}(t) = \bar{N}_0\, e^{-t/\lambda(B^0_d)} \cdot [1 + \sin(2\beta)\sin(\Delta m t)], \tag{A.2}$$

with $\bar{N}(t)$ the number of $\bar{B}^0_d$'s remaining and $\bar{N}_0$ the scale factor proportional to the number of $\bar{B}^0_d$'s produced. The distributions of $N(t)$ and $\bar{N}(t)$ are shown in Fig. A.1. It should be noted that due to the CP violating term, $\sin(2\beta)\sin(\Delta m t)$, the integral of the two distributions is not the same (in the case of $N_0 = \bar{N}_0$). This constitutes the non-zero time-integrated CP asymmetry in this decay.

In data, the modulations cancel out. This happens because the number of B^0_d's and $\bar{B}^0_d$'s produced is equal, that is, $N_0 = \bar{N}_0$, and we don't distinguish between $(J/\psi K^0_S)$ decays from B^0_d's and $\bar{B}^0_d$'s. We measure the sum of $N(t)$ and $\bar{N}(t)$.

In the Monte Carlo however, if we generate a $B^0_d \to J/\psi K^0_S$ sample with the branching ratio of the B^0_d to $(J/\psi K^0_S)$ set to 1, the decays are generated according to equation A.1 and we expect to reconstructed a proper time distribution that is described by $N(t)$, i.e. does not follow an exponential, and thus does not describe the reconstructed data. Setting the branching ratio of the $\bar{B}^0_d$'s to $(J/\psi K^0_S)$ to 1 as well will not remedy this shortcoming of the Monte Carlo. In this case,

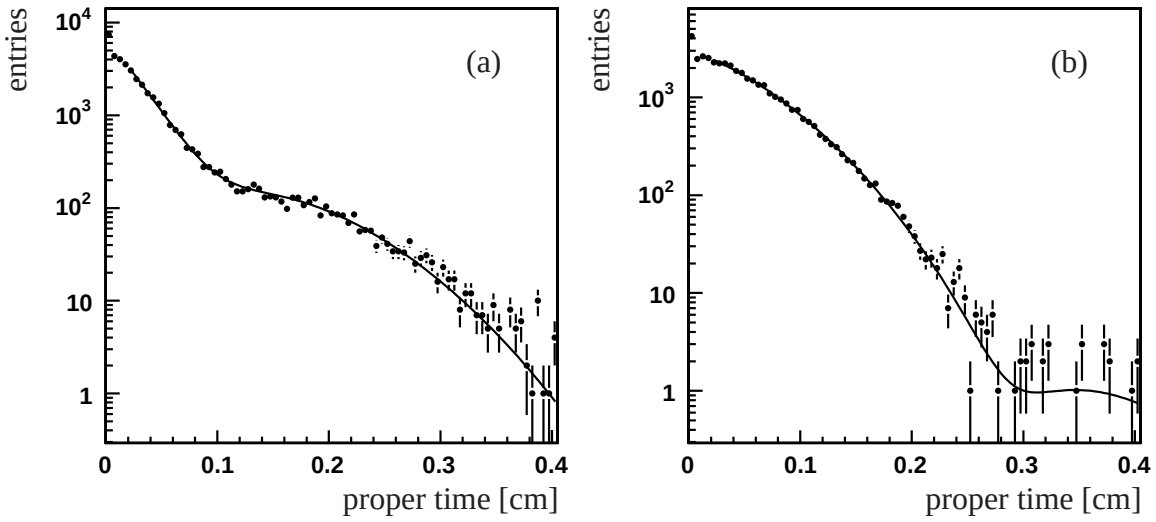


Figure A.1: The proper time distribution of (a) B_d^0's and (b) $\bar{B}_d^0$'s in the decay to $(J/\psi K_S^0)$, as generated by EvtGen. Due to CP violating effects, as properly simulated by EvtGen, the exponential shapes are modulated with a sine curve. The amplitude of the sine is $\sin(2\beta) = 0.7$, and the period is $\Delta m = 0.49$ ps^{-1}.

namely, the overall rate of $B_d^0 \to J/\psi K_S^0$ and $\bar{B}_d^0 \to J/\psi K_S^0$ are set equal, and the integral of the number of events is equal:

$$\int_0^\infty N(t)dt = \int_0^\infty \bar{N}(t)dt. \tag{A.3}$$

That is, there is no time-integrated asymmetry in the observed number of B_d^0 and $\bar{B}_d^0$ decays, as should be. Or, to say it differently, since the total number of generated B_d^0 and $\bar{B}_d^0$ decays is the same, and the events decay according to $N(t)$ and $\bar{N}(t)$, one is forced to conclude that $N_0 \neq \bar{N}_0$, which does not correspond to real life.

We have chosen to fix this problem by modifying EvtGen as follows. Every B_d^0 that is produced we force to decay to $(J/\psi K_S^0)$ with a probability p, and otherwise we allow it to decay according to the default branching ratios. The $\bar{B}_d^0$'s are forced to decay to $(J/\psi K_S^0)$ with a probability $\bar{p}$. The probabilities p and $\bar{p}$ are different, and chosen such that the time-integrated asymmetry that is observed in real life is reproduced. Given a set of $B_d^0 \to J/\psi K_S^0$ and $\bar{B}_d^0 \to J/\psi K_S^0$ decays, p is the fraction of events containing a $B_d^0 \to J/\psi K_S^0$ decay we would observe in real data. This is approximately 32%. The complementary probability $\bar{p}$ is $(1-p)$ or 68%.

After the generation stage of PYTHIA and EvtGen, a final selection was made using the DØMess package. This selection required the presence of a J/ψ and a K_S^0 from the decay of a B_d^0 or B_d^0, and the J/ψ was required to decay into a μ^+ and a μ^-. Each muon was required to have a p_T of at least 1.0 GeV and $|\eta| < 2.5$. The resulting sample consists of 20,000 events.

A.2 $B \to J/\psi X$ MC sample

To generate a sample of events with B-mesons decaying to channels with a J/ψ, we generated $b\bar{b}$ quark pairs, using the selector `MSEL=5`. Each quark was required to have a p_T greater than 100 MeV and $|\eta| < 4.2$. EvtGen was used to decay the B-mesons. The only modification to EvtGen relative to its default version, was that the coherence between the two B-hadrons in the event had been removed. A custom EvtGen decay table was used to force all J/ψ's from B-decays to decay to $(\mu^+\mu^-)$.

The DØMess package was used to select only those events containing a J/ψ, decaying to two muons, each with a p_T greater than 1.5 GeV and $|\eta| < 2.5$. The resulting sample contains 112,500 events.

A.3 Prompt J/ψ MC sample

The prompt J/ψ sample was generated by enabling only those production processes in PYTHIA producing J/ψ's. Enabling specific production processes is done with `MSUB` selectors in the PYTHIA cardfile (Ref. [59] contains the complete list). The `MSUB` processes that have been re-enabled are the dominant prompt J/ψ production processes at the Tevatron:

`MSUB(86):` $gg \to J/\psi g$

`MSUB(106):` $gg \to J/\psi\gamma$

`MSUB(107):` $g\gamma \to J/\psi g$

`MSUB(108):` $\gamma\gamma \to J/\psi\gamma$

Also the dominant processes for producing χ_{c0}, χ_{c1} and χ_{c2} were enabled:

`MSUB(87):` $gg \to \chi_{c0} g$

`MSUB(88):` $gg \to \chi_{c1} g$

`MSUB(89):` $gg \to \chi_{c2} g$

`MSUB(104):` $gg \to \chi_{c0}$

`MSUB(105):` $gg \to \chi_{c2}$

Using the PYTHIA cardfile, all χ_c mesons were forced to decay to $J/\psi\ \gamma$, and all J/ψ's were forced to decay to dimuons. Finally we used DØMess to select the events containing a J/ψ, decaying into two muons that each have a p_T greater than 2.5 GeV, and $|\eta| < 2.5$. The resulting sample contains 96,328 events.

A.4 Minimum bias events

In a single bunch crossing in the Tevatron, multiple $p\bar{p}$ interactions are possible. These interactions produce generally low-momentum particles that enter the detector, in addition to the particles that originate from the process we are studying. To simulate these additional particles, so-called "minimum bias events" (i.e. events that are only marginally biased by a trigger selection) are overlaid with the Monte Carlo events we want to study. The number of additional interactions that is added is drawn from a Poisson distribution with an average of 0.8, which best corresponds to the typical luminosity during collection of our data sample.

The minimum bias events are generated using PYTHIA, with just the elastic, diffractive and low-p_T QCD processes enabled (`MSEL=0`, with `MSUB(92)`, `MSUB(93)`, `MSUB(94)` and `MSUB(95)` set to 1). For these events, PYTHIA was tuned to produce events in better correspondence to CDF Run I dijet data [68]. This tune is usually referred to as the "CDF tune A", and comprises enabling a specific model of multiple scattering (`MSTP(81)=1` and `MSTP(82)=4`), with a smooth turn-off at a p_T =2.4 GeV (`PARP(82)=2.4`). A total of 100,000 minimum bias events have been produced.

Bibliography

[1] S.L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579 (1961); A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm, Almqvist and Wiksell, Stockholm, p. 367 (1968); S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967); M. Veltman, *Perturbation theory of massive Yang-Mills fields*, Nucl. Phys B **7**, 637 (1968); G.M. 't Hooft, *Renormalizable Lagrangians for massive Yang-Mills fields*, Nucl. Phys. B **35**, 167 (1971); G.M. 't Hooft, M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44**, 189 (1972); G.M. 't Hooft, M. Veltman, *Combinatorics of Gauge Fields*, Nucl. Phys. B **50**, 318 (1972).

[2] LEPS Collaboration; T. Nakano *et al.*, *Evidence For A Narrow S = +1 Baryon Resonance In Photoproduction From The Neutron*, Phys. Rev. Lett. **91**, 012002 (2003) [arXiv:hep-ex/0301020].

[3] M. Kobayashi, T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49**, 652 (1973).

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

[5] M. Neubert, *Introduction to B Physics*, arXiv:hep-ph/0001334 (2000).

[6] T. Mannel, *Thirteen years of heavy quark expansion*, Acta Phys. Polon. B **32**, 1857 (2001).

[7] K. Anikeev *et al.*, *B Physics at the Tevatron – Run II and beyond*, FERMILAB-Pub-01/197 (2001).

[8] J. Rademacker, *Heavy flavour lifetimes and lifetime differences*, eConf **C0304052**, WG114 (2003) [arXiv:hep-ex/0306054].

[9] The DØ Collaboration, *Measurement of the lifetime ratio for neutral and charged B-mesons*, DØ Note 4376 (2004) (submitted to XXXIXth Rencontres de Moriond).

[10] The CDF Collaboration, *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*, CDF Note 6387 (2004).

[11] K. Hagiwara *et al*, Phys. Rev. D 66, 010001 (2002) and 2003 off-year partial update for the 2004 edition available on the PDG WWW pages. URL: `http://pdg.lbl.gov`

[12] C.S. Wu, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105**, 1413 (1957).

[13] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the K_2 meson*, Phys. Rev. Lett. **13**, 138 (1964).

[14] P.F. Harrison and H.R. Quinn (editors), *The BaBar physics book*, SLAC-R-504 (1998).

[15] G.V. Borissov, S.Burdin, A. Nomerotski, *Flavor oscillations with same side tagging*, DØ Note 4337 (2004).

[16] M. Gronau, J.L. Rosner, *Identification of neutral B-mesons using correlated hadrons*, Phys. Rev. D **49**, 254 (1994) .

[17] The CDF Collaboration; F. Abe *et al.*, *Measurement of the CP-violation parameter $\sin(2\beta)$ in $B^0_d/\bar{B}^0_d \to J/\psi K^0_S$ decays*, Phys. Rev. Lett. **81**, 5513 (1998).

[18] The CDF Collaboration; T. Affolder *et al.*, *Measurement of $\sin(2\beta)$ from $B^0_d \to J/\psi K^0_S$ with the CDF detector*, Phys. Rev. D **61**, 072005 (2000).

[19] The OPAL Collaboration; K. Ackerstaff *et al.*, *Investigation of CP violation in $B^0 \to J/\psi K^0_S$ decays at LEP*, Eur. Phys. J. C **5**, 379 (1998).

[20] PEP-II, asymmetric B-factory. URL: `http://www.slac.stanford.edu/accel/pepii`

[21] The BABAR Collaboration; B. Aubert *et al.*, *Measurement of the CP-violating asymmetry amplitude $\sin(2\beta)$*, Phys. Rev. Lett. **89**, 201802 (2002) [arXiv:hep-ex/0207042].

[22] The Belle Collaboration; K. Abe *et al.*, *Measurement of the CP-violating parameter $\sin(2\phi_1)$ with 152 million $B\bar{B}$ pairs*, Belle-CONF-0353 (2003).

[23] Heavy Flavor Averaging Group, averages prepared for Winter 2004 conferences. URL: `http://www.slac.stanford.edu/xorg/hfag/triangle/winter2004/index.shtml`

[24] The BABAR Collaboration; B. Aubert *et al.*, *Measurements of CP-violating Asymmetries and Branching Fractions in B Meson Decays to $\eta' K$*, Phys. Rev. Lett. **91**, 161801 (2003) [arXiv:hep-ex/0303046]; T. Browder, *CKM phases: φ_1, β*, talk given at the XXI International Symposium on Lepton and Photon Interactions at High Energies (2003).

[25] The Belle Collaboration; K.Abe *et al.*, *Measurement of Time-Dependent CP-Violating Asymmetries in $B^0 \to \phi K^0_S, K^+K^-K^0_S$ and $\eta' K^0_S$ Decays*, Phys. Rev. Lett. **91**, 261602 (2003) [arXiv:hep-ex/0308035].

[26] Y. Grossman, M.P. Worah, *CP asymmetries in B decays with new physics in decay amplitudes*, Phys. Lett. B **395**, 241 (1997).

[27] The DØ Collaboration; S. Abachi *et al.*, *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995).

[28] The CDF Collaboration; F. Abe *et al.*, *The CDF detector: An overview*, Nucl. Instrum. Meth. A **271**, 387 (1988).

[29] T. LeCompte, H.T. Diehl, *The CDF and DØ upgrades for RunII* (2000). URL: `http://www-d0.fnal.gov/~diehl/Public/annrev/cdfd0.ps`

[30] The CDF Collaboration; F. Abe *et al.*, *Observation of Top Quark Production in $\bar{p}p$ Collisions*, Phys. Rev. Lett. **74**, 2626 (1995)

[31] *The Run II handbook*. URL: `http://www-bd.fnal.gov/runII/index.html`

[32] J. Brzezniak *et al.*, *Conceptual design of a 2 Tesla superconducting solenoid for the Fermilab DØ detector upgrade*, FERMILAB-TM-1886 (1994).

[33] The SVXII Design Group; R. Yarema *et al.*, *A Beginners' Guide to the SVXII*, FERMILAB-TM-1892 (1994).

[34] S. N. Ahmed, R. Lipton, B. Wijngaarden, S. Duensing, A. Naumann, S. de Jong, *DØ SMT Radiation Monitoring System*, DØ Note 4022 (2002).

[35] *DØ Fiber Tracker Technical Design Report* (1997). URL: `http://d0server1.fnal.gov/projects/SciFi/cft_home.html`

[36] P. Balm *et al.*, *Magnetic field monitors for the DØ solenoid*, DØ Note 3977 (2003).

[37] The DØ Collaboration; S. Abachi *et al.*, *The DØ Detector*, Nucl. Instrum. Meth. A **338**, 185 (1994).

[38] The DØ Collaboration, *The DØ Upgrade*, FERMILAB-PUB-96-357-E (1996).

[39] H.T. Diehl, S. Eno, G. Gomez, N.J. Hadley, A.S. Ito, A. Lyon, J. Thompson, *Proposal for Central Muon A-layer Scintillator Trigger Counters*, DØ Note 2213 (1995).

[40] V. Abramov *et al.*, *Technical Design Report for the DØ forward trigger scintillation counters*, DØ Note 3237 (1997).

[41] B.S. Acharya *et al.*, *Technical Design of the Central Muon System*, DØ Note 3365 (1997).

[42] G. Alexeev *et al.*, *Technical Design Report for the DØ Forward Muon Tracking Detector Based on Mini-drift Tubes*, DØ Note 3366 (1997).

[43] W. Busza, *Experience with Iarocci tubes produced on a large scale*, Nucl. Instrum. Meth. A **265**, 210 (1988).

[44] D. Denisov *et al.*, *FAMUS offline software for Run II*, DØ Note 3244 (1997).

[45] C. Miao, *The DØ Run II Luminosity Monitor*, DØ Note 3573 (1998).

[46] G.V. Borisov, *Technical details of AA Tracking*, talk given at the All DØ meeting (February 28, 2003). URL: `http://www-d0.fnal.gov/atwork/adm/d0_private/default.htm`

[47] G.V. Borisov, *VØ's in DØ* , talk given at the DØ B Physics meeting, December 19, 2002.

[48] T. Diehl *et al.*, *Tracking Algorithm Recommendation Committe (TARC) Report II* (2003). `http://www-d0.fnal.gov/phys_id/bid/d0_private/all_track_reprocess.html`

[49] G.V. Borisov, *Lifetime Tag of events $Z^0 \rightarrow b\bar{b}$ with the DELPHI detector. AABTAG program.*, DELPHI Note 94-125/PROG 208 (1994). URL: `http://home.cern.ch/~pubxx/tasks/btagging/www`

[50] M. Narain and F. Stichelbaut, *Vertex Reconstruction using the Impact Parameters Technique*, DØ Note 3560 (1998).

[51] C. Luo, *Muon PDT Time-to-Distance Studies* (2002). URL: `http://hep.physics.indiana.edu/~chuluo/time-to-distance`

[52] O. Peters, *Muon Segment Reconstruction — Linked List Algorithm*, DØ Note 3901 (2001).

[53] F. Déliot, *The Fit Algorithm in muo_trackreco* (2000). URL: `http://www-d0.fnal.gov/~deliot/fitalg.ps`

[54] D. Denisov *et al.*, *DØ Detector magnets for Run II: geometry, currents and B-H curves*, DØ Note 3792 (2000).

[55] G. Hesketh, *Track Extrapolation* (2004). URL: `http://www-d0.fnal.gov/~ghesketh/tracking`

[56] C. Clément *et al.*, *Muon ID Certification for p14*, DØ Note 4350 (2004).

[57] Run Quality Database. URL: `http://d0db.fnal.gov/qualitygrabber/qualQueries.html`

[58] A. Nomerotski, *Magnetic field in the DØ tracker volume*, DØ Note 4312 (2003).

[59] Torbjörn Sjöstrand, Leif Lönnblad, Stephen Mrenna, Peter Skands, *Pythia 6.2 Physics and Manual*, arXiv:hep-ph/0108264 (2002).

[60] J. Stark, *B physics results from DØ* , talk given at the XXXIX Rencontres de Moriond on Electroweak Interactions and Unified Theories (2004). URL: `http://moriond.in2p3.fr/EW/2004/electroweak_uk.html`

[61] U.S. Department of Energy Review of Accelerator Run II, Final Report (2003). URL: http://www-bd.fnal.gov/doereview03

[62] The BTeV project at Fermilab. URL: http://www-btev.fnal.gov

[63] The Large Hadron Collider. URL: http://lhc-new-homepage.web.cern.ch

[64] P. Ball, R. Fleischer, G.F. Tartarelli, P. Vikas, G. Wilkinson (conveners), *B Decays at the LHC*, hep-ph/0003238 (2000).

[65] A. Ryd *et al.*, *EvtGen, a Monte Carlo generator for B-physics*. URL: http://www.slac.stanford.edu/~lange/EvtGen

[66] KEKB, an asymmetric electron-positron collider for B-physics. URL: http://www-acc.kek.jp/WWW-ACC-exp/KEKB/

[67] D. Evans, W. Taylor, *DØMess, DØ Monte Carlo Event Selection System*. URL: http://www-clued0.fnal.gov/d0_mess

[68] R.D. Field, *The underlying event in hard scattering processes*, in *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)* ed. N. Graf, eConf **C010630**, P501 (2001) [arXiv:hep-ph/0201192].

Summary

This thesis describes a measurement of the B_d^0 lifetime in the decay to $(J/\psi K_S^0)$, using 114 pb^{-1} of data collected by the DØ experiment at the Tevatron from October 15, 2002, to June 10, 2003. The measurement is motivated by the tests of the Standard Model that it makes possible. These include tests of Heavy Quark Effective Theory predicting B-meson lifetimes, and of the complex phase in the CKM-matrix as the source of CP-violation in B_d^0 decays to $(J/\psi K_S^0)$.

The lifetime is measured by reconstructing the production and decay vertex of the B_d^0 meson and determining the proper decay length. The $B_d^0 \to J/\psi K_S^0$ decay mode is identified by selecting $J/\psi \to \mu^+\mu^-$ decays and $K_S^0 \to \pi^+\pi^-$ decays. Due to the extremely short lifetime of the J/ψ, the decay vertex of the J/ψ can be identified with the B_d^0 decay vertex. The K_S^0 meson will travel a macroscopic distance before decaying into two π-mesons. The algorithm used to reconstruct the events and to determine the primary vertex as well as the J/ψ and K_S^0 decay vertices are described. A procedure is then developed to select the $B_d^0 \to J/\psi K_S^0$ decays, using only selection cuts that leave the shape of the proper time spectrum undisturbed, as not to introduce any biases on the lifetime measurement. The cut values are optimized for the signal significance, defined as $N_S/\sqrt{N_S + N_B}$, with N_S the number of selected signal decays, and N_B the number of background events in a mass window around the B_d^0 mass. The event sample consists of 56 ± 6 B_d^0 mesons over a background of 133 events in a mass window from 5.16 to 5.4 GeV.

The background in the selected data is found to come from $B \to J/\psi X$ decays and from prompt J/ψ production. The latter background is characterized by a narrow proper time distribution with very little lifetime, the shape of which can be accurately determined. Although the $B \to J/\psi X$ background is the less dominant background in terms of the number of events in the data sample, its understanding is crucial for an accurate lifetime measurement. This background has an inherent lifetime since its origin is genuine B-decays and discriminating this sample from the signal events is critical.

It was identified that two classes of events contribute to the $B \to J/\psi X$ background, depending on the source of the K_S^0. The first category are those events where the K_S^0 is unrelated to the B-meson decay. This class is called "unrelated K_S^0" and distinguishes itself from the signal in two aspects. First, since the K_S^0 is generally produced at the primary vertex, the decay length will be reconstructed with a downward bias. Moreover, the reconstructed mass distribution will be uniform. The second category of events is called "missing decay products". These are $B \to J/\psi K_S^0 X$ events, where not all decay products of the B-meson were identified. This class of events is characterized by a correctly reconstructed decay length, but an underestimated p_T of the B-meson and thus an overestimated proper time. Moreover, the mass of the recon-

structed B-meson will be below the typical B-meson mass due to the missing decay products.

The analysis described in this thesis applies a two-dimensional likelihood fit, that exploits the difference in distribution in mass and in proper time to uniquely distinguish between the various background contributions and the signal. It enables us to disentangle all time constants that affect the proper time spectrum of the data, and determine the B_d^0 lifetime. We have studied all relevant sources of systematic uncertainty and found that the shape of the selection efficiency, as a function of proper time, is the dominant effect. It leads to an uncertainty of 0.04 ps toward larger values of the lifetime, and 0.08 ps toward lower values. Other sources of systematic uncertainty are radial alignment of the detector, the background model, and parameters describing the background whose values were determined from data. The result is a measurement of the B_d^0 lifetime $\tau(B_d^0)$ of:

$$\tau(B_d^0) = 1.38\ ^{+0.26}_{-0.20}\ (\mathrm{stat.})\ ^{+0.05}_{-0.09}\ (\mathrm{syst.})\ \mathrm{ps}.$$

This result is consistent with the current world average. The addition of more data, expected to be accumulated by DØ over the next few years, will allow this result to be greatly improved. The combination of this measurement with lifetime measurements from other B_d^0 decay modes will allow for sensitive tests of the lifetime predictions of Heavy Quark Effective Theory. The addition of flavor tagging to the analysis should allow for a $\sin(2\beta)$ measurement that is competitive with the current world average.

Samenvatting

Dit proefschrift beschrijft een meting van de levensduur van het B_d^0 meson. Mesonen zijn deeltjes die uit een quark en anti-quark bestaan. Een B_d^0 meson bestaat uit een anti-b-quark en een d-quark. De levensduur is bepaald met behulp van vervallen van B_d^0 mesonen naar een J/ψ en een K_S^0 meson. Daarvoor zijn in eerste instantie alle gegevens gebruikt, die zijn vastgelegd door het DØ experiment van 15 oktober 2002 tot 10 juni 2003. Het doel van deze meting is om tests mogelijk te maken van de theorie die het huidige begrip van de elementaire deeltjes en hun wisselwerking beschrijft, het Standaard Model. Dit doel is bereikt: Met behulp van het resultaat van dit proefschrift kunnen verschillende tests uitgevoerd worden.

De levensduur van het B_d^0 meson wordt gemeten door het productie- en vervalspunt te reconstrueren en de transversale impuls te bepalen. Het $B_d^0 \rightarrow J/\psi K_S^0$ vervalskanaal wordt geïdentificeerd door $J/\psi \rightarrow \mu^+\mu^-$ vervallen en $K_S^0 \rightarrow \pi^+\pi^-$ vervallen te selecteren. Door de extreem korte levensduur van de J/ψ kan het vervalspunt van de J/ψ gelijkgesteld worden aan het vervalspunt van de B_d^0. Het K_S^0 meson daarentegen vliegt een zekere afstand voor het in twee π mesonen vervalt. Een procedure is ontwikkeld om vervolgens de $B_d^0 \rightarrow J/\psi K_S^0$ vervallen te selecteren, waarbij alleen selectie sneden gebruikt zijn die de levensduurverdeling ongemoeid laten, om te voorkomen dat de levensduur meting verstoord wordt. De sneden worden zo gekozen, dat de significantie van het signaal optimaal is. De signaal significantie is gedefinieerd als $N_S/\sqrt{N_S + N_A}$, waarin N_S het aantal signaal events is en N_A het aantal achtergrond events in een venster rond de B_d^0 massa. Het resultaat van onze selectie is 56 ± 6 B_d^0 mesonen met een achtergrond van 133 events in een massa venster van 5,16 tot 5,4 GeV.

De achtergrond in de geselecteerde data blijkt te bestaan uit $B \rightarrow J/\psi X$ vervallen en directe J/ψ productie. Deze laatste wordt gekarakteriseerd door een smalle piek rond nul in de levensduurverdeling, die nauwkeurig bepaald kan worden. Hoewel de $B \rightarrow J/\psi X$ vervallen minder sterk aanwezig zijn in termen van aantallen events, is het begrip van deze achtergrond van cruciaal belang voor een nauwkeurige meting van de B_d^0 levensduur. Deze achtergrond heeft namelijk zelf een inherente levensduur, omdat hij voortkomt uit echte b-quark vervallen. De scheiding van deze achtergrond en het signaal is cruciaal.

Het is gebleken dat er twee typen events een bijdrage leveren aan de achtergrond van $B \rightarrow J/\psi X$ vervallen. Deze verschillen in de oorsprong van de K_S^0. Het eerste type bestaat uit die events waar de K_S^0 niet voort komt uit het B-meson verval. We noemen dit type "ongerelateerde K_S^0". Het onderscheid zichzelf van het signaal op twee punten. Ten eerste, aangezien de K_S^0 normaalgesproken bij de primaire vertex geproduceerd is, wordt de vervalslengte van het B-meson gemiddeld onderschat, hetgeen resulteert in een steilere levensduurverdeling. Ten

tweede is het massa spectrum een uniforme verdeling. Het tweede type events is “ontbrekende vervalsproducten” genoemd. Dit zijn $B \rightarrow J/\psi K^0_S X$ vervallen, waarin niet alle vervalsproducten geïdentificeerd zijn. Dit type events wordt gekenmerkt door een correct gereconstrueerde vervalslengte, maar een onderschatte impuls van het B-meson, en daardoor een onderschatte levensduur. Bovendien is de massa van het gereconstrueerde B-meson lager dan de typische B-meson massa.

De analyse, zoals die in dit proefschrift beschreven is, past een twee-dimensionale *likelihood fit* toe, die gebruik maakt van het verschil in de massa- en levensduurverdelingen van de verschillende achtergrond bijdragen en het signaal. Deze methode stelt ons in staat om alle tijdsconstanten te onderscheiden, die de levensduurverdeling van de data beïnvloeden, en de levensduur van het B^0_d signaal te bepalen. Bestudering van alle relevante bronnen van systematische onnauwkeurigheid heeft laten zien, dat de belangrijkste factor de vorm van de selectie efficiency van B^0_d mesonen (als functie van de vervalstijd) is. De onzekerheid op deze vorm heeft een effect op de nauwkeurigheid van 0.04 ps naar hogere waarden van de B^0_d levensduur en 0.08 ps naar lagere waarden. Andere bronnen van systematische onzekerheid zijn de radiële uitlijning van de detector, het model van de achtergronden, en fit parameters die met een zekere nauwkeurigheid zijn bepaald, maar vervolgens zijn vastgezet in de uiteindelijke fit. Het resultaat is een meting van de B^0_d levensduur $\tau(B^0_d)$ van:

$$\tau(B^0_d) = 1,38\ ^{+0,26}_{-0,20}\ (\mathrm{stat.})\ ^{+0,05}_{-0,09}\ (\mathrm{syst.})\ \mathrm{ps}.$$

Hierin hebben we zowel de onzekerheid die veroorzaakt wordt door de beperkte statistiek, als de systematisch onzekerheid aangegeven. Dit resultaat is in overeenstemming met het gewogen gemiddelde van eerdere metingen.

Acknowledgements

Today I look back on my period as a graduate student as one of the most impressive periods in my life. I want to thank Miek for convincing me to choose DØ and Chicago, and my promotor Frank Linde for providing me with the opportunity. Once started, in April 1999, Kors and Onne were there to jumpstart me into the DØ software, and to stake out the Level 3 Muon project for me. I had a lot of fun working on this project, and I'm grateful to them for offering it. I owe the Muon and the Level 3 groups for making it so much fun.

After about half a year I moved to the U.S. Again Onne was there to drive me around and to tease me for failing my driver's exam in the Netherlands. He also turned out a superior partner in complaining about the U.S. ("tgg!"), and eventually in getting over it. Lucio proved a great flatmate during that first period. More people started moving in from Europe: Axel, Lukas and Freya – the Americans most commonly being represented by Daniel and Andy. Your computer and physics related advice has been crucial, though I must say your parties slowed me down here and there. You were great company for some classic bagels, or later in the day, a beer at Rock Bottom and one or two movies at Cantera. Miek I'm thankful for her enthusiasm for moving to the city. That certainly made our stay more interesting.

Arriving back at NIKHEF two and a half years later, the analysis that is the topic of this thesis had to be finished. At this stage the B-physics group back at Fermilab, with Brad, Daria, Guennadi, and Vivek in particular, was a great help. You were always willing to provide me with advice and answers to my questions. The Dutch DØ group was invaluable as a critical peer, testing my arguments for all choices in the analysis. In the final stage of writing it all up, I found Lukas a great help through endless discussions of the final issues of my analysis, but also simply with the process of getting done. Here again, I owe Miek for relentlessly chasing me around with her thesis-schedules (the column "date finished" would be left open for me to fill in). At the time I may not have been very receptive, to say the least, but in retrospect I must admit it helped a great deal. Also at this stage, I appreciated how Paul de Jong and Marcel Demarteau meticulously read and criticized my writings, from structure to grammar, in never-ending review cycles.

All of this helped me finish this work, and I am grateful for it. I feel, however, that in describing this period step by step, I left out the constant factors. At DØ, not just the B-physics group, but the experiment as a whole has been a stimulating place to work. For their moral support, for example by listening to my (rare) frustrations, I'm indebted to my parents and Miek's, my friends, and the guys at NIKHEF (that includes Freya and Mieke). Finally, Marcel Demarteau fully deserves the honor of the last words of these acknowledgements. Discussing my work with

me he has spent more hours than anyone, he has gotten me back on track more than just a number of times, and during the process, he was better able than anyone to make me believe again that one sweet day, I would finish this thesis. He was right.