Search for Rare Decays of the B_s^0 Meson with the DØ Experiment

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der

Universität Zürich

von

Ralf Patrick Bernhard

aus
Deutschland

Promotionskomitee

Prof. Dr. Ulrich Straumann (Vorsitz)
Dr. Frank Lehner

Zürich 2005

To the memory of my grandparents
Alois and Maria Magdalena Bauer

Abstract

This document presents the searches for the flavour-changing neutral current decays $B_s^0 \to \mu^+\mu^-$ and $B_s^0 \to \phi\,\mu^+\mu^-$. A data set with integrated luminosity of 300 pb^{-1} of proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV collected with the DØ detector in Run II of the FERMILAB Tevatron collider is used. The former decay mode is particularly sensitive to supersymmetric extensions of the Standard Model. For the latter mode, a measurement of the branching ratio could validate the prediction of the Standard Model. In the absence of an apparent signal, a limit on the branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ can be computed by normalising the upper limit on the number of events in the B_s^0 signal region to the number of reconstructed $B^\pm \to J/\psi\,K^\pm$ events. An upper limit on the branching fraction of

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 3.7 \times 10^{-7}$$

at a 95% CL is obtained. This limit can be used to constrain models beyond the Standard Model. In models where the lightest supersymmetric particle is considered to be a dark matter candidate the limit aids in restricting the dark matter scattering cross section on nucleons. For the decay $B_s^0 \to \phi\,\mu^+\mu^-$ also no signal has been observed and an upper limit on the branching ratio normalised to $B_s^0 \to J/\psi\,\phi$ events of

$$\frac{\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} < 4.4 \times 10^{-3}$$

at a 95% CL is obtained. In addition, the rare decay $B_s^0 \to \psi(2S)\,\phi$ has been observed. To measure a branching ratio, the $B_s^0 \to J/\psi\,\phi$ mode was used for normalisation, while $B^\pm \to \psi(2S)\,K^\pm$ and $B^\pm \to J/\psi\,K^\pm$ modes were used as control samples. The relative branching ratio has been measured to be

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24\ \text{(stat)}\ \pm 0.09\ \text{(sys)}.$$

Zusammenfassung

Die vorliegende Arbeit beschreibt die Suche nach flavour-ändernden neutralen Strömen der seltenen Zerfälle $B_s^0 \to \mu^+\mu^-$ und $B_s^0 \to \phi\,\mu^+\mu^-$. Dazu wurde ein Datensatz mit einer integrierten Luminosität von 300 pb^{-1} von Proton-Antiproton-Kollisionen benutzt, der bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96$ TeV gesammelt wurde. Dieser Datensatz wurde mit dem DØ-Detektor im Run II des Tevatron-Beschleunigers aufgezeichnet. Der Zerfall $B_s^0 \to \mu^+\mu^-$ ist besonders auf supersymmetrische Erweiterungen des Standardmodells der Teilchenphysik sensitiv. Mithilfe des Zerfalls $B_s^0 \to \phi\,\mu^+\mu^-$ könnte durch eine Messung des Verzweigungsverhältnisses die Vorhersage des Standardmodells überprüft werden. Da in der ersten Suche kein Signal gefunden wurde, kann eine obere Schranke auf das Verzweigungsverhältnis $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ bestimmt werden, indem die beobachtete Anzahl von Ereignissen in der B_s^0 Signalregion auf die Anzahl der rekonstruierten $B^\pm \to J/\psi\,K^\pm$ Ereignisse normiert wird. Es wurde eine obere Grenze auf das Verzweigungsverhältnis von

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 3.7 \times 10^{-7}$$

bei einem Konfidenzniveau (CL) von 95% ermittelt. Diese obere Schranke kann benutzt werden, um mögliche Erweiterungen des Standardmodells der Teilchenphysik auszuschliessen. In bestimmten supersymmetrischen Theorien, in denen das leichteste supersymmetrische Teilchen ein Kandidat für dunkle Materie darstellt, kann nun der Wechselwirkungsquerschnitt zwischen Nukleonen und dunkler Materie eingeschränkt werden. In der Suche nach dem Zerfall $B_s^0 \to \phi\,\mu^+\mu^-$ wurde ebenfalls kein Signal beobachtet. Daher konnte eine obere Schranke auf das Verzweigungsverhältnis normiert auf den Zerfall $B_s^0 \to J/\psi\,\phi$ von

$$\frac{\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} < 4.4 \times 10^{-3}$$

bei einem CL von 95% ermittelt werden. Ausserdem konnte der seltene Zerfall $B_s^0 \to \psi(2S)\,\phi$ beobachtet werden. Um ein Verzweigungsverhältnis zu messen, wurde der Zerfall $B_s^0 \to J/\psi\,\phi$ als Normierung benutzt. Ferner wurden die bekannten Zerfälle $B^\pm \to \psi(2S)\,K^\pm$ und $B^\pm \to J/\psi\,K^\pm$ als Kontrollkanäle studiert. Das relative Verzweigungsverhältnis wurde schliesslich zu

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24\ \text{(stat)}\ \pm 0.09\ \text{(sys)}$$

bestimmt.

Contents

Introduction **1**

1 Standard Model of Particle Physics **3**

1.1 Electroweak Interactions within the Standard Model 4

1.2 Quantum Chromodynamics 6

1.3 The CKM Matrix 7

1.4 Beyond the Standard Model 8

2 Flavour-Changing Neutral Current Decays of B Mesons **11**

2.1 Introduction 11

2.2 Theoretical Framework 12

2.2.1 Effective FCNC Vertices 12

2.2.2 Operator Product Expansion Approach 12

2.2.3 Effective Hamiltonian 13

2.3 Radiative Decays $b \to s\,\gamma$ 16

2.4 The Decays $b \to s\,\ell^+\ell^-$ 17

2.4.1 Inclusive Decays 19

2.4.2 Exclusive Decays 20

2.5 The Decays $B^0_{s,d} \to \ell^+\ell^-$ 22

2.5.1 Beyond the Standard Model 24

3 Experimental Environment **32**
3.1 Tevatron 32
3.1.1 The Tevatron Accelerator Complex 33
3.2 The DØ Detector 35
3.2.1 Coordinate System 35
3.2.2 Luminosity System 37
3.2.3 Central Tracking System 38
3.2.4 Calorimeter System 42
3.2.5 Muon System 45
3.2.6 Trigger System 49

4 Event Reconstruction and Object Identification **51**
4.1 Tracks of Charged Particles 51
4.2 Primary Vertex 52
4.3 Secondary Vertex 53
4.4 Muons 53
4.4.1 Muon Identification Criteria 54

5 Data and Event Selection **56**
5.1 Data Sample 56
5.2 Data Quality Selection 56
5.3 Trigger Selection 57
5.4 Monte Carlo Simulation 58

6 Optimisation **61**
6.1 Discriminating Variables 61
6.1.1 Pointing Angle 61
6.1.2 Isolation 62
6.1.3 Transverse Decay Length Significance 62
6.2 Random Grid Search 63
6.3 Probability Density Ratio 64
6.4 A Multivariate Linear Classifier 65

7 Calculating Limits **67**

7.1 Probabilities 67

7.2 Confidence Intervals 67

7.3 Confidence Intervals for Poisson Distributions 69

7.4 Confidence Intervals for the Bayesian approach 70

7.5 Confidence Intervals and Systematic Uncertainties 71

7.6 Observing fewer Events than Expected 72

7.7 Expected Upper Limit 73

8 Search for the Decay $B_s^0 \to \mu^+\mu^-$ **74**

8.1 Pre-selection Requirements 74

8.2 Optimisation of Discriminating Variables 76

8.2.1 Random Grid Search 78

8.2.2 Likelihood Ratio 78

8.3 The Normalisation Channel $B^\pm \to J/\psi\, K^\pm$ 80

8.4 Deriving an Upper Limit 83

8.5 Sensitivity 85

8.6 Systematic Uncertainties 86

8.7 Results 87

9 Search for the Rare Decay $B_s^0 \to \phi\, \mu^+\mu^-$ **90**

9.1 Introduction 90

9.2 Pre-selection Requirements 91

9.3 Optimisation of Discriminating Variables 93

9.4 The Normalisation Channel $B_s^0 \to J/\psi\, \phi$ 94

9.5 Deriving an Upper Limit 97

9.6 Sensitivity 99

9.7 Systematic Uncertainties 99

9.8 Results 101

10 Observation of the Decay $B_s^0 \to \psi(2S)\,\phi$ **103**
10.1 Introduction . . . 103
10.2 Event Selection . . . 103
10.2.1 Pre-selection . . . 104
10.3 Reconstruction of the Decays $B^\pm \to J/\psi\,K^\pm$ and $B^\pm \to \psi(2S)\,K^\pm$ . . . 105
10.4 Reconstruction of the Decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\,\phi$ . . . 106
10.5 Optimisation Procedure . . . 107
10.5.1 Random Grid Search . . . 109
10.5.2 A Multivariate Linear Classifier . . . 111
10.6 Relative Branching Ratio Measurements . . . 112
10.6.1 Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ . . . 112
10.6.2 Measurement of $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$ . . . 114
10.7 Systematic Uncertainties . . . 114
10.7.1 Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ . . . 115
10.7.2 Measurement of $\mathcal{B}(B^\pm \to J/\psi\,K^\pm)/\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)$ . . . 116
10.8 Conclusions . . . 116

11 Outlook and Summary **118**
11.1 Outlook for $B_s^0 \to \mu^+\mu^-$ . . . 118
11.2 Outlook for $B_s^0 \to \phi\,\mu^+\mu^-$ . . . 119
11.3 Summary . . . 119

A Deriving the Master Formula **122**

B Signal Monte Carlo Data Matching for $B_s^0 \to J/\psi\,\phi$ Events **124**

C Depletion Voltage Studies **127**
C.1 Radiation Damage . . . 127
C.2 Method for Measuring the Depletion Voltage . . . 128
C.3 The Model . . . 129
C.4 Results . . . 130

List of Figures **133**

List of Tables **138**

Bibliography **141**

Acknowledgments **149**

Curriculum Vitae **150**

Introduction

Since the origin of mankind, man has tried to understand its environment and the manifold processes that are related with it. He has tried to find rules for regularities underlying his observations. As time evolved the techniques to investigate and describe observation evolved, too. At all times, the goal was to find one coherent description for all processes.

Today we observe very small as well as very large-scale objects and try to find one coherent description for both. The very large-scale objects are described by the theory of gravitation. The very small objects are the subject of the field of particle physics and the theory which describes best our findings is called the Standard Model of particle physics. Within the Standard Model, there exist descriptions of the most basic particles and their properties. Also described are the means by which these particles interact with each another, giving rise to the forces observed in nature. This approach of predicting particle interactions has proven to be a very powerful tool for understanding matter on its most fundamental level. However, the Standard Model is an approximate theory and is not believed to be the ultimate description of nature. These issues reinforce the belief that the Standard Model only describes a portion of physics and that a more complete theory could someday be formulated. Therefore, physicists continue to test the Standard Model's predictive ability in an effort to identify physics that cannot be described by the model. Such searches for new phenomena will continue to expand the unknown horizon of particle physics.

The most important method for testing the Standard Model is the acceleration of particles to high energies. If two accelerated particles are collided they can annihilate or scatter. If the center-of-mass energy of the colliding particles is sufficiently large the resulting annihilation or scattering can create final-state particles that are different from the original particles. There are two main types of particle accelerators: linear accelerators and synchrotrons. Linear accelerators are machines which accelerate electrically charged particles to a specific energy over a straight path. Collisions between particles can be arranged by directing the particle beams of two opposite-facing linear accelerators at a single interaction point. Synchrotrons are circular accelerators. Such accelerators allow for two counter-rotating beams in the accelerator at the same time, and these beams can be brought into collision at different points along the ring.

Around the interaction point a detector is built to observe the properties of the new particles created. The most useful information is the energy and momentum of the final state particles created in the interaction of the initial particles. The detectors contain subdetectors which allow tracking of the particle to measure their direction of flight and their momentum and calorimetry to measure deposited energies.

This thesis is organised as follows:

The first Chapter provides a brief introduction into the Standard Model. The theoretical background on flavour changing neutral decays of B mesons is given in the second Chapter.

An overview of the experimental environment, the FERMILAB collider and the DØ detector are presented in Chapter 3. This is followed by an introduction into event reconstruction and object identification in Chapter 4 and data and event selection in Chapter 5.

The optimisation of the analyses and various multivariate techniques used are presented in Chapter 6. A short introduction to the calculation of limits using different approaches is given in Chapter 7.

The analyses are presented in Chapters 8 to 10. In Chapter 8 the search for the rare flavour changing neutral current decay $B_s^0 \to \mu^+\mu^-$ is presented followed by the search for the rare decay $B_s^0 \to \phi\,\mu^+\mu^-$ in Chapter 9. In Chapter 10 the observation of the decay $B_s^0 \to \psi(2S)\,\phi$ and a measurement of its relative branching ratio with respect to $B_s^0 \to J/\psi\,\phi$ is presented. This Chapter reports also on the measurement of the relative branching ratio of $B^\pm \to \psi(2S)\,K^\pm$ with respect to $B^\pm \to J/\psi\,K^\pm$.

This thesis is concluded with an outlook and a summary.

Chapter 1

Standard Model of Particle Physics

The interactions between the known smallest constituents of matter, namely the quarks and leptons, are described by the Standard Model[1] of particle physics. Although the Standard Model is very successful in making predictions for physical observables with high accuracy, it is not thought to be a final theory as it contains many free parameters whose values cannot be predicted but have to be measured by experiments.

The Standard Model is in fact a composition of two theories: the *Glashow-Salam-Weinberg* (GSW) model of electroweak interactions, describing the electromagnetic and weak forces, and *Quantum Chromodynamics* (QCD), describing the strong force. Still separated from the other forces is Gravitation. No quantum field theory has yet been formulated to describe it, the best description being Einstein's theory of General Relativity.

The Standard Model contains two kinds of fundamental elementary particles: matter particles and force carriers. The matter particles can be divided into two types of fermions (particles with intrinsic spin of 1/2): quarks and leptons. These fermions all have mass and are the constituents of the physical universe. Each fermion has an associated anti-particle with equal mass and lifetime but opposite charge quantum numbers. The fermions can be classified in generations according to their rising mass. Corresponding particles from different generations have different mass but otherwise identical quantum numbers and interactions. The Standard Model does neither constrain nor predict the number of generations. Until now three generations have been observed, as shown in Table 1.1. The force carriers have integral spin and are called bosons. Interactions between the matter particles are mediated via exchanges of these bosons. These exchanges give rise to the three Standard Model forces: electromagnetic, weak, and strong. Although not yet included in the framework of the Standard Model, Gravitation is thought to be mediated by a massless spin two boson called the *graviton*. The observed five force carriers are listed in Table 1.2.

[1] In the following description of the Standard Model, the conventions of [1] have been used.

Table 1.1: The spin-1/2 particles (fermions) of the Standard Model of particle physics.

Leptons			Quarks		
Particle Type	Symbol	Charge	Particle Type	Symbol	Charge
electron neutrino	ν_e	0	up quark	u	2/3
electron	e	-1	down quark	d	-1/3
muon neutrino	ν_μ	0	charm quark	c	2/3
muon	μ	-1	strange quark	s	-1/3
tau neutrino	ν_τ	0	top quark	t	2/3
tau	τ	-1	bottom quark	b	-1/3

Table 1.2: The integral-spin particles (bosons) of the Standard Model of particle physics.

Bosons			
Particle Type	Symbol	Charge	Force Mediated
photon	γ	0	Electromagnetic
W^+ boson	W^+	+1	Weak
W^- boson	W^-	-1	Weak
Z boson	Z	0	Weak
gluon	g	0	Strong

1.1 Electroweak Interactions within the Standard Model

In the Standard Model, electroweak interactions are described by an $SU(2)_L \otimes U(1)_Y$ gauge theory, combining the two groups $SU(2)_L$ of weak isospin and $U(1)_Y$ of weak hypercharge. In this theory, the left-handed fermion fields form weak isospin doublets while the right-handed components f_R are $SU(2)_R$ singlets. The first generation can be represented as

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L, \; e_R \qquad \text{and} \qquad \begin{pmatrix} u \\ d \end{pmatrix}_L, \; u_R, \; d_R. \tag{1.1}$$

This classification ensures that the parity violating $V-A$ structure of the weak interactions will be obtained. The weak hypercharge Y is related to the electric charge Q and the third component of the isospin I_3 via

$$Q = I_3 + \frac{Y}{2}. \tag{1.2}$$

By imposing local gauge invariance on $SU(2)_L \otimes U(1)_Y$ the GSW gauge theory is obtained. The $SU(2)_L$ weak isospin gauge group will have three gauge bosons, namely W^1_μ, W^2_μ and W^3_μ, with one coupling constant g. The $U(1)_Y$ hypercharge gauge group will have one gauge boson B_μ, with a coupling constant $g'/2$. However, the Lagrangian that is obtained by requiring this gauge invariance does not describe the electroweak interactions correctly. Moreover, all gauge bosons and fermions in the theory are massless. This is obviously not what is observed in nature.

The GSW model addresses the problem of gauge boson masses by introducing an $SU(2)$ doublet of complex scalar (spin-0) fields to the Standard Model Lagrangian

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \text{with} \quad \begin{array}{c} \phi^+ = (\phi_1 + i\phi_2)/\sqrt{2} \\ \phi^0 = (\phi_3 + i\phi_4)/\sqrt{2} \end{array} \tag{1.3}$$

which interacts with itself via a potential $V(\phi) = \mu^2\, \overline{\phi}\phi + \lambda \left(\overline{\phi}\phi\right)^2$. To force non-zero boson masses, we consider the case of $\mu^2 < 0$ and $\lambda > 0$, giving $V(\phi)$ two minima at $\phi_0 = \pm\sqrt{-\mu^2/2\lambda} = \pm v$. We have the freedom to choose either solution as the ground state for ϕ and in this discussion we shall choose the positive solution. We can then rewrite the scalar field as an excitation relative to this non-zero vacuum expectation value

$$\phi(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.4}$$

which manifestly breaks the $SU(2)_L \otimes U(1)_Y$ gauge symmetry and introduces a single real scalar field, $h(x)$, which is referred to as the Higgs field. Via interactions with the Higgs field, the Standard Model Lagrangian obtains a new term which generates the gauge boson masses.

By defining the orthogonal combinations

$$W^\pm_\mu = \frac{1}{\sqrt{2}} \left(W^1_\mu \mp i W^2_\mu\right) \quad \text{with} \quad m_W = \frac{v\, g}{2} \tag{1.5}$$

$$Z^0_\mu = \frac{1}{\sqrt{g^2 + k^2}} \left(g W^3_\mu - g' B_\mu\right) \quad \text{with} \quad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \tag{1.6}$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} \left(g' W^3_\mu + g B_\mu\right) \quad \text{with} \quad m_A = 0 \tag{1.7}$$

one can now interpret the weak gauge fields and the photon field. The three extra degrees of freedom obtained from $\phi(x)$ appear as longitudinal polarisations of the original W^i_μ fields, allowing them to become massive. The $W^\pm_\mu$ can be identified with the charged W bosons, the Z^0_μ field with the Z^0 boson and the A_μ field with the photon. A weak mixing angle θ_W

$$\tan\theta_W = \frac{g'}{g}, \tag{1.8}$$

called Weinberg angle, can be defined, which parametrises the mixing of the neutral gauge bosons.

The interaction terms in the Lagrangian will give the electromagnetic and weak interactions, whose coupling constants are given as:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \tag{1.9}$$

$$G_F = \frac{v}{\sqrt{2}} \tag{1.10}$$

The introduction of the Higgs field, although it spontaneously breaks $SU(2)_L \otimes U(1)_Y$ symmetry upon the choice of a ground state for the Higgs potential, solves the problem of weak boson masses while keeping the theory renormalisable. Fermion masses are obtained in a similar manner through their interactions with the Higgs field. One also gets a mass term for the Higgs field itself given by $m_h = \sqrt{2}\mu^2$. The only unsatisfactory feature of this new Higgs field is that its mass is not predicted by the theory, even though the $W^{\pm}$ and Z^0 boson masses can be determined by measurements of weak decays. Until now, no experimental evidence has shown any hint that this model may be not complete, although one crucial component, the Higss boson, has not been experimentally observed.

1.2 Quantum Chromodynamics

The third Standard Model force is the strong force. This force is incorporated into the Standard Model Lagrangian via the requirement of $SU(3)$ gauge invariance. The fundamental representation of this group requires three quantum numbers, which are referred to as "colours" in the quark sector, thus motivating the name quantum chromodynamics (QCD). As $SU(2)$ gave rise to $2^2 - 1 = 3$ gauge bosons, QCD adds $3^2 - 1 = 8$ new fields, referred to as gluons. Quarks are defined as containing one of three colour fields (chosen as red (r), green (g), and blue (b)), defining three quark and three anti-quark quantum states. The gluons occupy eight colour states:

$$r\bar{g},\ r\bar{b},\ g\bar{r},\ g\bar{b},\ b\bar{r},\ b\bar{g},\ \left(r\bar{r} - g\bar{g}\right)/\sqrt{2},\ \text{ and } \left(r\bar{r} + g\bar{g} - 2b\bar{b}\right)/\sqrt{6}.$$

These gluons are massless and are the carriers of the strong force. Quarks and anti-quarks are bound by gluons into colour singlet combinations to form colourless mesons and hadrons.

1.3 The CKM Matrix

CP is the combined transformation of the charge conjugation C and the parity transformation P. The parity transformation is a simple inversion of all space coordinates and the charge conjugation converts a particle into its antiparticle. In 1957, it was shown that C and P are separately violated in weak interactions. It was then thought for some time that the combined operation CP was conserved. However, in 1964, Christenson, Cronin, Fitch, and Turlay observed [2] CP violation in the neutral Kaon system, by observing the CP-forbidden two body decay $K_L \to \pi^+\pi^-$.

To introduce into the Standard Model a mechanism to generate CP violation Kobayashi and Maskawa proposed in 1973 [3] the introduction of a third generation of quarks. The idea was to extend the existing mixing matrix

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = U \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{1.11}$$

with θ_C the quark mixing angle (also known as Cabibbo angle) from two generations to three generations. At that time there was evidence for the existance of three quarks only. The primes in Eq. 1.11 indicate that the weak eigenstates (d', s') are not equal to the corresponding mass eigenstates (d, s), but rather a linear combination of the latter.

As the quark mixing matrix has to be unitary, the complex $n \times n$ matrix for n generations of quarks is initially defined by n^2 parameters. After defining the quark phases $(n-1)^2$ independent parameters are left. They are split in $n(n-1)/2$ angles and $(n-1)(n-2)/2$ phases. For two quark generations, a single angle, the Cabibbo angle, describes the matrix. For three quark generations, the so called *Cabibbo-Kobayashi-Maskawa* (CKM) mixing matrix is described by three angles and one phase. To emphasize the physical transitions associated with the CKM matrix, it is usually written as

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.12}$$

so that the elements of the matrix are labelled by the quark flavours. The parameterisation favoured by the Particle Data Group [4] is

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.13}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The real angles θ_{ij} may be chosen such that $0 \leq \theta_{ij} \leq \pi/2$, and the phase δ_{13} such that $0 \leq \delta_{13} < 2\pi$. In this parameterisation the CKM matrix is the product of three complex matrices, of which only one has a non-zero phase. This

has the advantage that each of the rotation angles (θ_{12}, θ_{23}, θ_{13}) relates to the mixing of two specific generations. The non-zero value of the phase (δ_{13}) causes CP violation in the weak interactions.

Another convenient parameterisation of the CKM matrix is due to Wolfenstein [5]. It stems from the observation that the measured matrix obeys a hierarchy, with diagonal elements close to 1, and progressively smaller elements away from the diagonal. This hierarchy can be formalised by defining λ, A, ρ, and η via

$$\lambda \equiv s_{12}\,, \qquad A \equiv s_{23}/\lambda^2\,, \qquad \rho + i\eta \equiv s_{13}e^{i\delta_{13}}/A\lambda^3\,. \tag{1.14}$$

From experiment, $\lambda \approx 0.22$, $A \approx 0.8$, and $\sqrt{\rho^2+\eta^2} \approx 0.4$, and it is phenomenologically useful to expand V_{CKM} in powers of λ:

$$V_{CKM} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{1.15}$$

The unitarity relation for the CKM matrix ($V_{CKM}V^\dagger_{CKM} = V^\dagger_{CKM}V_{CKM} = 1$) leads to nine equations among its elements. Six of these have a result of zero and can be represented as triangles in the complex plane. Most of the relations yield flat triangles, except for equations involving the first and the last columns, where the three terms to be summed are of the same order in λ ($\mathcal{O}(\lambda^3)$):

$$\begin{gathered} V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0 \\ [(\rho + i\,\eta) + (-1) + (1-\rho - i\,\eta)]A\lambda^3 + \mathcal{O}(\lambda^5) = 0. \end{gathered} \tag{1.16}$$

The base of this triangle is usually normalised by dividing the equation by the second term of Eq. 1.16, $V_{cd}V^*_{cb}$. This yields the so called *unitarity triangle* with a base of unity, as illustrated in Figure 1.1. The coordinate of the apex in the Wolfenstein parameterisation is simply (ρ,η).

1.4 Beyond the Standard Model

The Standard Model has about twenty free parameters that are not fixed by the gauge principles. These are the strengths of the gauge couplings, the Yukawa couplings of quarks and leptons, the mixing angles and the mass of the Higgs particle. The development of particle physics in the past twenty years was marked by the accurate experimental determination of most of these parameters. The Standard Model has proven to be a sufficient theory describing fundamental particles and their interactions up to the energies presently achievable in an era of precision electroweak measurements. The robustness of this theory is due largely to the fundamental simplicity it uses to approach a description of matter.

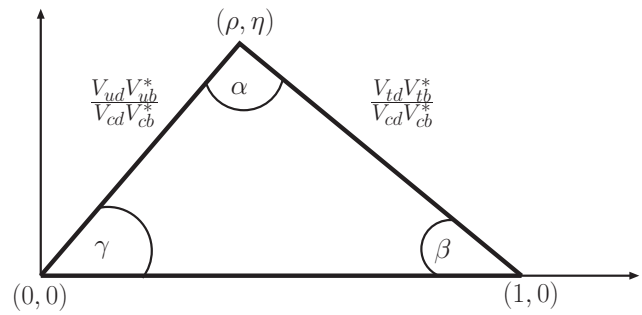

Figure 1.1: The unitarity triangle.

However, one of the most interesting parameters, the Higgs mass and indeed the Higgs particle itself have eluded experimental discovery so far. Furthermore, questions about the origin of the symmetry breaking and the great disparity between the mass scales of the Standard Model and gravitational interactions lead to believe that the Standard Model is an effective theory. That is, that at some higher energy scale it will be incorporated in an even more fundamental theory.

The electroweak theory postulated a single interaction to describe electromagnetic and weak processes, and spontaneous symmetry breaking to account for their different strengths in the energy region below the masses of the mediating bosons. The so-called Grand Unified Theories (GUTs) appeal to a further broken symmetry in order to adapt the relatively strong coupling of strong interactions at low energies to a unique intrinsic coupling for all three interactions at the unification energy. There are many ways in which the $SU(2)_L$, $U(1)_Y$ and $SU(3)$ symmetries could be incorporated into a more global gauge symmetry. The simplest is the $SU(5)$ group proposed by Georgi and Glashow [6]. This incorporates the known fermions in multiplets, inside which quarks can transform to leptons and antiquarks, via the mediation of very massive (10^{15} GeV/c^2) bosons called X and Y. In this model there are a total of 24 gauge bosons, the Standard Model ones, plus 12 varieties of X and Y. The model makes a number of predictions, for example the decay of the proton which has not been observed.

Another approach is Supersymmetry (SUSY) which provides a stabilisation of the Higgs sector at high energies and solves the hierarchy problem in the gauge sector. Furthermore, it predicts a light dark matter candidate. This model introduces a symmetry relating particles of different spin, i.e., fermions and bosons. Particles are combined in a *superfield* containing two fields differing by one-half unit of spin. Thus, each fermion of the Standard Model is given a bosonic superpartner (forming a chiral superfield) and each boson of the Standard Model is given a fermionic superpartner (forming a vector superfield). Particles in a superfield have the same masses and quantum numbers aside from the 1/2 unit of spin. This is a problem as no scalar particles with the (small) masses of the

Standard Model leptons have been observed, directly or indirectly. Thus, SUSY must be a broken symmetry, and the mechanism for this symmetry breaking is not well described. The currently accepted means of breaking the symmetry requires the superpartners to the Standard Model particles to have masses less than roughly 1 TeV. Furthermore, the simplest anomaly-free SUSY model requires two $SU(2)$ Higgs-type doublets of complex scalar fields, therefore predicting not one but five scalar Higgs fields. This introduces a second unpredicted parameter to the Higgs sector namely the ratio of the vacuum expectation values ($\tan\beta = v_1/v_2$) for the two Higgs doublets. Despite its deficiencies, SUSY provides a badly needed substructure for a Standard Model-like gauge field theory. The predictions of this model provide an excitingly rich array of new physics which can be tested at energies in the TeV scale.

Chapter 2

Flavour-Changing Neutral Current Decays of B Mesons

2.1 Introduction

Flavour-changing neutral current (FCNC) transitions, such as $b \to s$ and $b \to d$, are forbidden in the Standard Model of electroweak interactions at tree level. They can proceed at very low rate through higher order box and penguin diagrams.

In the Standard Model, neutral currents are flavour conserving. The unitarity of the CKM matrix ensures that. There is no direct coupling between the b and the s or d quarks. If all quarks of same isospin were of equal mass, higher order processes would also cancel, as explained by the *Glashow-Iliopoulos-Maiani* (GIM) mechanism [7], and FCNC processes would not exist at any order. Thus, in the B system, consisting of a heavy b quark and a light u, d, or s anti-quark, it is due to the high mass of the top quark and the fact that the coupling of the b quark to the t quark is close to unity, that these processes are not entirely negligible. Moreover, their relative importance is enhanced by the small coupling of the b quark to the c quark, V_{cb}, which suppresses the otherwise CKM-favoured tree diagram decays. Precision measurements of these rare FCNC processes represent a sensitive probe for new physics, as many processes of extensions to the Standard Model enhance or suppress FCNC decay rates. This permits to indirectly probe these models while measurements of direct effects could still be beyond reach at present day accelerators.

In the following, a brief overview of the theoretical framework of FCNC decays will be given. An in-depth review by Buras and Fleischer [8] and the references therein give a good introduction into this subject.

2.2 Theoretical Framework

2.2.1 Effective FCNC Vertices

At the one loop level, new structures appear which were absent at tree level. Amongst these are the flavour-changing neutral current transitions which can be summarised by a set of basic triple and quartic effective vertices. In the literature they appear under the names of penguin and box diagrams, respectively. Typical diagrams are depicted in Figure 2.1.

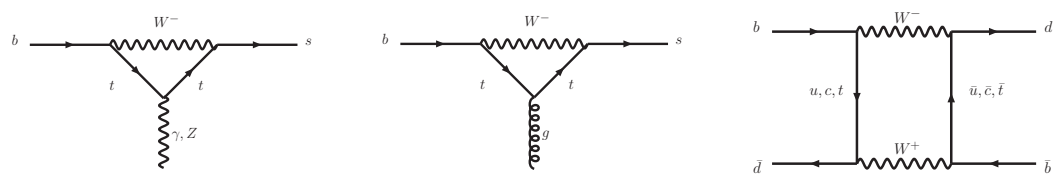

Figure 2.1: Example of penguin and box diagrams

These effective vertices can be calculated by using elementary vertices and propagators, and effective "Feynman rules" can be derived [8]. The higher order in the gauge couplings contributes to the suppression of these diagrams. Their relative importance depends on the mass of the internal fermion lines, which explains the importance of the contribution of the top quark, and on the relevant CKM matrix element.

2.2.2 Operator Product Expansion Approach

To predict the decay rate of a given meson M into some final state f, one must calculate the transition amplitude $\mathcal{M}$ for $M \to f$. In general there are many contributions to $\mathcal{M}$, each of which is at the quark level represented by Feynman diagrams.

The most often used theoretical tool to calculate decay rates for FCNC decays is the Operator Product Expansion approach (OPE). The basic idea behind the OPE approach is that the energy scale of the relevant interactions (weak decays of hadrons containing u, d, s, c and b quarks) is low compared to the mass of the propagator, the W boson. As the W boson propagator is of the form $\frac{1}{q^2-M_W^2}$ (where q is the momentum transferred by the W boson), the amplitude can be written as an expansion in terms of q^2/M_W^2. Schematically the decay amplitude $\mathcal{M}$ is then expressed as

$$\mathcal{M} = -\frac{4\,G_F}{\sqrt{2}}\,V_{\mathrm{CKM}} \sum_j C_j(\mu)\,\langle f|Q_j(\mu)|M\rangle \left[1 + \mathcal{O}\left(\frac{q^2}{M_W^2}\right)\right], \tag{2.1}$$

where μ is a renormalisation scale. Physics from distances shorter than μ^{-1} is contained in the Wilson coefficients C_j, and physics from distances longer than μ^{-1} is accounted for by the hadronic matrix elements $\langle f|Q_j|B\rangle$ of the local operators Q_j. In principle, there are infinitely many terms in the OPE, but higher dimension operators yield contributions suppressed by powers of q^2/m_W^2. Therefore, when calculating a decay amplitude, terms with $\mathcal{O}(q^2/M_W^2)$ and higher can be neglected. This is equivalent to writing an effective Hamiltonian, $\mathcal{H}_{\text{eff}}$, where the short range contributions of a heavy boson can be replaced by a point-like interaction. To get $\mathcal{H}_{\text{eff}}$, the product of two charged-current operators is expanded in a series of local operators. The contribution of these operators is weighted by effective coupling terms, called Wilson coefficients. $\mathcal{H}_{\text{eff}}$ can then be written as

$$\mathcal{H}_{\text{eff}} = -\frac{4\,G_{\text{F}}}{\sqrt{2}} V_{\text{CKM}} \sum_j C_j(\mu) Q_j, \tag{2.2}$$

where Q_j are the local operators and C_j the Wilson coefficients.

This approach factorises the problem into two parts. Low-energy (long-distance) interactions (hadronisation effects) are contained in the operators. High-energy (short-distance) interactions (hard gluon exchanges) are represented by the Wilson coefficients. The renormalisation scale μ represents the limit between the two different regimes. For B decays it is chosen to be at the order of a few GeV. The Wilson coefficients are functions of the strong coupling constant α_s, M_W and μ. They are calculated perturbatively at $\mu_W = M_W$ and then evolved to the chosen low-energy scale μ using renormalisation group equations. At μ_W perturbative calculations can be made, as the effective QCD couplings are small.

2.2.3 Effective Hamiltonian

The similarity of FCNC processes, in decays or mixing, allows to describe all of them with a set of basic effective vertices. The effective Hamiltonian for FCNC decays can then be written as an expansion in terms of the four-fermion local operators that describe these effective vertices. These local operators can be classified in six classes [8]

Current–Current (Figure 2.2a):

$$\begin{aligned} Q_1 &= (\bar{s}_{L\beta}\gamma_\mu c_{L\alpha})\,(\bar{c}_{L\alpha}\gamma^\mu b_{L\beta}) \\ Q_2 &= (\bar{s}_L\gamma_\mu c_L)\,(\bar{c}_L\gamma^\mu b_L) \end{aligned}$$

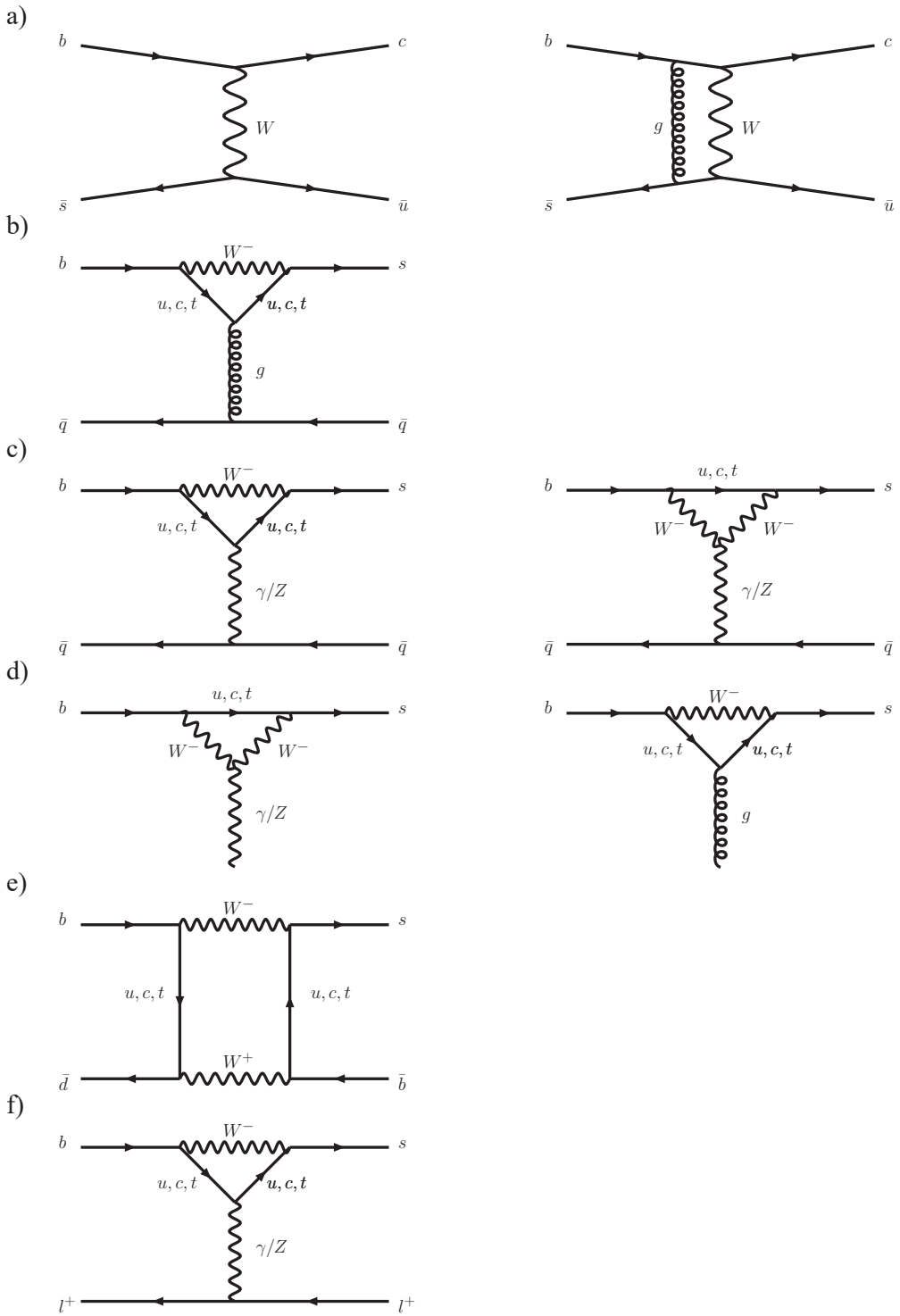

Figure 2.2: Typical diagrams for FCNC decays for the six different classes as described in the text.

QCD–Penguins (Figure 2.2b):

$$
\begin{aligned}
Q_3 &= (\bar{s}_L\gamma_\mu b_L)\sum_{q=u,d,s,c,b}(\bar{q}\gamma^\mu q)\\
Q_4 &= (\bar{s}_{L\alpha}\gamma_\mu b_{L\beta})\sum_{q=u,d,s,c,b}(\bar{q}_{L\beta}\gamma^\mu q_{L\alpha})\\
Q_5 &= (\bar{s}_L\gamma_\mu b_L)\sum_{q=u,d,s,c,b}(\bar{q}_R\gamma^\mu q_R)\\
Q_6 &= (\bar{s}_{L\alpha}\gamma_\mu b_{L\beta})\sum_{q=u,d,s,c,b}(\bar{q}_{R\beta}\gamma^\mu q_{R\alpha})
\end{aligned}
$$

Electroweak–Penguins (Figure 2.2c):

$$
\begin{aligned}
Q_{3Q} &= \frac{3}{2}\,(\bar{s}_L\gamma_\mu b_L)\sum_{q=u,d,s,c,b} e_q\,(\bar{q}_R\gamma^\mu q_R)\\
Q_{4Q} &= \frac{3}{2}\,(\bar{s}_{L\alpha}\gamma_\mu b_{L\beta})\sum_{q=u,d,s,c,b} e_q(\bar{q}_{R\beta}\gamma^\mu q_{R\alpha})\\
Q_{5Q} &= \frac{3}{2}\,(\bar{s}_L\gamma_\mu b_L)\sum_{q=u,d,s,c,b} e_q(\bar{q}_L\gamma^\mu q_L)\\
Q_{6Q} &= \frac{3}{2}\,(\bar{s}_{L\alpha}\gamma_\mu b_{L\beta})\sum_{q=u,d,s,c,b} e_q\,(\bar{q}_{L\beta}\gamma^\mu q_{L\alpha})
\end{aligned}
$$

Magnetic–Penguins (Figure 2.2d):

$$
\begin{aligned}
Q_7 &= \frac{e}{16\pi^2}m_b\bar{s}_{L\alpha}\sigma^{\mu\nu}b_{R\alpha}F_{\mu\nu}\\
Q_8 &= \frac{g}{16\pi^2}m_b\bar{s}_{L\alpha}\sigma^{\mu\nu}T^a_{\alpha\beta}b_{R\beta}G^a_{\mu\nu}\\
Q_{7'} &= \frac{e}{16\pi^2}m_s\bar{s}_{R\alpha}\sigma^{\mu\nu}b_{L\alpha}F_{\mu\nu}\\
Q_{8'} &= \frac{g}{16\pi^2}m_s\bar{s}_{R\alpha}\sigma^{\mu\nu}T^a_{\alpha\beta}b_{L\beta}G^a_{\mu\nu}
\end{aligned}
$$

$\Delta S = 2$ **and** $\Delta B = 2$ **Operators (Figure 2.2e):**

$$
\begin{aligned}
Q(\Delta S = 2) &= (\bar{s}_L\gamma_\mu d_L)(\bar{s}_L\gamma^\mu d_L)\\
Q(\Delta B = 2) &= (\bar{b}_L\gamma_\mu d_L)(\bar{b}_L\gamma^\mu d_L)
\end{aligned}
$$

Semi–Leptonic Operators (Figure 2.2f):

$$
\begin{aligned}
Q_9 &= \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)(\bar{l}\gamma^\mu l) \\
Q_{10} &= \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)(\bar{l}\gamma^\mu\gamma_5 l) \\
Q_{9'} &= \frac{e^2}{16\pi^2}(\bar{s}_R\gamma_\mu b_R)(\bar{l}\gamma^\mu l) \\
Q_{10'} &= \frac{e^2}{16\pi^2}(\bar{s}_R\gamma_\mu b_R)(\bar{l}\gamma^\mu\gamma_5 l) \\
Q_{\nu\bar{\nu}} &= \frac{\alpha_{em}}{4\pi}(\bar{s}_L\gamma_\mu b_L)(\bar{\nu}_L\gamma^\mu\nu_L)
\end{aligned}
$$

where the subscripts L and R refer to left- and right-handed components of the fermion fields. The quark colour indices are represent by α and β and are omitted for colour singlet currents, $F_{\mu\nu}$ and $G^a_{\mu\nu}$ are the electromagnetic and strong interaction tensors and e_q is the electric charge of the relevant quark.

Not all processes are sensitive to all Wilson coefficients even after they have been mixed by the renormalisation. The decay $b \to s\,\gamma$ is only sensitive to the absolute value of C_7, and to a lesser extend to C_8. The transition $b \to s\,\ell^+\ell^-$ is sensitive to the values and signs of C_7, C_9 and C_{10}. Measurements of these processes will therefore impose limits on the permitted values of the Wilson coefficients and translate directly into bounds for extensions to the Standard Model.

2.3 Radiative Decays $b \to s\,\gamma$

The radiative decay is a magnetic dipole transition and is therefore mainly mediated by the operator Q_7. The effective Hamiltonian is given by

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}V_{ts}^*V_{tb}\sum_{j=1}^{8}C_j(\mu)Q_j(\mu) \tag{2.3}$$

with the basis for this decay consisting of the first eight operators in the expansion. The inclusive decay $B \to X_s\,\gamma$ is a perfect test-ground for perturbative QCD. Non-perturbative effects are well under control thanks to the heavy quark expansion and play only a subordinate role. The dominant short-distance QCD corrections enhance the partonic decay rate by a factor of more than two. At next-to-leading-log the theoretical Standard Model branching ratio is predicted to be [9]

$$\mathcal{B}(B \to X_s\,\gamma\,)_{SM} = (3.73 \pm 0.30)\ \times 10^{-4}$$

where the full photon energy spectrum is included. The first exclusive radiative decays observed were $B \to K^*\gamma$ and $B^\pm \to K^{*\pm}\gamma$ by the CLEO collaboration [10]. The inclusive $b \to s\,\gamma$ decay rate was later measured by the CLEO collaboration and their most precise measurement using their complete data set is $\mathcal{B}(B \to X_s\,\gamma) = (3.21 \pm 0.43 \pm^{0.18}_{0.10}) \times 10^{-4}$ [11]. The ALEPH collaboration measured the branching fraction $B \to X_s\,\gamma$ to be $\mathcal{B}(B \to X_s\,\gamma) = (3.11 \pm 0.80 \pm 0.72) \times 10^{-4}$ [12]. Including results from the B-factories and taking into account correlations due to the model dependence a world average [13] of

$$\mathcal{B}(B \to X_s\,\gamma)_{exp.} = (3.34 \pm 0.38) \times 10^{-4}$$

is obtained, which is in good agreement with the Standard Model prediction.

Recently the Belle collaboration reported the observation of the CKM-suppressed inclusive $b \to d\,\gamma$ decay [14] and a measurement of the CKM matrix element $|V_{td}/V_{ts}|$. The obtained results are in a good agreement with the predictions from the Standard Model and a fit to the unitarity triangle. An improved measurement of the decay rate $b \to d\,\gamma$ will help to reduce the currently allowed region of the CKM Wolfenstein parameters ρ and η. It is also important with respect to new physics, because its CKM suppression ($|V_{td}/V_{ts}|^2$) in the standard model might not hold in extended models.

2.4 The Decays $b \to s\,\ell^+\ell^-$

The decays $b \to s\,\ell^+\ell^-$ receive short-distance contributions from W-box diagrams (Figure 2.3 (a)) and from electromagnetic and Z^0-penguin diagrams (Figure 2.3(b)). Long-distance contributions arise from the process $b \to s\,(c\bar{c}) \to s\,\ell^+\ell^-$, with both resonant (mainly J/ψ and $\psi(2S)$) and continuum intermediate ($c\bar{c}$) states (Figure 2.3(c)).

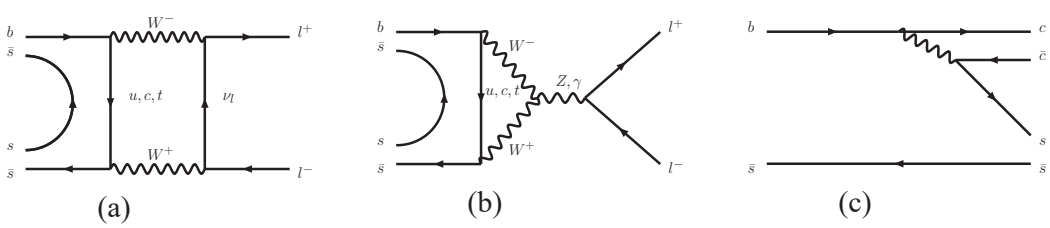

Figure 2.3: Main diagrams contributing to the decay $B_s^0 \to \phi\,\ell^+\ell^-$. The diagrams (a) and (b) are short distance processes and (c) represents the long distance processes leading to $c\bar{c}\,(J/\psi,\ \psi(2S),\ \ldots,\ \psi^{(n)})$ intermediate states.

The operators contributing to the effective Hamiltonian are the current-current operators (Q_1, Q_2), the QCD-penguin operators ($Q_3, \ldots, Q_6$), the magnetic penguin operators

$(Q_{7\gamma}, Q_{8\gamma})$ and the semi-leptonic Z^0 and γ penguin operators (Q_{9V}, Q_{10A}). The effective Hamiltonian will then be

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{ts}^* V_{tb} \sum_{j=1}^{10} C_j(\mu) Q_j(\mu). \tag{2.4}$$

The coefficient $C_j(\mu)$ are systematically calculable in perturbation theory. Standard Model values of the C_j at $\mu = 4.8$ GeV calculated in the naive regularisation scheme at next-to-leading-order (NLO) are given in Table 2.1 [15].

Table 2.1: Standard Model values of the Wilson coefficients at NLO ($C_7^{\text{eff}} \equiv C_7 - C_5/3 - C_6$).

C_1	C_2	C_3	C_4	C_5	C_6	C_7^{eff}	C_9	C_{10}
-0.248	$+1.107$	$+0.011$	-0.026	$+0.007$	-0.031	-0.313	$+4.344$	-4.669

Diagrams where the internal virtual quark is a top quark are - due to the large top quark mass - largely dominant over diagrams with virtual up or charm quarks. Entirely negligible are diagrams with an internal up quark, due to its small mass and the small value of the relevant CKM matrix elements ($|V_{us}^* V_{ub}/V_{ts}^* V_{tb}| < 0.02$). A small correction is caused by the diagrams with an internal charm quark. Using the unitarity constraint $V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0$ together with the smallness of $V_{us}^* V_{ub}$ allows the replacement $V_{cs}^* V_{cb} = -V_{ts}^* V_{tb}$.

The long-distance contributions are generated by the current-current operators Q_1 and Q_2. In the di-muon invariant mass regions around the J/ψ and $\psi(2S)$ resonances, the decay is largely dominated by the long-distance contributions, while away from these resonances, the short-distance contributions dominate. The branching fraction of the long-distance decay is about two to three orders of magnitude larger than the short-distance decay. For the calculation of the non-resonant branching fraction the large distance processes have to be excluded because of a failure of quark-hadron duality. For most final states, the long-distance decay $B \to X_s\, J/\psi$ is well established (with the J/ψ usually seen via its decay to two muons), and the main interest lies in the observation of the short distance processes.

The main contribution to the short-distance processes comes from the photonic penguin diagram, the Z^0-penguin and box diagrams. As a result the Wilson coefficients C_7, C_9 and C_{10} can be completely determined in $b \to s\,\ell^+\ell^-$ transitions. The measurement of the branching fraction of the radiative penguin decay $b \to s\,\gamma$ strongly constraints the magnitude of the effective Wilson coefficients C_7 but not its sign. The latter could, however, be changed due to non-Standard Model contributions without changing the branching fraction of $b \to s\,\gamma$. The determination of the sign of C_7 (as well as the sign of C_9

and C_{10}) in $b \to s\,\ell^+\ell^-$ relies on precise measurements of the distribution of the squared di-lepton momentum and the forward-backward asymmetry in these decays.

The forward-backward asymmetry for leptons as a function of the squared di-lepton mass $m_{\ell\ell}^2 = q^2$ is defined as

$$A_{FB}(q^2) = \frac{\displaystyle\int_0^1 \frac{d^2\Gamma}{dxdq^2}\,dx - \int_{-1}^0 \frac{d^2\Gamma}{dxdq^2}\,dx}{\dfrac{d\Gamma}{dq^2}}\,, \tag{2.5}$$

where $x \equiv \cos\theta$ parameterises the angle between the ℓ^+ and the B meson in the di-lepton centre-of-mass frame.

To measure the Wilson coefficients the di-lepton invariant mass distribution can only be used in two small regions, at low q^2 and at high q^2. This is due to the fact that elsewhere the long distance contributions have a non negligible interference.

2.4.1 Inclusive Decays

Inclusive decays can be calculated perturbatively up to power corrections, using quark-hadron duality and the optical theorem, and summing over all final states. Schematically, the decay width is given by

$$\sum_{X_s} \langle B|Q|X_s\rangle\langle X_s|Q^\dagger|B\rangle = \sum_p \langle B|Q|p\rangle\langle p|Q^\dagger|B\rangle \tag{2.6}$$

where X_s is any final state and the sum over p represents the sum over quarks and gluons. The equality in Eq. 2.6 does not hold on the $c\bar{c}$ resonances as the quark-hadron duality fails there. The matrix element is usually written as an expansion in terms of $1/m_b$. To first order, the hadronic matrix element $\langle X_S|Q_i|B\rangle$ corresponds to the matrix element of the free quarks $\langle s|Q_i|b\rangle$, which can be calculated perturbatively. Corrections due to virtual and real gluons need to be added. Power corrections ($\mathcal{O}(1/m_b^2)$) describe the difference between initial b-quarks and B mesons. This correction causes an estimated suppression of the decay rate of the order of 1.5% [16].

The inclusive decay $B \to X_s\,\ell^+\ell^-$ was first measured by the Belle collaboration [17] and confirmed by the BaBar collaboration [18], using electrons (e) and muons (μ) as leptons in the final states. The measurements agree well with the Standard Model predictions calculated at NLO including long distance contributions. Table 2.2 gives an overview over the theoretical predictions and the measurements from the B factories for the branching fractions with a lower cut of 0.2 GeV/c^2 on the di-lepton invariant mass.

Table 2.2: Standard Model branching fraction predictions for the decays $B \to X_s \ell^+\ell^-$ and experimental measurements for $m_{\ell^+\ell^-} > 0.2$ GeV$/c^2$.

	SM prediction $\mathcal{B}(\times 10^{-6})$	$\mathcal{B}(\times 10^{-6})$ Belle [19]	$\mathcal{B}(\times 10^{-6})$BaBar [18]
$\ell = e, \mu$	4.6 ± 0.8 [20]	$4.11 \pm 0.83^{+0.85}_{-0.81}$	$5.6 \pm 1.5 \pm 0.6 \pm 1.1$
$\ell = e$	6.89 ± 1.01 [21]	$4.04 \pm 1.30^{+0.87}_{-0.83}$	$6.0 \pm 1.7 \pm 0.7 \pm 1.7$
$\ell = \mu$	4.15 ± 0.70 [21]	$4.13 \pm 1.05^{+0.85}_{-0.81}$	$5.0 \pm 2.8 \pm 0.6 \pm 1.0$
$\ell = \tau$	0.26 ± 0.05 [16]	-	-

2.4.2 Exclusive Decays

Theoretical calculations of the decay rates for exclusive decays require the calculation of the decay rate into a specific hadronic final state. For semi-leptonic decays , the matrix elements of the effective Hamiltonian $\mathcal{H}_{\text{eff}}$ can be factorised into a hadronic and a leptonic current:

$$\langle h\, l|Q_j|B\rangle = \langle h|Q_j^H|B\rangle \cdot \langle l|Q_j^L|0\rangle \tag{2.7}$$

with $Q_i = Q_j^H \cdot Q_j^L$. The matrix element of the leptonic current can be explicitly calculated and the hadronic current is expressed in terms of meson form factors. From the theoretical point of view there are large uncertainties arising from decay form factors, making predictions for the exclusive decay rates difficult. As these form factors describe strong interaction effects, they have to be calculated using non-perturbative models, Light Cone-QCD sum rules or vector meson dominance, for each individual configuration of initial and final state. Different methods have been used to calculate the form factors. Usually they are calculated in a small region of q^2 (momentum transfer to the di-lepton system) and then extrapolated to the entire region by means of some parametrisation. Theoretical predictions therefore depend on the model used.

The decays $B \to K\,\ell^+\ell^-$ and $B \to K^*\,\ell^+\ell^-$

The decays $B \to K\,\ell^+\ell^-$ and $B \to K^*\,\ell^+\ell^-$, induced by $b \to s\,\ell^+\ell^-$ transitions at quark level, are experimentally easier to measure than the inclusive processes $B \to X_s\,\ell^+\ell^-$. The Standard Model predictions for different form factor calculations are shown in Table 2.3. The calculations by Ali *et al.* used the Light Cone-QCD sum rule approach, while Melikhov *et al.* used the quark model. As the branching fractions obtained with different form factor calculations differ, it is difficult to determine fundamental parameters of the Standard Model, such as CKM elements, from exclusive decays.

Table 2.3: Branching fractions for $B \to K^{(*)}\,\ell^+\ell^-$ predicted in the framework of the Standard Model.

Mode	Predicted branching fraction $\mathcal{B}[\times 10^{-6}]$	
	Ali *et al.* [22]	Melikhov *et al.* [23]
$B \to K\,\ell^+\ell^-$	$0.57^{+0.16}_{-0.10}$	0.42 ± 0.09
$B \to K^*\,e^+e^-$	$2.3^{+0.7}_{-0.4}$	1.4 ± 0.5
$B \to K^*\,\mu^+\mu^-$	$1.9^{+0.5}_{-0.3}$	1.0 ± 0.4

Experimental results on exclusive decays are also coming from the B-factories. The Belle collaboration reported in 2002 the first observation of $B \to K\,\ell^+\ell^-$ decays [24] and in 2003 the first observation of $B \to K^*\,\ell^+\ell^-$ decays [25]. Shortly after, also the BaBar collaboration reported on measurements of these exclusive decays [26]. In Table 2.4, the most recent measurements are summarised.

Table 2.4: Measurements of the branching fractions for $B \to K^{(*)}\,\ell^+\ell^-$.

Mode	Measured branching fraction $\mathcal{B}[\times 10^{-7}]$	
	Belle [27]	BaBar [28]
$B \to K\,\ell^+\ell^-$	$5.50^{+0.75}_{-0.70} \pm 0.27 \pm 0.02$	$3.4 \pm 0.7 \pm 0.3$
$B \to K^*\,\ell^+\ell^-$	$16.50^{+2.3}_{-2.2} \pm 0.9 \pm 0.04$	$7.8^{+1.9}_{-1.7} \pm 1.2$

In addition to the measurements of branching fractions, the Belle collaboration reported in [27] the first measurement of the forward–backward asymmetry as a function of q^2 for $B \to K^*\,\ell^+\ell^-$. Within their limited statistical precision, the measured asymmetry is consistent with the Standard Model as well as with both signs of C_7. In a more recent publication [29] they now directly fit the ratio of Wilson coefficients, namely C_9/C_7 and C_{10}/C_7 and find that the sign of $C_9 \cdot C_{10}$ has to be negative at a 95% CL. The sign of $C_7 \cdot C_{10}$ could not yet be determined.

The Decay $B_s^0 \to \phi\,\mu^+\mu^-$

The exclusive decay $B_s^0 \to \phi\,\mu^+\mu^-$ has not yet been observed. A measurement of this decay is complementary to measurements of the decays $B \to K\,\ell^+\ell^-$ and $B \to K^*\,\ell^+\ell^-$, which in the Standard Model occur through the same loop processes. Presently, the only existing experimental bound on this decay is given by CDF from a Run I search [30].

With an integrated luminosity of 91 pb^{-1} they set an upper limit at a 95% CL of $\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-) < 6.7 \times 10^{-5}$.

2.5 The Decays $B_{s,d}^0 \to \ell^+\ell^-$

The decays $B_{s,d}^0 \to \ell^+\ell^-$ are dominated by the Z^0-penguin (also called vertical or annihilation penguin) and box diagrams involving top quark exchanges, as shown in Figure 2.4. With long-distance contributions being negligible in this case, they are, after the $b \to s\,\nu\bar{\nu}$ decay, the theoretically cleanest decays in the field of rare B-decays.

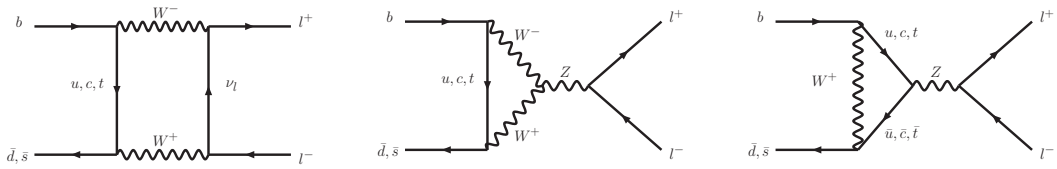

Figure 2.4: Decay processes contributing to $B_{s,d} \to \ell^+\ell^-$ in the Standard Model.

The effective Hamiltonian for $B_{s,d}^0 \to \ell^+\ell^-$ decays is given as follows:

$$\mathcal{H}_{\text{eff}} = -\frac{4G_{\text{F}}}{\sqrt{2}} V_{tb}^* V_{tq} [C_{10}Q_{10} + C_S Q_S + C_P Q_P] \tag{2.8}$$

were q stands either for a s-quark for the B_s^0 or a d-quark in the case of the B_d^0, and the two operators Q_S (scalar) and Q_P (pseudoscalar) are defined as

$$\begin{aligned} Q_S &= \frac{e^2}{16\pi^2}(\bar{q}_{L\alpha} b_{R\alpha})(\bar{\ell}\ell) \\ Q_P &= \frac{e^2}{16\pi^2}(\bar{q}_{L\alpha} b_{R\alpha})(\bar{\ell}\gamma_5\ell). \end{aligned}$$

The diagrams in Figure 2.4 contribute only to the Wilson coefficient C_{10}. There is no contribution from a photonic penguin because of the photon's pure vector coupling to leptons. There are also contributions to the Wilson coefficient C_S from a Standard Model Higgs penguin [31] and to the Wilson coefficient C_P from the would–be neutral Goldstone boson penguin [32], but these contributions to the amplitude are suppressed by a factor of m_b^2/M_W^2 relative to the dominant contributions and can be ignored. For completeness C_S and C_P have been kept because they can be significant in some extensions of the Standard Model.

The Wilson coefficients are evaluated at a high energy scale $\sim \mathcal{O}(M_W)$ and then propagated to the low scale $\sim \mathcal{O}(m_B)$, where the hadronic matrix elements of the operators

are evaluated. This propagation in general leads to QCD corrections enhanced by large logarithms of the ratio of the two scales, which must be re-summed. The operator Q_{10} has zero anomalous dimension because it is a $(V-A)$ quark current, which is conserved in the limit of vanishing quark masses. Therefore the renormalisation group evolution of C_{10} is straightforward. The operators Q_S and Q_P have the same form as a quark mass term and thus have the anomalous dimension of a quark mass. In the Standard Model and many extensions, C_S and C_P are proportional to m_b. The running of these Wilson coefficients is therefore properly taken into account by replacing $m_b(M_W)$ with $m_b(m_B)$ in C_S and C_P.

Evaluating the hadronic matrix elements, the resulting branching ratio for $B_{q=s,d}$ is

$$\begin{aligned}\mathcal{B}(B_q \to \ell^+\ell^-) &= \frac{G_F^2\alpha^2 m_{B_q}^3 \tau_{B_q} f_{B_q}^2}{64\pi^3}|V_{tb}^* V_{tq}|^2 \sqrt{1-\frac{4m_\ell^2}{m_{B_q}^2}} \\ &\times \left[\left(1-\frac{4m_\ell^2}{m_{B_q}^2}\right)\left|\frac{m_{B_q}}{m_b+m_q}C_S\right|^2 + \left|\frac{2m_\ell}{m_{B_q}}C_{10} - \frac{m_{B_q}}{m_b+m_q}C_P\right|^2\right], \end{aligned} \tag{2.9}$$

where τ_{B_q} is the B_q lifetime, f_{B_q} is the B_q decay constant, normalised according to $f_\pi = 132$ MeV, and the Wilson coefficients C_S and C_P are retained for completeness.

The Standard Model decay amplitude is given by the Wilson coefficient

$$C_{10} = -Y(x_t)/\sin^2\theta_W = -4.2. \tag{2.10}$$

The function $Y(x_t)$ represents the dependence on the top mass ($x_t = m_t^2/M_W^2$) and is given by

$$Y(x_t) = \eta_\gamma \cdot Y_0(x_t) \tag{2.11}$$

$$Y_0(x_t) = \frac{x_t}{8}\left[\frac{x_t-4}{x_t-1} + \frac{3x_t}{(x_t-1)^2}\ln x_t\right] \tag{2.12}$$

where η_Y summarises the NLO corrections with $\eta_Y = 1.012$.

Purely leptonic decays are helicity-suppressed by a factor m_l^2, as, by conservation of angular momentum, the spin-less B^0 forces both the lepton and the anti-lepton to be in the same helicity state.

The Standard Model leptonic branching fractions have been calculated including QCD corrections in Ref. [33]. The latest Standard Model predictions [34][1] for the various branching fractions are given in Table 2.5, where the errors are dominated by non-perturbative uncertainties. For the calculation of the branching fractions of B_s^0 mesons, $\Delta M_s = (18.0 \pm 0.5)$/ps was assumed. The experimental upper limits are given in Table 2.6.

[1] In Ref. [34] only the Standard Model prediction for the decay $B_s^0 \to \mu^+\mu^-$ is given. With the given formulae the other Standard Model predictions have been calculated.

Table 2.5: Standard Model branching fraction predictions for the decays $B^0_{s,d} \to \ell^+\ell^-$.

	$\mathcal{B}(B^0_d \to \ell^+\ell^-)$	$\mathcal{B}(B^0_s \to \ell^+\ell^-)$
$\ell = e$	$(2.40 \pm 0.34) \times 10^{-15}$	$(8.15 \pm 1.29) \times 10^{-14}$
$\ell = \mu$	$(1.00 \pm 0.14) \times 10^{-10}$	$(3.42 \pm 0.54) \times 10^{-9}$
$\ell = \tau$	$(2.90 \pm 0.41) \times 10^{-8}$	$(9.86 \pm 1.55) \times 10^{-7}$

Table 2.6: Experimental upper limits for the decays $B^0_{s,d} \to \ell^+\ell^-$ at a 90% CL.

	$\mathcal{B}(B^0_d \to \ell^+\ell^-)$	$\mathcal{B}(B^0_s \to \ell^+\ell^-)$
$\ell = e$	$< 6.1 \times 10^{-8}$ [35]	$< 5.4 \times 10^{-5}$ [36]
$\ell = \mu$	$< 3.2 \times 10^{-8}$ [37]	$< 1.2 \times 10^{-7}$ [37]
$\ell = \tau$	$< 3.1 \times 10^{-3}$ [38]	$< 5.0 \times 10^{-2}$ [39]

Since the ratio of the decay constants $f_{B^0_d}$ and $f_{B^0_s}$ can be calculated reliably by non-perturbative methods or measured in leptonic decays, the ratio

$$\frac{\mathcal{B}(B^0_s \to \mu^+\mu^-)}{\mathcal{B}(B^0_d \to \mu^+\mu^-)} = \frac{\tau(B^0_s) m_{B^0_s} f^2_{B^0_s} |V_{ts}|^2}{\tau(B^0_d) m_{B^0_d} f^2_{B^0_d} |V_{td}|^2} \tag{2.13}$$

will provide a useful measurement of $|V_{ts}|^2/|V_{td}|^2$ with a smaller theoretical uncertainty than the individual branching ratio measurements. As this ratio is larger than one, it can be seen that the branching fraction for B^0_d-decays is expected to be smaller (by about one order of magnitude) than for B^0_s-decays.

The limits on the branching ratios $B^0_s \to \mu^+\mu^-$ and $B^0_d \to \mu^+\mu^-$ can be compared to the values expected in the Standard Model. Currently the 90% CL limit on $B^0_s \to \mu^+\mu^-$ is approximately a factor of 35 larger than the Standard Model branching ratio, while the limit on $B^0_d \to \mu^+\mu^-$ is approximately 320 times larger. This makes it rather clear why a search for the decay $B^0_s \to \mu^+\mu^-$ is the preferred choice.

2.5.1 Beyond the Standard Model

The decay amplitude of $B^0_{s,d} \to \mu^+\mu^-$ can be significantly enhanced in some extensions of the Standard Model. The effects of additional non-Standard Model diagrams would

modify the magnitude and sign of the Wilson coefficients and could even introduce new operators in more exotic models.

Models with more than one Higgs doublet (HDM)

In general models with more than one Higgs doublet, flavour-changing neutral currents can occur at tree-level leading to a severe enhancements of $K - \overline{K}$ and $B - \overline{B}$ mixing not seen in experimental data [4]. These enhancements are suppressed "naturally" if all the down-type quarks acquire their masses through their coupling to the same Higgs boson doublet, H_d, and all the up-type quarks through their coupling to a second Higgs boson doublet, H_u. Under the requirement of this "natural flavour conservation", different configurations are possible for the coupling of the two Higgs doublets H_u and H_d of a 2 Higgs doublet model (2HDM) to quarks. In the type-I 2HDM all quarks couple only to one Higgs doublet and not to the other. In the type-II 2HDM the down-type right-handed quark singlets couple to H_d while the up-type right-handed quark singlets couple to H_u. The most general 2HDM without natural flavour conservation, in which the right handed fermion singlets with different values of hypercharge couple to Higgs doublets, is called the type-III 2HDM.

The Minimal Supersymmetric Standard Model (MSSM) is a supersymmetric extension of the Standard Model with a minimal number of new particles. For each particle, there is a superpartner with the same internal quantum numbers, but with a spin that differs by half a unit. The Higgs sector contains two doublets, corresponding to five physical particles: one CP-odd neutral Higgs boson, A, two neutral CP-even bosons, h^0 and H^0, and the charged $H^\pm$. At tree level, the Higgs masses can be described by two parameters, namely $\tan\beta$, the ratio of the two neutral Higgs field vacuum expectation values, and m_A. The low energy effective theory of the MSSM is a type-III 2HDM.

In this model the Wilson coefficients C_S and C_P in Eq. 2.9, neglected in the Standard Model, have to be taken into account due to Higgs penguin diagrams. The scalar and pseudoscalar Wilson coefficients $C_{S,P}$ grow like $\tan^3\beta$, and therefore are roughly $(m_b^2/M_W^2)\tan^3\beta$ times bigger than C_{10} and dominate for large values of $\tan\beta$. Other electroweak SUSY box and Z-penguin contributions grow with at most the second power of $\tan\beta$ and thus are subdominant in this region. This gives

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \propto (\tan\beta)^6, \tag{2.14}$$

leading to an enhancement of up to three orders of magnitude [40, 41] compared to the Standard Model, even if the MSSM with minimal flavour violation (MFV) is considered in which the CKM matrix is the only source of flavour violation. An observation of $B_s^0 \to \mu^+\mu^-$ would then yield an upper bound on the heaviest mass in the MSSM Higgs sector [42] if MFV applies.

In type-II 2HDM, charged Higgs bosons can be exchanged in addition to W bosons in the loop diagrams. In the limit of large $\tan\beta$ the 2HDM contributions for the decay

$B^0_{s,d} \to \mu^+\mu^-$ are significant. The Wilson coefficients C_S and C_P receive contributions from the box and penguin diagrams involving W and H^+ bosons. There are also contributions due to fermion self-energy diagrams with a neutral Higgs boson exchange. In this model there are no new contributions to C_{10}, which therefore retains its Standard Model value. The obtained Wilson coefficients in the type-II 2HDM are

$$C_S = C_P = \frac{m_l}{2M_W^2} \tan^2\beta \frac{\log r}{r-1}$$

with $r = x_{H^+}/x_t = M^2_{H^+}/m_t^2$. The dependence on the masses of the neutral Higgs bosons from penguin and self-energy diagrams drop out without invoking any relations between the mixing angle and the Higgs masses. The branching fraction depends only on two 2HDM parameters, the charged Higgs mass M_{H^+} and $\tan\beta$, with the branching fraction growing like $\tan^4\beta$ [43]. The predicted value for $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ as a function of the charged Higgs mass M_{H^+} and for various values of $\tan\beta$ is shown in Figure 2.5. For large $\tan\beta$ and light M_{H^+}, the 2HDM contribution dominates, enhancing the branching fraction significantly. As the 2HDM contribution becomes smaller due to decreasing $\tan\beta$ or increasing M_{H^+} the branching fraction drops, eventually falling below the Standard Model prediction due to destructive interference.

$B^0_s \to \mu^+\mu^-$ and dark matter

The general superpotential of the MSSM contains terms that violate baryon and lepton numbers. In order to cure this, a new symmetry, called R-parity, is introduced. It is a multiplicative quantum number which is positive for all particles of the Standard Model and negative for their superpartners. The R-parity quantum number is given by

$$R = (-1)^{3(B-L)+2S} \tag{2.15}$$

for a particle with spin S and baryon- and lepton numbers B and L. This symmetry has a great influence on the phenomenology of the MSSM, since it requires that supersymmetric particles are always produced in pairs. The conservation of R-parity means that the Lightest Supersymmetric Particle (LSP) is stable. In supergravity models this is the neutralino. This particle could turn out to be a good candidate for cold dark matter (DM). Recent data from the WMAP satellite indicate a dark matter density in the universe of $\Omega_{\rm DM}h^2 \simeq (0.095 - 0.13)$ at a $\pm 2\sigma$ level [46] which could be dominated by the relic density of neutralinos within SUSY models.

There has been experimental progress in the direct detection of neutralino dark matter through (in)elastic scattering on various nuclei. Such experiments can be sensitive to a neutralino with mass $O(100$ GeV/c$^2)$, which is in the mass region predicted by various supergravity scenarios. A particularly interesting region of the neutralino nucleon cross section $\sigma_{\tilde{\chi}p}$ in which one experiment (DAMA) has claimed a dark matter signal [47] has

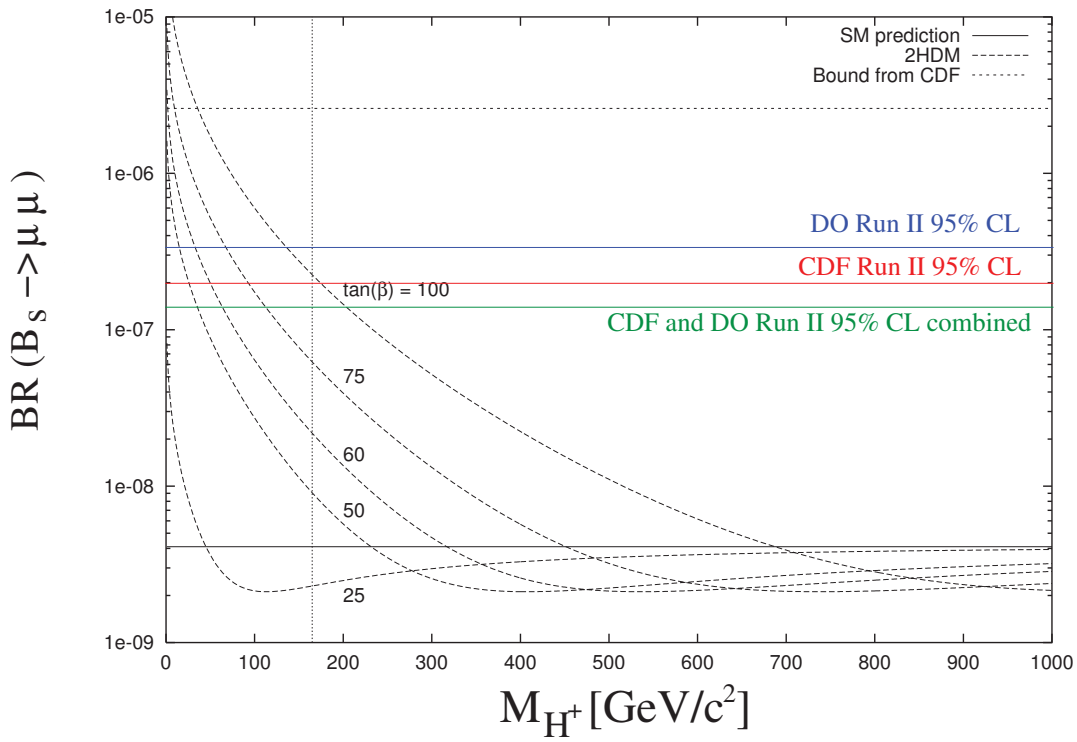

Figure 2.5: The branching fraction $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ in the 2HDM as a function of the charged Higgs mass M_{H^+} and various values of $\tan\beta$ [43]. For comparison the Standard Model prediction and the current experimental bounds [37, 44, 45] are shown where the result from DØ represents the limit obtained in the analysis presented in this thesis. The vertical line is the lower bound on M_{H^+} in the type-II 2HDM from $b \to s\,\gamma$ decays.

been recently excluded by the CDMS cryogenic dark matter search experiment [48] in the range of

$$\sigma_{\tilde{\chi}p} = (10^{-6} - 10^{-5})\ \mathrm{pb}. \tag{2.16}$$

For a given value of $\tan\beta$ there is a strong correlation between the dark matter scattering cross section of the neutralino $\sigma_{\tilde{\chi}p}$ and $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ [49]. In the large $\tan\beta$ limit, the heavy neutral Higgs H^0 exchange contribution to the dark matter scattering becomes important because of its enhanced coupling to down type quarks. This is relevant for the heavy Higgs interaction with the strange quark content inside nucleons, and the dark matter scattering cross section becomes enhanced. Therefore, the dark matter scattering amplitude increases linearly as $\tan\beta$ increases, and decreases as m_A increases. On the left side in Figure 2.6, the correlation between $\sigma_{\tilde{\chi}p}$ and $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ within mSUGRA

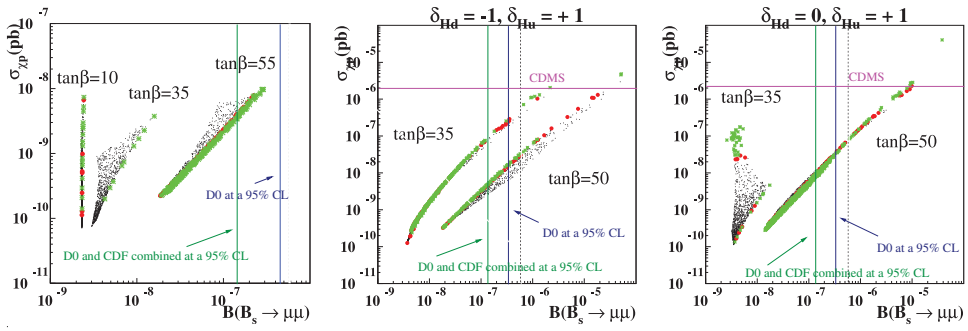

Figure 2.6: $\sigma_{\tilde{\chi}p}$ vs. $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ within mSUGRA with universal Higgs mass parameters (left side) and non-universal Higgs mass parameters (centre and right side) for various values of $\tan\beta$ [49]. The black dots are for $\Omega_\chi h^2 \geq 0.13$ (greater than 2σ of the preferred density of WMAP data), red dots for $0.095 \leq \Omega_\chi h^2 \leq 0.13$ (within 2σ of WMAP data) and green dots for $\Omega_\chi h^2 \leq 0.095$ (less than 2σ of WMAP data). Also indicated are the current experimental bounds from the analysis presented in this thesis [44] as well as the combined limit with CDF [37] as vertical lines. The experimental bound on $\sigma_{\tilde{\chi}p}$ from CDMS[48] is also shown in the centre and right plot.

is shown for various values of the universal Higgs mass and for $\tan\beta = 10$, 35 and 55, respectively. For large $\tan\beta$, there is a strong correlation between the two observables.

Giving up the restriction on the universal mass parameter for the Higgs doublet one can assume the universal soft masses as

$$m_{H_u}^2 = m_0^2\,(1+\delta_{H_u}), \quad m_{H_d}^2 = m_0^2\,(1+\delta_{H_d}), \tag{2.17}$$

whereas other scalar masses are still universal. The correlation between $\sigma_{\tilde{\chi}p}$ vs. $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ for two parameter sets of δ_{H_u} and δ_{H_d} is shown in the centre and on the right side of Fig. 2.6. The current experimental limit on $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ already puts a strong constraint on $\sigma_{\tilde{\chi}p}$ in the large $\tan\beta$ region, even stronger than the CDMS limit. Thus, the decay $B_s^0 \to \mu^+\mu^-$ could give information not only on SUSY breaking mediation mechanisms but also places a severe limit on the neutralino dark matter scattering cross section within a large class of supergravity models in the large $\tan\beta$ region.

mSUGRA

Even if R-parity conservation is imposed on the MSSM there is still a large number of parameters. The construction of a model with only a few parameters is therefore desirable. An elegant solution to this problem is the minimal supergravity (mSUGRA) scenario. In mSUGRA, local supersymmetry (supergravity) is spontaneously broken in the hidden sector, where the minimum of the scalar potential violates supersymmetry. The superpartner of the graviton acquires a mass from the so called super Higgs effect. This

is a phenomenon similar to the Higgs mechanism through which the W boson gains mass by absorbing the charged component of a Higgs doublet in the Glashow-Salam-Weinberg model. Supersymmetry breaking is mediated to the observable sector via gravity interactions. In this model all supersymmetric parameters relate to just five real quantities: the universal scalar and gaugino masses M_0 and $M_{1/2}$, the trilinear term A_0, $\tan\beta$ and the sign of μ, where μ is the Higgsino mass parameter. The first three quantities are defined at the grand unification scale, the others at the electroweak scale.

An enhancement of $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ is correlated [40, 50] with a sizeable positive shift in $(g-2)_\mu$, the anomalous magnetic momentum of the muon. Recently, an excess of the muon anomalous magnetic moment $a_\mu = (g-2)_\mu/2$ has been reported [51]. Taking into account e^+e^- collision data in order to calculate the $\pi^+\pi^-$ spectral functions for the hadronic vacuum polarisation for the magnetic moment of the muon, the measurement corresponds to a 2.6σ deviation [52] from the Standard Model predicted value. In the mSUGRA interpretation this anomaly implies $\mu > 0$. For not too large values of M_0, $M_{1/2} \lesssim$500 GeV/c^2 and $A_0 \simeq 0$ GeV/c^2 the branching fraction for $B^0_s \to \mu^+\mu^-$ can be approximated by

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) \approx 10^{-6} \cdot \tan^6\beta \frac{M^2_{1/2}\,\mathrm{GeV}^4}{(M^2_{1/2}+M^2_0)^3} \tag{2.18}$$

outside the vicinity of $M_{1/2} = 0.4M_0$. An estimation of the supersymmetric contributions yields

$$(\delta a_\mu)_{SUSY} \propto \tan\beta \frac{f(M_0)}{M_{1/2}} \tag{2.19}$$

with $\delta a_\mu = a^{exp}_\mu - a^{SM}_\mu$. This shows that both $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $(\delta a_\mu)_{SUSY}$ grow with $\tan\beta$ and decrease with increasing $M_{1/2}$. If the SUSY explanation for $\delta a_\mu \neq 0$ holds and is due to a large value of $\tan\beta$, one would expect $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ to be enhanced by more than one order of magnitude. For a small value of $M_{1/2}$ and a moderate value of $\tan\beta$, however, the enhancement is expected to be smaller.

Grand Unified Theories

The minimal SO_{10} supersymmetric model [MSO$_{10}$SM] based on the SO(10) Lie group [53, 54, 55] is one candidate for a grand unified theory (GUT). In this model, quarks and leptons, and their supersymmetric partners, of one family reside in the **16** dimensional representation, while the two Higgs doublets of the MSSM reside in one **10** dimensional representation. It is assumed that the couplings obey a hierarchical mass matrix for the first two generations and their mixing with the third generation, due to effective higher dimensional operators. Soft SUSY breaking parameters are a universal gaugino mass $M_{1/2}$, a universal squark and slepton mass[2] m_{16}, a universal scalar Higgs

[2] SO_{10} does not require all sfermions to have the same mass.

mass m_{10}, and a universal parameter A_0. In addition there is the supersymmetric (soft SUSY breaking) Higgs mass parameter μ ($B\mu$), that is interchangeable with $\tan\beta$. All these parameters are independent. To fit the low energy electroweak data, including the third generation fermion masses, it has been shown that $A_0,\ m_{10},\ m_{16}$ must satisfy the constraints [53, 54]

$$A_0 \approx -2\, m_{16}; \quad m_{10} \approx \sqrt{2}\, m_{16} \tag{2.20}$$

$$m_{16} > 1.2 \text{ TeV/c}^2; \quad \mu,\ M_{1/2} \ll m_{16} \tag{2.21}$$

with

$$\tan\beta \approx 50. \tag{2.22}$$

In this model, the cosmological dark matter density $\Omega_\chi h^2$ and $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ both depend on m_A and $\tan\beta$. In Figure 2.7 contours of constant χ^2 for m_{16} = 3 TeV/c^2 and

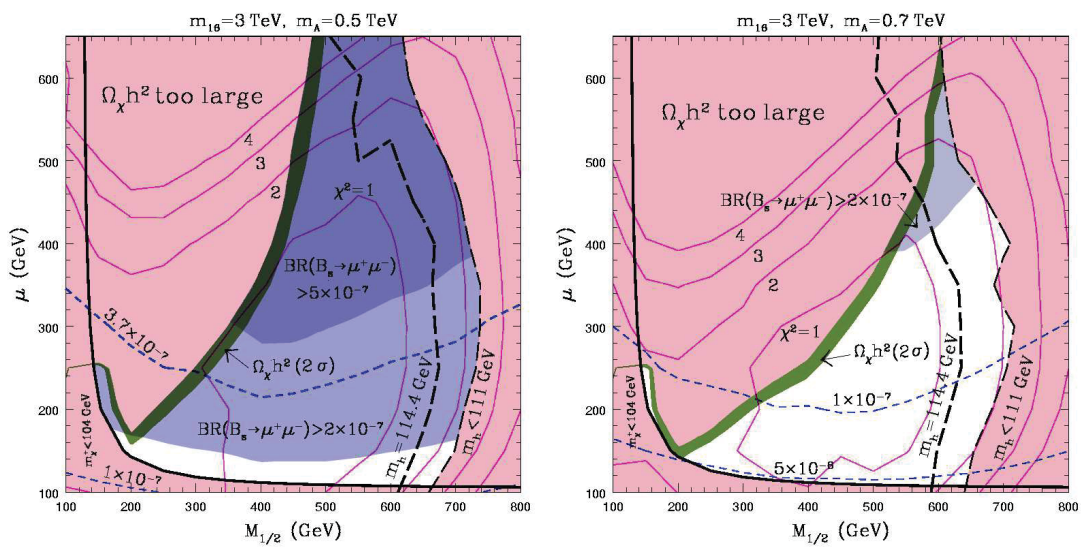

Figure 2.7: Contours of constant χ^2 for m_{16} = 3 TeV/c^2 and m_A = 500 GeV/c^2 (left side) and 700 GeV/c^2 (right side), indicating the allowed region for $B_s^0 \to \mu^+\mu^-$ [55].

m_A = 500 GeV/c^2 (left plot) and m_A = 700 GeV/c^2 (right plot) are shown. The red (light shaded) regions are excluded by $m_{\chi^+} < 104$ GeV/c^2 (below and to the left of the solid black curve), $m_h < 111$ GeV/c^2 (on the right) and by $\Omega_\chi h^2 > 0.129$. To the right of the thick broken black line one has $m_h < 114.4$ GeV/c^2 which is the CERN e^+e^- Collider (LEP) exclusion limit on the Higgs mass. The green (darkest shaded) band corresponds to the preferred 2σ range $0.094 < \Omega_\chi h^2 < 0.129$, while the white regions below it

correspond to $\Omega_\chi h^2 < 0.094$. The region excluded by the DØ experimental bound on $\mathcal{B}(B^0_s \to \mu^+\mu^-) < 5.0 \times 10^{-7}$ at 95% CL is marked in dark blue (dark shaded), while the region affected by the new preliminary CDF bound $\mathcal{B}(B^0_s \to \mu^+\mu^-) < 2.0 \times 10^{-7}$ at 95% is marked in light blue (light shaded). Contours of constant $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ are given by the blue dashed lines. The branching ratio $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ depends on the value of the CP-odd Higgs mass m_A, as m_A^{-4}. For m_A = 500 GeV/c^2 and for acceptable values of $\Omega_\chi h^2$ (Figure 2.7 (left)) the branching ratio $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ is below the published DØ bound, but is almost excluded by the preliminary CDF bound. Increasing the CP-odd Higgs mass to $m_A = 1.25$ TeV/c^2 shifts the allowable $B^0_s \to \mu^+\mu^-$ decay rate down to $\mathcal{B}(B^0_s \to \mu^+\mu^-) > 10^{-8}$. Hence, acceptable regions of parameter space in the minimal SO(10) model lead to predictions of observable rates for $B^0_s \to \mu^+\mu^-$ at the Tevatron.

Other models

In R-parity violating supersymmetry, lepton and baryon number violation can occur. Therefore also lepton violating terms have to be considered when calculating the decay $B^0_s \to \mu^+\mu^-$. Unlike in R-parity conserving models, the SUSY contribution can now enter at tree level, and can become considerable [56].

The idea of extra dimensional (ED) models is that extra compactified space-like dimensions at a scale larger than the Planck scale of $\sim 10^{-33}$ cm [57] exist. This idea led to exciting perspectives for particle phenomenology, cosmology and string theory. In ED models, Standard Model particles are allowed to travel in all available extra dimensions. Such a higher-dimensional theory can be described as an effective theory in four dimensions which contains an infinite tower of Kaluza-Klein modes of the higher dimensional field. The decay $B^0_q \to \ell^+\ell^-$ will proceed via the same box and penguin diagrams as in the Standard Model. However, each boson line will now involve several Kaluza-Klein modes. An enhancement of the branching ratio depends on the compactification scale and the number of extra dimensions [58].

In technicolour models, there is no fundamental Higgs scalar particle. Instead, additional scalar particles (techni-particles) are introduced to connect the technicolour condensate to the ordinary fermions. The branching fraction for $B^0_s \to \mu^+\mu^-$ can be enhanced in this models up to a factor of five over the Standard Model prediction [59].

Chapter 3

Experimental Environment

The Tevatron accelerator at the *Fermi National Accelerator Laboratory* (FERMILAB) in Batavia, Illinois, USA, is a proton-antiproton ($p\overline{p}$) collider. During the first data-taking period from 1992 to 1996 (Run I), the Tevatron experiments CDF and DØ each collected about 125 pb^{-1} of $p\overline{p}$ collision data at a center-of-mass energy of 1.8 TeV, leading to the discovery of the top quark and the measurement of its mass, a precision measurement of the mass of the W boson, detailed analyses of gauge boson couplings, studies of jet production and vastly improved limits on new phenomena, such as leptoquarks and supersymmetric particles, among many other accomplishments.

After a break of about five years, during which the accelerator and the experiments have undergone significant upgrades, the second phase of data taking, called Run II, started in 2001. The Tevatron is expected to deliver between 4 fb^{-1} and 9 fb^{-1} by the year 2009.

3.1 Tevatron

Today, the Tevatron accelerator, operated at a center-of-mass energy of 1.96 TeV, is the most energetic particle collider in the world and will remain so until the Large Hadron Collider (LHC) at CERN will be turned on.

The main ring has a circumference of about $6\,\mathrm{km}$. The accelerated protons collide with antiprotons moving in the opposite direction in the synchrotron ring. The collisions take place in two interaction regions which are surrounded by the two detectors CDF and DØ.

3.1.1 The Tevatron Accelerator Complex

Generally, collision experiments involve in three different steps: the initial production and injection of the particles, a chain of successive acceleration and finally the collision itself. The Tevatron ring is only the last part of a cascade of pre-accelerators at FERMILAB as schematically illustrated in Figure 3.1.

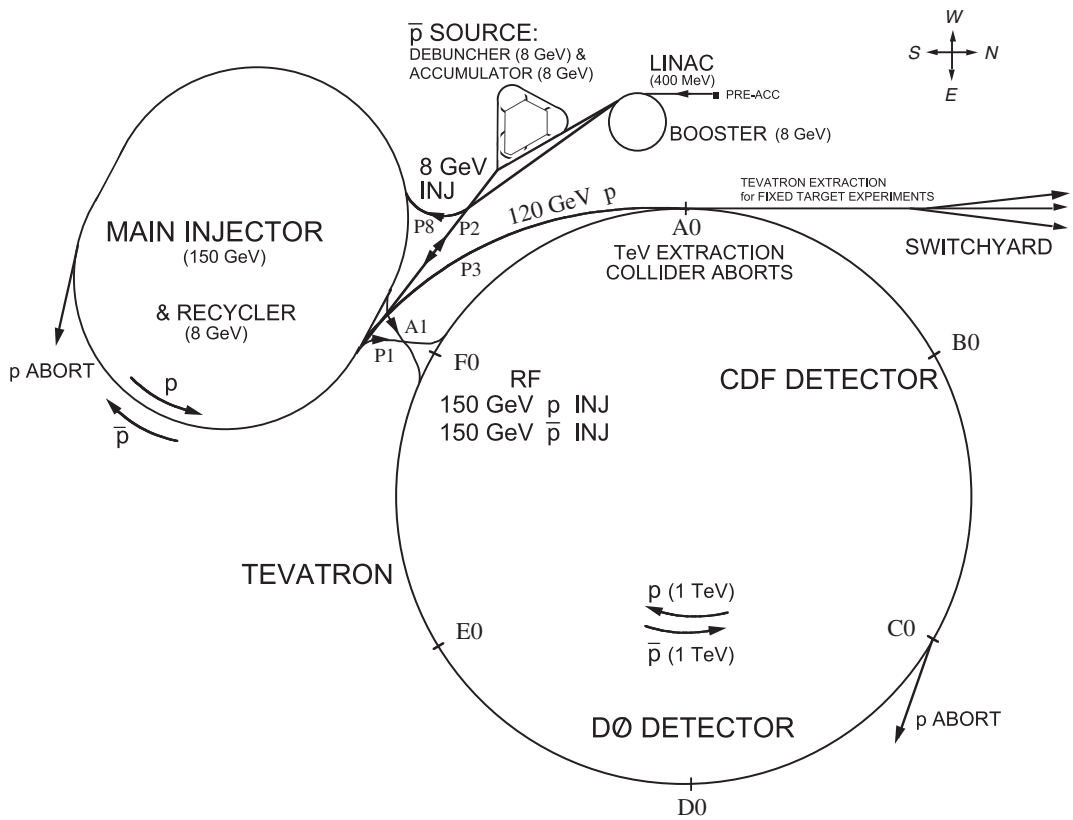

Figure 3.1: The accelerator facilities at FERMILAB with the two multi purpose detectors CDF and DØ.

The protons used in the collisions are extracted from negatively charged hydrogen ions. The ions are accelerated to 750 keV by a Cockroft-Walton accelerator and injected into a 165 m long linear accelerator which boosts their energy to 400 MeV. The ions are then stripped off their electrons as they pass through a sheet of graphite and are injected into the Booster, a synchrotron which brings their energy to 8 GeV.

Protons from the Booster are sent to the Main Injector, where they are further accelerated to 150 GeV. Antiprotons used in the collisions are collected from the interaction

products of a fraction of the 120 GeV proton beam incident on a Nickel-Copper target. The collision energy is chosen such that the energy spectrum of the emerging antiprotons ($\bar{p}$) has its maximum at about 8 GeV. On average, around 50 000 protons are necessary to produce one antiproton in the required energy range. The Tevatron complex is designed in a way that new antiprotons can be produced in parallel to collisions taking place in the main Tevatron ring. The produced antiprotons are cooled and debunched in the Debuncher and Accumulator, and once the number of antiprotons is sufficiently large they are passed to the Main Injector where they are accelerated to 150 GeV for transfer to the Tevatron. Once the protons and antiprotons are injected into the Tevatron, they are accelerated to their final energy of 980 GeV before colliding at the centre of the CDF and DØ detector.

Protons and antiprotons are stored in bunches, with 36 bunches separated by 396 nsec. The 36 bunches in the proton and anti-proton beam are organised into three super-bunches, with a 2 μs gap between each super-bunch.

Figure 3.2 shows the integrated luminosity per week and the total integrated luminosity accumulated in Run II from May 2001 until August 2005. The initial luminosity for each fill is shown in Figure 3.3.

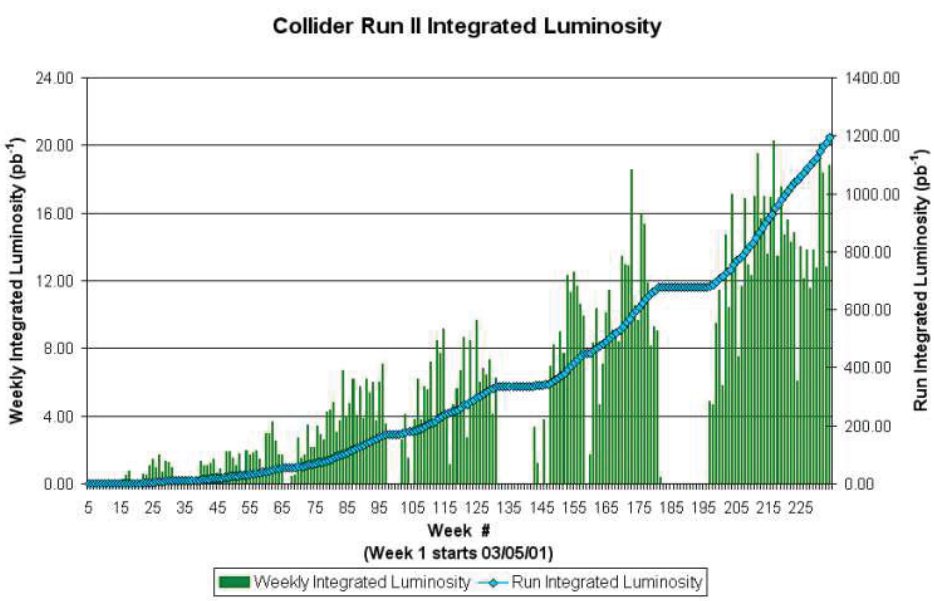

Figure 3.2: The integrated luminosity per week and total integrated luminosity for Run II from May 2001 until August 2005 in pb^{-1}.

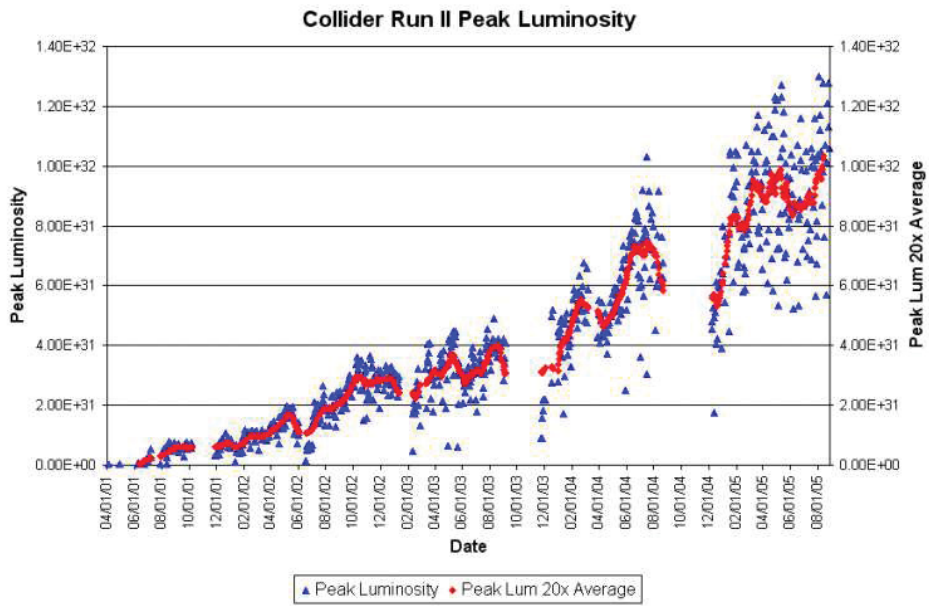

Figure 3.3: The initial luminosity for each fill for the Tevatron Run II from May 2001 until August 2005.

3.2 The DØ Detector

The DØ detector was built in the late 1980's and early 1990's [60], operated from 1992 to 1996 in Run I of the Tevatron, and was then upgraded to handle the increased luminosities in Run II of the Tevatron [61, 62]. It is a large multi-purpose detector and has been built to measure precisely muons, electrons, jets and missing transverse energy. To serve this purpose, the DØ detector consists of three major subsystems. At the core of the detector, a magnetised tracking system records precisely the angles of charged particles and measures their transverse momenta. A hermetic, finely grained Uranium and Liquid Argon calorimeter measures the energies of electromagnetic and hadronic showers, and a muon spectrometer detects and measures the momenta of escaping muons. Figure 3.4 shows a schematic overview of the detector.

3.2.1 Coordinate System

The Cartesian coordinate frame of DØ is defined with its origin at the nominal $p\overline{p}$ interaction point and the z-coordinate pointing along the proton beam-line (the protons circle

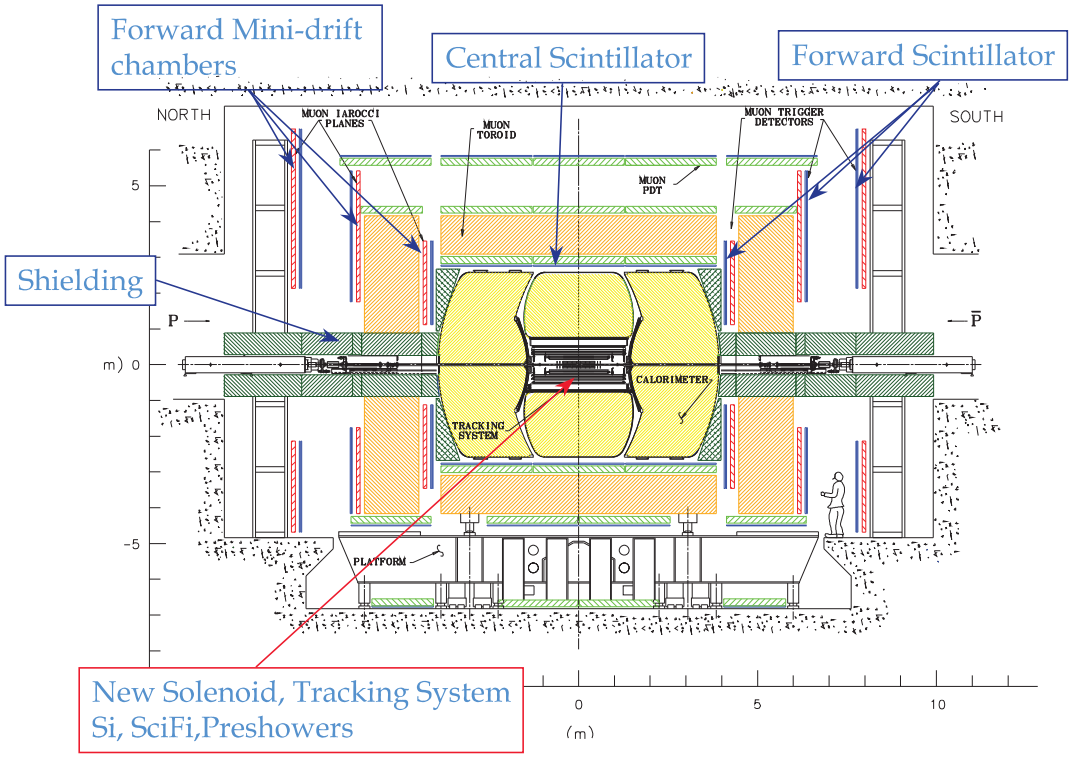

Figure 3.4: Cross-sectional view of the DØ Run II detector.

clockwise in the Tevatron). The x-coordinate is pointing towards the centre of the Tevatron ring and the y-coordinate is pointing upwards. Polar coordinates (r,ϕ) are defined by:

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \phi &= \arctan \frac{y}{x} \end{aligned} \tag{3.1}$$

A convenient parameter to express the polar angle θ at hadron colliders is the pseudo-rapidity η, defined as

$$\eta = -\ln \tan \frac{\theta}{2}\,, \tag{3.2}$$

as the multiplicity of high energy particles in QCD production is roughly constant in η.

3.2.2 Luminosity System

The primary purpose of the Luminosity Monitor (LM) [63] is to make an accurate determination of the Tevatron collider luminosity at the DØ interaction region [64].

The LM detector consists of two arrays of twenty-four plastic scintillation counters with photomultiplier readout. A schematic drawing of an array is shown in Figure 3.5. The arrays are located in front of the end calorimeters at $z = \pm 140$ cm, and occupy the

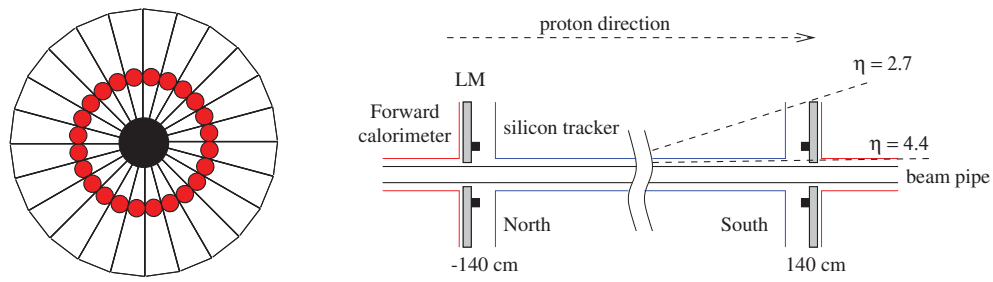

Figure 3.5: Luminosity Monitor layout. The $r - \phi$ view is shown on the left, the $r - z$ view of the two arrays is shown on the right.

region between the beam pipe and the Forward Preshower Detector (see Sect. 3.2.4). The counters are 15 cm long and cover the η range $2.7 < |\eta_{det}| < 4.4$.

The luminosity measurement is accomplished by detecting inelastic $p\overline{p}$ collisions in the dedicated LM detector. The luminosity $\mathcal{L}$ is calculated as

$$\mathcal{L} = \frac{N}{\varepsilon A \sigma_{inel}} \tag{3.3}$$

The known inelastic cross section σ_{inel} has to be corrected for the efficiency ε and the acceptance A of the LM detector for inelastic $p\overline{p}$ collisions.

Multiple $p\overline{p}$ collisions can occur in a single beam crossing. The number of interactions per bunch crossing is given by Poisson statistics. For the instantaneous luminosities at which the data for this analysis were recorded, an average[1] of 0.5 inelastic $p\overline{p}$ collisions are expected per event in addition to the hard-scatter interaction. Collision products arrive at each set of scintillators roughly in coincidence, while products from beam halo interactions pass through the detectors at distinctly separated times. Time-of-flight information from the two luminosity arrays and the z vertex distribution (see Figure 3.6) is utilised to separate these processes. The number of events N is corrected for these two effects.

[1] The average refers to the mean of a Poisson distribution.

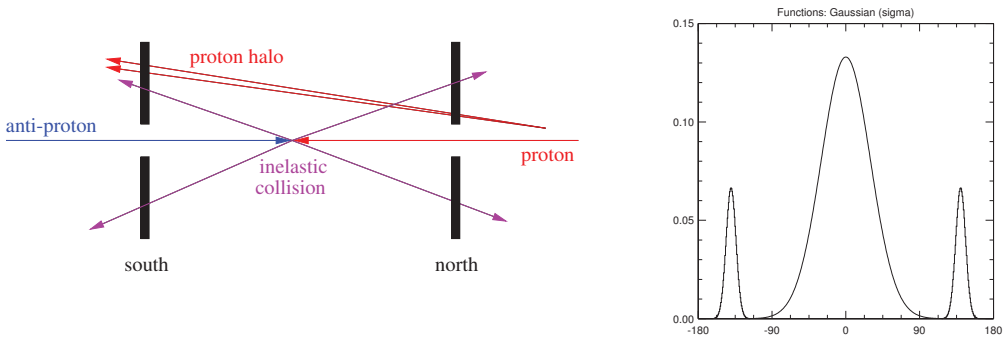

Figure 3.6: The sketch on the left shows the differentiation between inelastic collisions and beam halo. Expected z vertex distribution for inelastic collisions, centred at $z = 0$ cm, p halo centred at $z = -140$ cm and $\bar{p}$ halo centred at $z = 140$ cm (right).

3.2.3 Central Tracking System

The central tracking system is surrounded by a solenoid magnet which provides a nearly uniform (magnetic) field of $B = 2$ T parallel to the beam axis. Charged particles produced in the collision are bent around the field lines. The radius r of the curvature allows for a measurement of the transverse momentum through:

$$p_T[\mathrm{GeV}] = 0.3 \cdot r[\mathrm{m}] \cdot B[\mathrm{T}]\,. \tag{3.4}$$

Closest to the beam pipe itself is the Silicon Microstrip Tracker (SMT), which allows for the precision measurements that are crucial for an accurate measurement of impact parameters and the identification of secondary vertices. Surrounding the SMT is the Central Fibre Tracker (CFT), consisting of 16 layers of scintillating fibre. The CFT extends to a radius of 50 cm, giving a lever arm long enough to provide sufficient transverse momentum resolution. The DØ central tracking system is illustrated in Figure 3.7.

Silicon Microstrip Tracker

Charged particles passing through the 300 μm thick wafers of n-type silicon which constitute the SMT [65] produce pairs of electrons and holes. The ionised charge is collected by strips of p^+-type or n^+-type silicon strips with a pitch between $\sim 50\,\mu$m and $\sim 150\,\mu$m.

The length of the interaction region ($\sigma \approx 25$ cm) determines the length of the device in z. With a long interaction region, it is difficult to deploy detectors such that the tracks are generally perpendicular to detector surfaces for all η. This led to a design using barrel modules interspersed with disks in the centre and assemblies of disks in the forward and

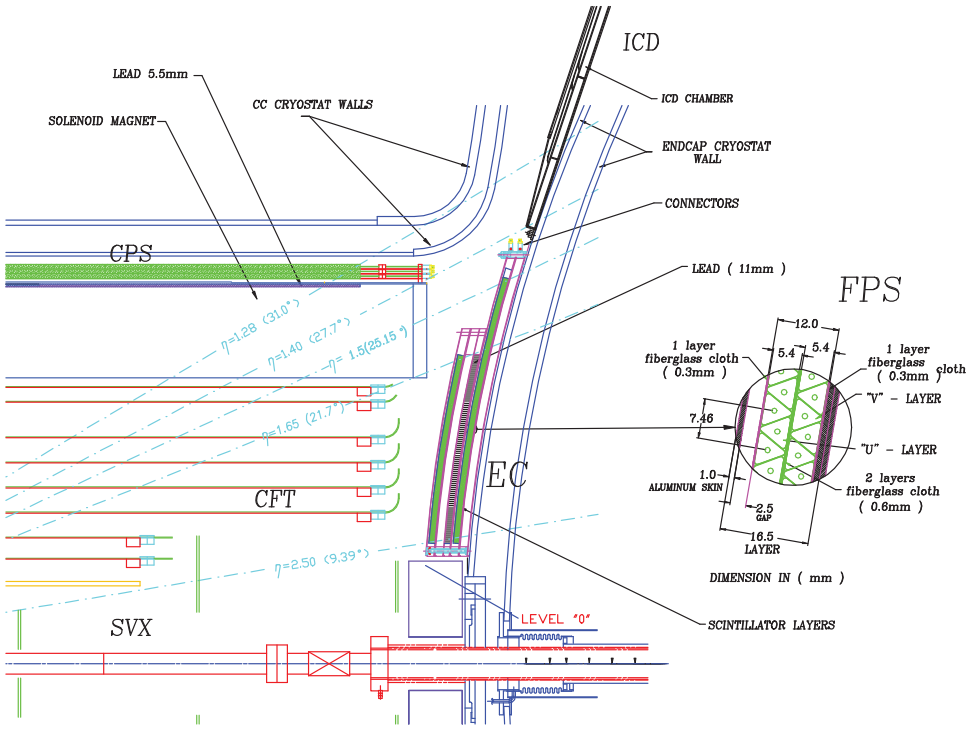

Figure 3.7: Cross-sectional view of the DØ tracking and preshower system.

backward regions. The barrel detectors measure primarily the r - ϕ coordinate and the disk detectors measure r - z as well as r - ϕ. Thus vertices for high η particles are reconstructed in three dimensions by the disks, and vertices of particles at small values of η are measured in the barrel.

An isometric view of the SMT is shown in Figure 3.8. The detector has six barrels in the central region. Each of these consists of four silicon readout layers, each layer having two staggered and overlapping sub-layers, see Figure 3.9. The outer barrels have single-sided and double-sided $2°$ stereo ladders. The four inner barrels have double-sided $90°$ stereo and double-sided $2°$ stereo ladders. Each barrel is capped at high $|z|$ with a so called "F-disk". In the far forward and backward regions, a unit consisting of three F-disks and two large-diameter "H-disks" provides tracking at $|\eta_{det}| < 3.0$. The F-disks are made of twelve wedges of double sided stereo detectors. The H-disks are made of 24 pairs of single-sided detectors glued back to back. Table 3.1 [66] summarises relevant SMT design parameters.

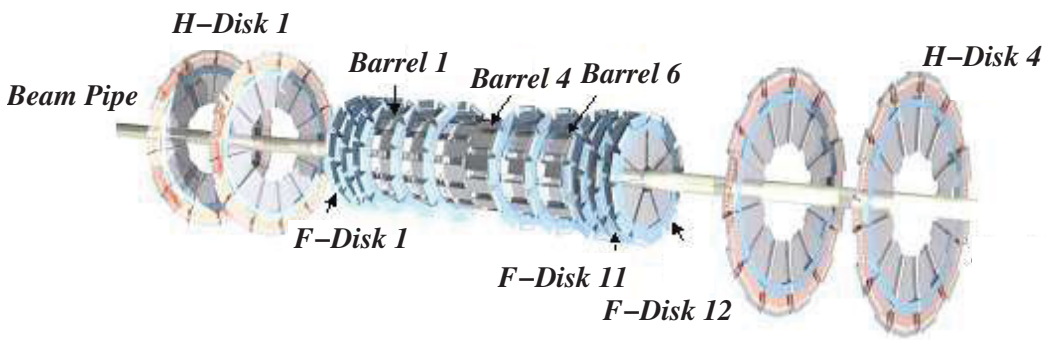

Figure 3.8: Isometric view of the DØ silicon vertex detector.

	Barrels	F-Disks	H-Disks
#Channels	387 072	258 048	147 456
Sensors	s/d sided	double sided	single sided
Stereo	$0°, 2°, 90°$	$\pm 15°$	$\pm 7.5°$
#Modules	432	144	96 pairs
Si area	1.3 m^2	0.4 m^2	1.3 m^2
Inner radius	2.7 cm	2.6 cm	9.5 cm
Outer radius	9.4 cm	10.5 cm	26 cm
Maximal $\|z\|$	38.4 cm	54.8 cm	120 cm

Table 3.1: Specifications of the silicon vertex detector, where “s,d” means single, double sided.

Central Fibre Tracker

The Central Fibre Tracker [61, 62] consists of 835 μm diameter scintillating fibres mounted on eight concentric support cylinders and occupies the radial space from 20 to 52 cm from the centre of the beam pipe. The two innermost cylinders are 1.66 m long, the outer six cylinders are 2.52 m long. Each cylinder supports one double layer of fibres oriented along the beam direction and a second double layer at a stereo angle of alternating $+3°$ and $-3°$. The two layers of fibres are offset by half a fibre width to ensure full acceptance coverage. The small fibre diameter gives the CFT a cluster resolution of about 100 μm per double layer.

Light production in the fibres is a multi-step process. When a charged particle traverses one of the fibres, the scintillator emits light at $\lambda = 340$ nm through a rapid fluorescence decay. A wave-shifting dye efficiently absorbs the light well at $\lambda = 340$ nm and emits at $\lambda = 530$ nm. The light is then transmitted by total internal reflexion to the

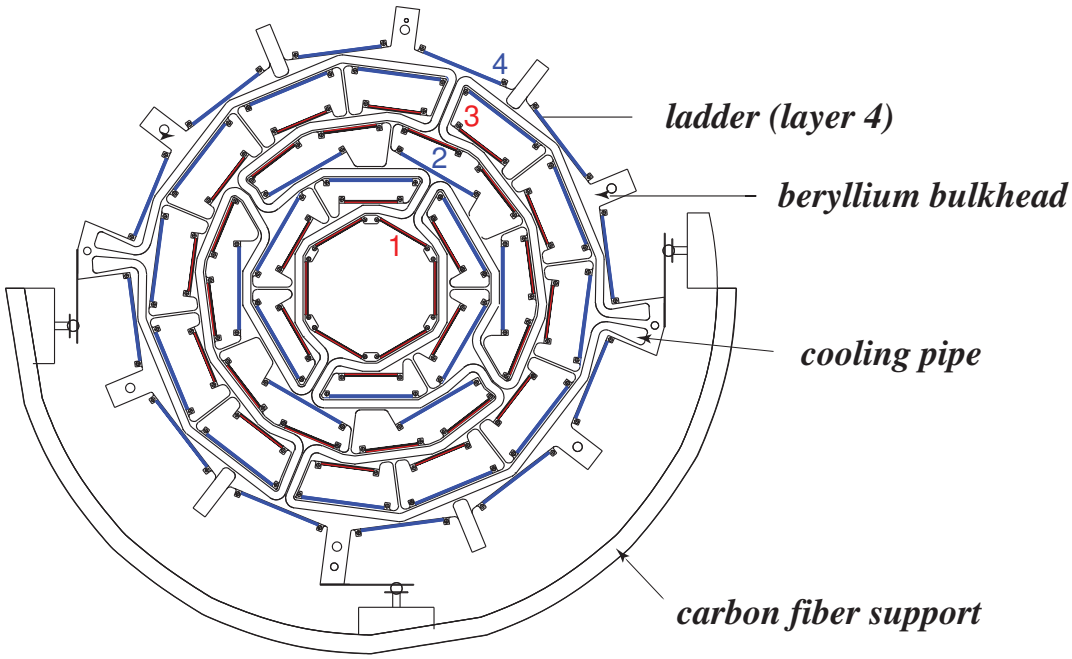

Figure 3.9: Cross section of the silicon vertex detector barrel structure.

end of the scintillating fibre, where it is transfered through an optical connection to clear fibre waveguides of identical diameter, which are 7.8 to 11.9 m long. The light is only observed from one end of each scintillating fibre. The opposite end of the scintillating fibres is sputtered with an aluminium coating that provides a reflectivity of 85 to 90 %. The clear fibre waveguides carry the scintillation light to visible light photon counters (VLPCs) which convert it into an electronic pulse.

The visible light photon counters are situated in a liquid Helium cryostat and operate at a temperature of 9 K. They detect photons with a quantum efficiency of 85 % and provide about 30 k- 60 k electrons per photon. A minimum ionising particle creates an average of eight photo-electrons per layer, depending on the angle between the scintillating fibre and the particle trajectory.

Solenoid Magnet

The superconducting solenoid magnet [67] is designed to optimise the tranverse momentum resolution, $\Delta p_T/p_T$, and track recognition. It is 2.73 m in length and 1.42 m in diameter, corresponding to maximum of 0.9 radiation lengths, X_0, and provides a 2 T magnetic field with a field homogeneity of 0.5 %. It operates at 10 K, the current is 4820 A and the stored energy is 5.6 MJ.

3.2.4 Calorimeter System

The task of the calorimeter system is to measure the energy of particles by inducing them to produce electromagnetic and hadronic showers. Passive layers of dense material in which the shower begins are followed by active layers, where the surviving fraction of the shower energy is sampled through ionisation.

Preshower Detectors

The preshower detectors (see Figure 3.7) are designed to improve electron identification and the energy resolution, which is degraded compared to Run I by the presence of the solenoid. They function as calorimeters as well as tracking detectors, enhancing the spatial matching between tracks and calorimeter showers [68].

Central Preshower detector
The Central Preshower detector (CPS) consists of three concentric cylindrical layers of triangular scintillator strips (axial and stereo $\pm 23°$). It is located in the 5 cm gap between the solenoid and the central calorimeter, and covers the of region $|\eta_{det}| < 1.3$. Installed between the solenoid and the CPS is a Pb radiator approximately one radiation length thick which serves to discriminate photons and electrons from pions by converting them into showers.

Forward Preshower detector
The Forward Preshower detectors (FPS) are each mounted on the spherical heads of the end calorimeter cryostats (see Sect. 3.2.4). They occupy the region between the luminosity counters (see Sect. 3.2.2) at the inner edge and the intercryostat detectors (see Sect. 3.2.4) at the outer edge, and cover the region of $1.5 < |\eta_{det}| < 2.5$. Each detector is made from two double layers of scintillator strips (stereo $\pm 22.5°$), separated by a 2 X_0-thick lead-stainless-steel absorber. All charged particles passing through the detector register a three-dimensional hit in the first, so-called MIP layer. Photons do not generally interact in the MIP layer. Electrons and photons shower in the absorber, leading to a cluster of energy in the second, so-called shower layer. These signatures can be used to distinguish muons, electrons and photons.

Scintillating light produced from the passage of charged particles is collected by wavelength-shifting fibres and transmitted through clear fibres to photon counters in the same way as explained for the central fiber detector in Section 3.2.3.

The preshower detectors are discussed in more detail in [69] and [70].

Liquid Argon Calorimeter

The Liquid Argon calorimeter [61, 62, 71] is designed to provide precise energy measurements, to assist in the identification of electrons, photons, taus and jets and to establish the transverse energy balance in an event. The device is also sensitive to minimum ionising particles and therefore can serve to identify muons. The calorimeter itself is unchanged from Run I and is shown in Figure 3.10. However, there is significantly more material in front of the calorimeter (2-4 X_0, depending on η) and faster readout electronics were build to cope with the reduced bunch crossing time from Run I to Run II.

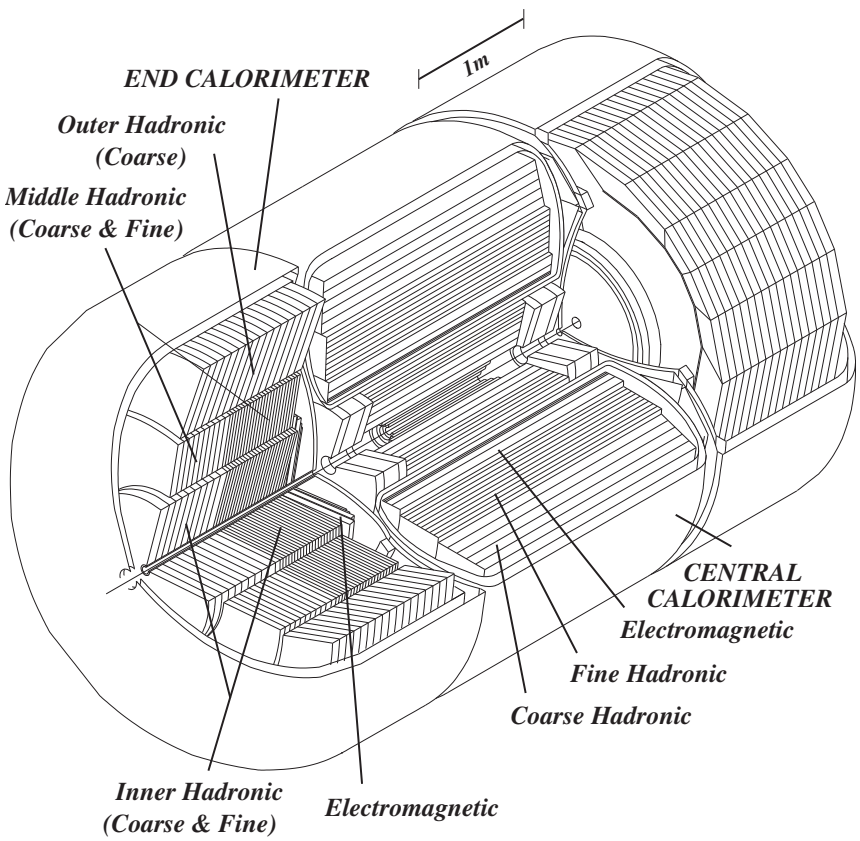

Figure 3.10: Isometric view of the central and two end calorimeters.

As shown in Figure 3.11, the Liquid Argon calorimeter is subdivided into the central calorimeter (CC) covering roughly $|\eta_{det}| < 1$ and two end calorimeters (EC) extending the coverage to $|\eta_{det}| \approx 4$. Each calorimeter contains an electromagnetic section closest to the interaction region followed by fine grained and coarse grained hadronic sections whose module size increases with the distance from the interaction region. The active medium

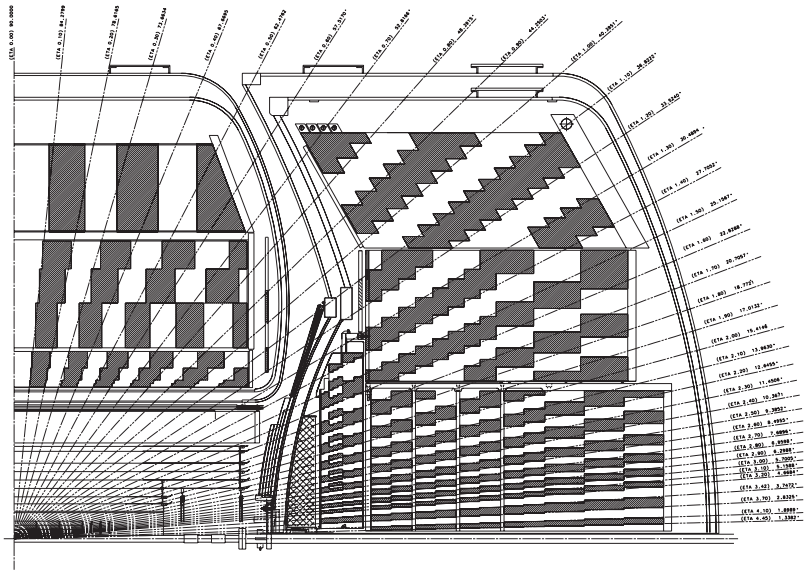

Figure 3.11: Schematic view of a quarter of the DØ calorimeter showing the transverse and longitudinal segmentation pattern. The shading pattern indicates cells for signal readout. The rays indicate the pseudo-rapidity intervals seen from the centre of the detector.

for all calorimeters is liquid Argon, and each of the three calorimeters is located within a cryostat that maintains the temperature at approximately 80 K. In order to achieve the same energy response for electromagnetic and hadronic particles, $e/h \approx 1$ (compensating calorimeter), different absorber plates are used at different locations. The e/π ratio was measured for charged pions in test beam data of Run I and the calorimeter was found to be nearly compensating, with $1 < e/\pi < 1.05$ for energies above 30 GeV [72]. The electromagnetic sections (EM) use thin (3 or 4 mm) plates made from nearly pure depleted Uranium. The fine hadronic sections are made from 6 mm thick Uranium-Niobium alloy. The coarse hadronic modules contain 46.5 mm thick plates of either Copper (in the CC) or stainless steel (in the EC).

The elements are combined in a basic unit, a calorimeter cell, depicted in Figure 3.12, which contains the absorber plates, the active medium and a Copper readout pad, laminated to G10 and covered in resistive epoxy, for collecting the charge generated by ionisation. An electric field is established by grounding the absorber plates and holding the resistive surface of the pad at typically 1.6 kV. The electron drift time across the 2.3 mm gap is approximately 450 ns. Several such pads are connected together in depth to form a readout cell (see Figure 3.11).

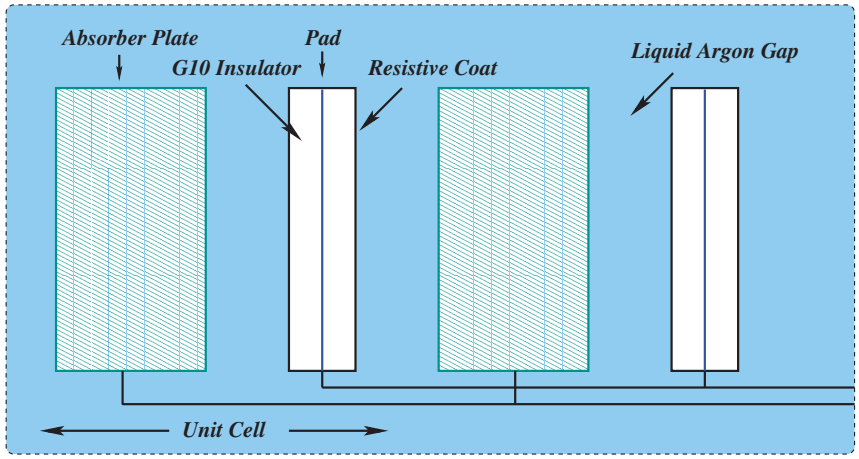

Figure 3.12: Schematic view of a calorimeter cell.

Inter-Cryostat Detectors

Since the calorimeter system is contained in three separate cryostats, it provides an incomplete coverage in the pseudorapidity region $0.8 < |\eta_{det}| < 1.4$, as can be seen in Figure 3.11. Additional sampling layers are added in the form of scintillating counters between the CC and EC cryostats (called Inter-Cryostat Detector or ICD [61, 62, 71]). The segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ matches exactly the Liquid Argon calorimeter geometry. In addition, separate single-cell structures, called *massless gaps* [61, 62, 71] are installed, both in the central calorimeter and in the end cap calorimeters.

3.2.5 Muon System

The outermost part of the DØ detector is the muon system (Figures 3.13 and 3.14). The detection of muons relies on their penetration power through material. Several meters of high-density material absorb almost all electrons and hadrons. Any charged particle that penetrates this material is defined to be a muon. The DØ muon detection system serves to identify and trigger on these muons and measure their momenta and charge.

The muon spectrometer consists of three subsystem: Proportional Drift Tubes (PDTs), Mini Drift Tubes (MDTs) and scintillation counters. The PDTs cover the region of $|\eta_{det}| < 1.0$ and the planes of MDTs extend the muon detection to $|\eta_{det}| = 2.0$. The scintillation counters are used for triggering and for the rejection of cosmic and beam related muons. Toroidal magnets and a special shielding complete the muon system. Each sub-system has three layers, called A, B, and C. The A layer is located between the calorimeter and the iron of the toroid magnet, the B and C layers are located outside of

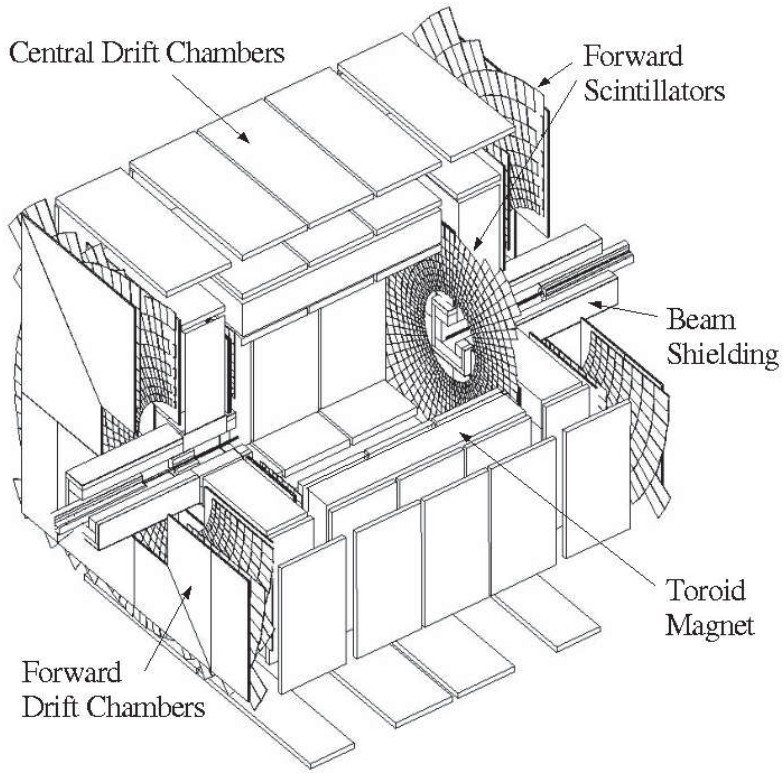

Figure 3.13: A cut-away view of the muon system.

the iron. In the region directly below the calorimeter, only partial coverage by muon detectors is possible since the support structure for the DØ detector and readout electronics is located in this region.

The thickness of the detector is in the range of 5-9 interaction lengths in the calorimeter, and in the range of 7-9 interaction lengths in the iron. This results in a most probable value for the energy loss of a muon of 1.6 GeVin the calorimeter, and about 1.7 GeV in the iron. The momentum measurement is corrected for this energy loss.

In the following, the subsystems of the muon spectrometer are discussed.

Toroid Magnet

The 1973 ton toroid magnet [67], located between layers A and B, allows a measurement of the muon momentum. The central toroid is centred at a distance of 317.5 cm from the Tevatron beam line. In order to permit access to the inner portions of the DØ detector, it

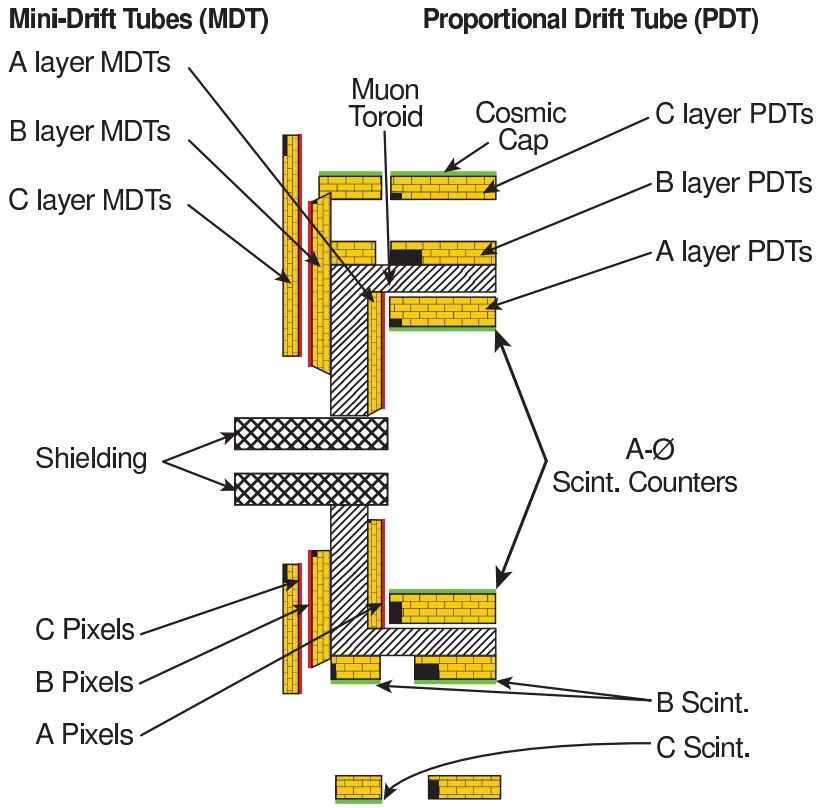

Figure 3.14: Side view of the DØ muon system with Proportional Drift Tubes (PDTs), Mini Drift Tubes (MDTs), scintillator counters, toroid magnet and shielding.

is split into three parts. A bottom section gives support for the calorimeter and tracking detectors. The central toroid is completed by two movable c-shaped shells. The two forward toroid magnets are located at $447 \leq |z| \leq 600$ cm. The coils carry currents of 1500 A and produce an internal field of 1.8 T.

Proportional Drift Tubes

The central muon system is assembled from drift chambers that consist of a row of 24 cells, each with a sense wire running parallel to the magnetic field lines in the toroid magnet. The cells have a cross section of 5.5 cm x 10 cm and are 240 cm long. The inner layer (A-layer) is four chambers deep and the outer layers (B and C-layer) are three chambers deep. The drift direction is parallel to the z-coordinate, with an expected resolution of around 0.5 mm. The PDTs are discussed in more detail in [73].

Mini Drift Tubes

The forward muon system is made from individual tubes with 8 cells of 1 cm x 1 cm cross section, each holding one wire. As in the central system, the A-layer is four layers deep. It is located at $|z| = 447$ cm. The B and C-layer are each three layers deep, and are located at $|z| = 630$ cm and $|z| = 827$ cm. Shielding around the beam-pipe has been implemented to reduce the radiation dose from beam remnants. The MDTs are discussed in more detail in [74].

Scintillation Counters

Scintillator counters are installed both in the central and forward regions: two or three layers of scintillator pixels in the forward, A-layer counters and BC counters in the central region. Scintillators are read out with both a 15-20 ns "trigger" gate, and a 80-100 ns "read out" gate. The counters have a ϕ segmentation of 4.5 degrees and are used for triggering, cosmic ray veto, beam related muon rejection and track reconstruction. The scintillator counters are discussed in more details in [75].

Shielding

Three sources contribute to backgrounds in the central and forward muon system:

- Scattered proton and antiproton fragments that interact with the end of the calorimeter or with the beam pipe produce background in the central and forward A layer.
- Proton and antiproton fragments, mostly muons from the decays of pions created by proton and antiproton interactions upstream of the detector, interacting with the low beta quadrupole magnets produce hits in the B and C layers of the forward system.
- Beam halo interactions affect both the central and the forward muon system.

The shielding consists of layers of iron, polyethylene and lead in a steel structure, surrounding the beam pipe and the low beta quadrupole magnets. Iron is used as a hadronic and electromagnetic absorber, polyethylene is a good absorber for neutrons due to its high hydrogen content, and lead is used to absorb gamma rays.

3.2.6 Trigger System

The overwhelming majority of proton antiproton collisions results in collisions of little interest. Collisions which produce massive particles such as W, Z bosons and top quarks, or those which might provide evidence for new physics occur extremely rarely. To accumulate a large sample of events of interest without having to store and reconstruct a staggering number of uninteresting collisions, DØ employs an event trigger which decides whether to store an event or to disregard it. At the Tevatron, an input bunch crossing rate of 2.5 MHz must be reduced to a final data taking rate of 50 Hz, a limit given by the offline reconstruction capabilities. The trigger system is a three-tiered pipelined system; each tier examines the event in more detail than lower tiers and restricts the input rate of events to higher tiers.

Level 1 Trigger

Collisions occur at a rate of 1.7 MHz. The first trigger stage (Level 1 or L1) has a pipeline, which allows to take a decision within 4.2 μs, and has a trigger accept rate of about 2 kHz. The trigger algorithm is carried out in a framework of field programmable gate arrays (FPGAs), which receive as input simple objects created in the luminosity monitor, the calorimeter and the muon system.

Level 2 Trigger

In the second stage (Level 2 or L2), hardware engines associated with specific subdetectors process information which is then used in a global processor to determine correlations between different detectors, e.g. to match tracks and leptons. Level 2 has an accept rate of 1 kHz with a maximum dead-time of 5%, and a maximal latency of 100 μs.

Figure 3.15 illustrates the design of the DØ Level 1 and Level 2 trigger system. The components currently being commissioned are represented as hatched boxes.

Level 3 Trigger

The third stage (Level 3 or L3) uses a collection of approximately 100 PC farm nodes to perform a partial reconstruction of the event and take a trigger decision using the full event information (complete physics objects as well as their correlations). The nominal 1 kHz input rate is reduced to a rate of 50 Hz at which data are recorded for offline analyses. The online version of the event reconstruction is optimised to fit into the timing requirements ($\sim$50 ms) at Level 3.

When an event passes the Level 2 decision, the data from the roughly 80 readout crates are sent to a single farm node, where the event reconstruction takes place. Filters

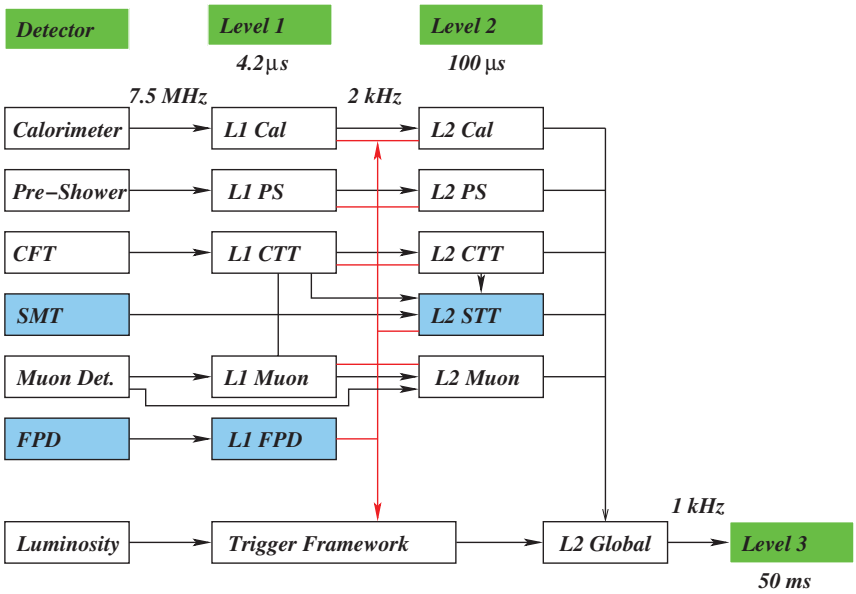

Figure 3.15: The design of the DØ Level 1 and Level 2 trigger system.

can select events based on physics quantities. A single farm node handles events at a rate of 10-20 Hz. Events that pass a Level 3 filter are sent from the farm node to a collector, which in turn sends the events over Gbit/s optical Ethernet to FERMILAB's Feynman Computing Center, where the events are written to tape for offline analysis.

Chapter 4

Event Reconstruction and Object Identification

The raw event data collected from the complete detector consist of hit information from nearly a million detector channels. This information has to be processed in order to extract the information required to reconstruct the physics event.

Various algorithms are used to process the huge amount of raw detector information and to define basic physics objects and their properties, representing the particles generated in the collision. The detector design allows to distinguish and to define the following fundamental objects: tracks, primary vertex, electrons, photons, muons, jets and their flavour, and missing transverse energy $\not\!\!E_T$.

Generally, the object reconstruction and identification is optimised for efficiency and purity, and to provide the best possible measurement of the magnitude and direction of the object's momentum. Other objects in the event, such as particles that do not originate from the primary $p\overline{p}$ collision or noise generated in the detector components or the readout electronics can mimic the object signature, thereby degrading the purity.

In this chapter, the objects used in this analysis, namely tracks of charged particles, primary vertices and muons are described.

4.1 Tracks of Charged Particles

Charged particles, bent in the magnetic field, generate hits in the central tracking system from which tracks are reconstructed. The reconstruction is abstracted in two parts. The first part is hit clustering, which groups individual channels that are likely to have been generated by the passage of an individual particle through the detector. If particles are passing through the silicon vertex detector, they will deposit charge in a number of strips.

A particle traversing the central fibre detector will cause a number of fibres to scintillate. In both cases, a cluster is defined as a group of adjacent strips or fibres above a noise threshold. The second part of the reconstruction is the track finding, which is again subdivided into two algorithms: pattern recognition and track fitting. The pattern recognition creates sets of clusters which lie along physical paths. The track fitting uses sophisticated algorithms (Kalman fitter) to fit a candidate charged particle track to a physical path, using a χ^2 test. For the track fitting the algorithm described in [76] has been used.

A good calibration for the tracking algorithm is the reconstruction of known meson and baryon resonances. As examples, the process $K_s^0 \to \pi^+\pi^-$ is measured with a mass resolution of $\sigma = 7.3$ MeV and $\Lambda^0 \to p^+\pi^-$ with $\sigma = 2.6$ MeV (Figure 4.1).

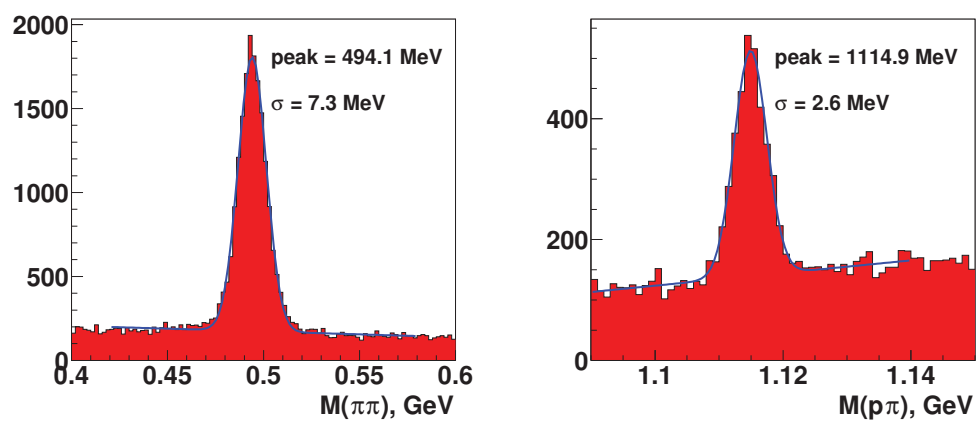

Figure 4.1: Invariant mass distributions for the processes $K_s^0 \to \pi^+\pi^-$ (left) and $\Lambda^0 \to p^+\pi^-$ (right). Also shown is a Gaussian distribution plus a linear fit to the signal and background.

4.2 Primary Vertex

A precise determination of the primary vertex position is crucial for all B physics analyses. The long lifetime of the B meson distinguishes B events from background events. A method that follows the procedure outlined in [76] has been used to determine the primary vertex.

A primary vertex is determined for each event by minimising a χ^2 function that depends on all tracks in the event and a term that represents the beam spot constraint. The beam spot is stable during the periods of time during which the proton and the antiproton beams are kept colliding continously, and can be used as a constraint for the primary

vertex fit. Therefore, the nominal beam spot is defined as the average beam position run-by-run, where a run typically lasts several hours. The initial primary vertex candidate and its χ^2 are obtained using all tracks in the event. Next, each track is removed temporarily from the vertex fit and the χ^2 is recalculated. If the χ^2 decreases by nine or more, this track is discarded from the primary vertex fit. This procedure is repeated until no more tracks are discarded. The achieved primary vertex resolution by using this procedure is 18 μm in the x and y directions.

4.3 Secondary Vertex

The secondary vertex algorithm begins by creating a list of all tracks that are not included in the primary vertex. Then the same procedure as for finding the primary vertex is applied. The minimum requirement to define a secondary vertex is that it contains two tracks of oppositely charged particles. If such a vertex is found, it is tested on the so-called V^0 hypothesis. The compatibility of the two tracks with the decays $K_s^0 \to \pi^+\pi^-$, $\Lambda^0 \to p^+\pi^-$ (where the higher p_T track is assumed to be the proton, according to observations from Monte Carlo studies) and with photon conversion ($\gamma \to e^+e^-$) is tested. Secondary vertices are rejected if the invariant di-track mass for a given V^0 hypothesis is consistent with the corresponding V^0 mass in a mass window defined by $\pm 3\sigma$ of the measured V^0 mass resolution.

The impact parameter resolution for single muon events is shown as a function of transverse momentum in Figure 4.2. The impact parameter denotes the minimal distance of the track to the primary vertex of the event.

As examples for the reconstruction of secondary vertices, the reconstruction of the processes $\Xi^\pm \to \Lambda^0\pi^\pm$ and $\Omega^\pm \to \Lambda^0 K^\pm$ is shown in Figure 4.3, where multiple tracks with impact parameters of the order of centimetres are reconstructed.

4.4 Muons

Muons are reconstructed using information from two detector systems: the muon detector and the central tracker. A 'local' track in the muon system forms the basis for muon identification. Muons are required to have hits in all layers of the muon system (both inside and outside of the toroid steel). The exact muon identification criteria are presented in Sect. 4.4.1. The superior track resolution of the central tracker is used to improve the knowledge of the kinematic properties of the muon.

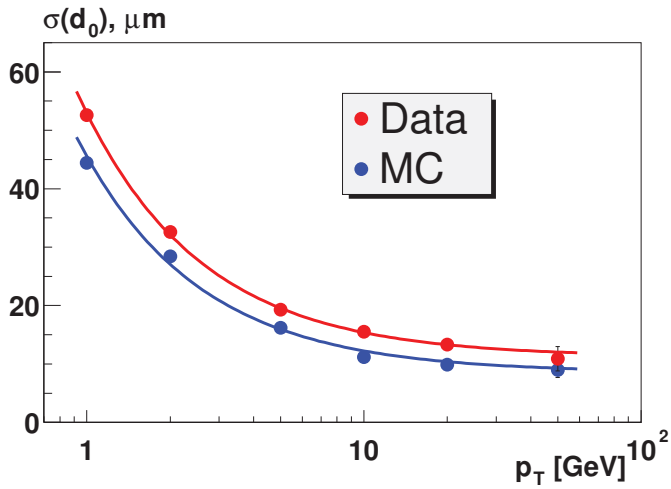

Figure 4.2: Impact parameter resolution measured in data and in simulated single muon events. The lines show fits to data and Monte Carlo events.

4.4.1 Muon Identification Criteria

The following standard Muon ID requirements are used in the analyses presented in this thesis:

- Muon candidates are required to be of '$|nseg| = 3$ medium' quality, according to the certified Muon ID criteria [77]:
  - at least two wire hits in the A segment,
  - at least one scintillator hit in the A segment,
  - at least two wire hits in the BC segment,
  - at least one scintillator hit in the BC segment, (except for central muons with less than four wire hits in the BC segment).
- The standard 'loose' cut against cosmics is applied, based on timing information from scintillator hits associated with the muon. The default timing cuts from the certified Muon ID [77] are used, requiring A-layer and BC-layer scintillator times of $|t_{\mathrm{A}}| < 10$ ns and $|t_{\mathrm{BC}}| < 10$ ns.

In addition a 'central track match' is required, which denotes that the muon tracks are extended to the point of closest approach (PCA) to the beam and their parameters are compared with those of central tracks at the PCA. For each muon track a global track fit is

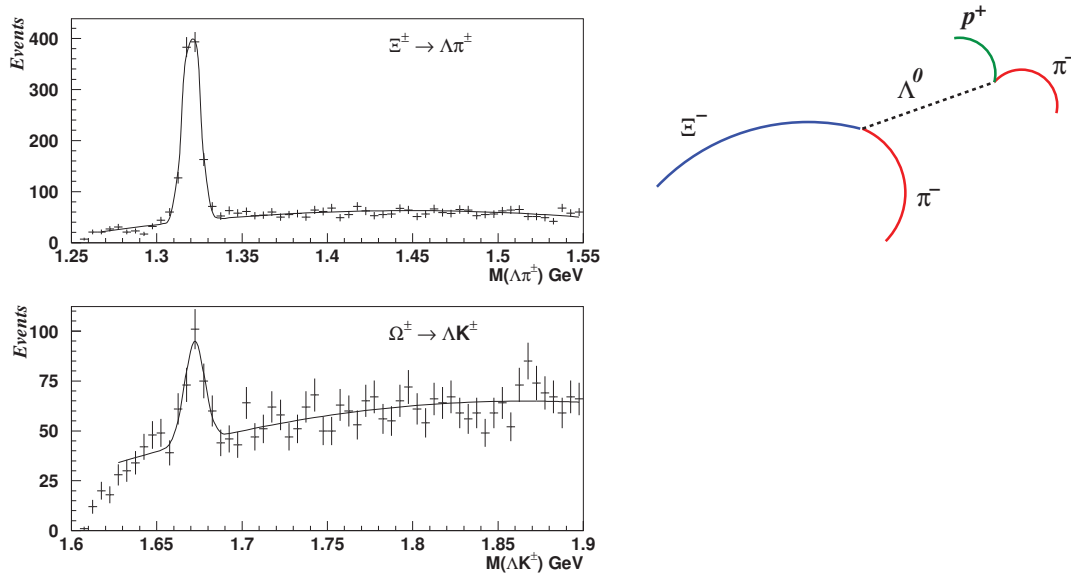

Figure 4.3: Invariant mass distributions for the processes $\Xi^{\pm} \rightarrow \Lambda^0\pi^{\pm}$ (top left) and $\Omega^{\pm} \rightarrow \Lambda^0 K^{\pm}$ (bottom left). The reconstruction of the processes are examples of events with multiple tracks with impact parameter in the order of centimetres. A sketch of the decay $\Xi^- \rightarrow \Lambda^0\pi^- \rightarrow p^+\pi^-\pi^-$ is also shown (right).

performed with all central tracks that fall within one radian in azimuthal and polar angle around the muon track at PCA. The central track that gives the highest χ^2-probability is then considered as the muon candidate. The muon track parameters are then taken from the central tracking system, since it gives a more precise measurement of these parameters than the muon system.

Chapter 5

Data and Event Selection

5.1 Data Sample

The data samples used in this thesis were collected with the DØ detector in $p\overline{p}$ collisions at a center-of-mass energy of 1.96 TeV. The data set corresponds to an integrated luminosity of 300 pb^{-1} and was collected from August 2002 to June 2004. The data sample was collected using five different trigger list versions: v8 to v12 (a trigger list version is defined for a certain range of runs and consists of a set of well defined triggers). After the 28th of June 2004 the trigger list was changed significantly in order to cope with the higher initial luminosities delivered by the Tevatron. Since no reliable simulation was available to estimate the efficiencies of the new triggers it was decided not to use data recorded with the new version of the trigger list.

5.2 Data Quality Selection

Only a fraction of the collisions delivered by the Tevatron is actually recorded by the DØ detector. Out of that data only the fraction for which all detector systems were functioning well is used for the data analyses.

Data quality monitoring is performed online and offline. It is crucial for a high data taking efficiency to catch any malfunctioning of detector components, of the readout or of the triggering as early as possible. Online data quality monitoring makes it possible to react fast and thus maximise the amount of heigh quality data taken.

However, there are data quality issues which cannot be recognised online. The remaining deficient data are eliminated by offline data quality monitoring. It is based on the comparison of distributions of basic physics objects or other variables that describe the detector performance with nominal distributions of the same quantities. In a second

step, a list of possible data corruption methods is identified, quality measures are defined with a maximum discrimination power between usable and unusable data, and criteria are defined to classify the data quality.

Since the corruption of data can occur on time scales much smaller than the length of a run, the classification is preferably done on the smallest possible units of data. The data quality selection is therefore broken down into a run based, a luminosity block number (LBN) based, corresponding to approximately one minute of data taking and several thousand events recorded, and an event-by-event based selection. The analyses presented here, however, did not make use of the luminosity block quality criteria.

Run quality selection:
The selection of good runs is based on information stored in the DØ Run Quality Database [78]. The requirements are summarised in Table 5.1. It is guaranteed that no hardware failures occurred and that all readout crates are included in the readout for the muon system [79], the SMT and the CFT. The tracking quality is particularly important for the identification of muons originating from the secondary vertex and for building the event variables.

Detector	Quality
MUON	Not Bad
SMT	Not Bad
CFT	Not Bad

Table 5.1: Run quality requirements.

5.3 Trigger Selection

All decay channels studied in this thesis contain two muons in the final state. Therefore, only events passing the following di-muon triggers mu2ptxatxx_fz, MU2_A_L2M0, MU2_A_L3L0 and 2MU_A_L2ETAPHI were considered. These triggers were found to be sufficiently robust and stable over the course of the runs and have high efficiencies. At Level 1, the trigger mu2ptxatxx_fz requires a tight muon scintillator coincidence in two layers of the DØ muon system (c.f. section 3.2.5). The trigger 2MU_A_L2M0 requires in addition at least one medium quality muon in the Level 2 trigger, consisting of a reconstructed track based on hits in the proportional wire counters inside and outside the magnetic toroid, as described in section 3.2.5. At least one additional Level 2 muon of loose quality, separated from the first Level 2 muon by at minimum of 13.5° in ϕ and 0.15 in η, is required by the 2MU_A_L2ETAPHI trigger.

The Level 1 and Level 2 trigger conditions and their implementation have not changed over the period of data taking using the trigger list v8.x to v12.x. The same trigger efficiency estimates extracted from a trigger simulation [80] applied to signal Monte Carlo and normalisation Monte Carlo samples were therefore used for all trigger list versions. The invariant mass spectrum for the full data set is shown in Figure 5.1. All expected resonances in the low invariant mass region are clearly visible.

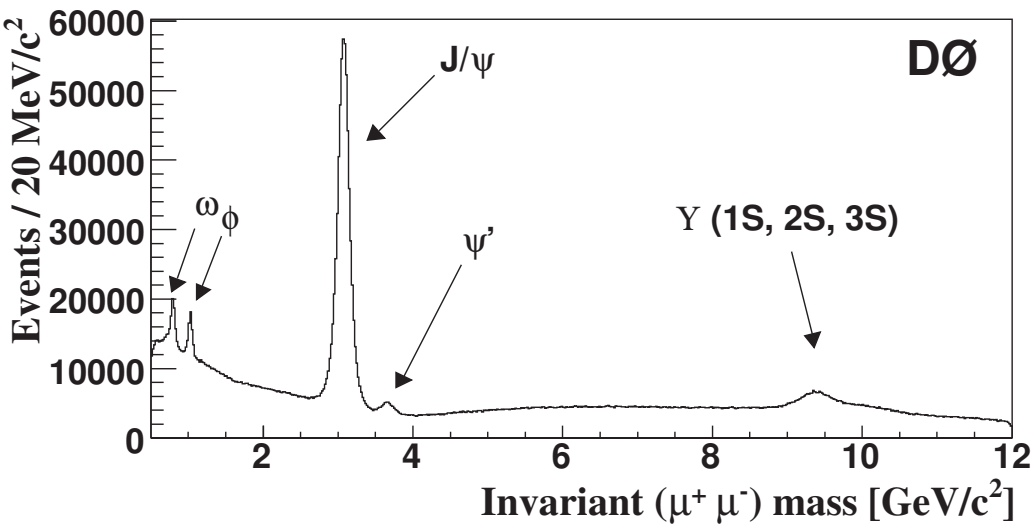

Figure 5.1: Invariant di-muon mass spectrum of various meson resonances: ω, ϕ, J/ψ, ψ' and Υ.

5.4 Monte Carlo Simulation

In general, Monte Carlo simulation methods are applied to calculate approximate solutions to a variety of mathematical problems by performing statistical sampling experiments. In particle physics, the term Monte Carlo simulation refers to the random generation of physical final states from the collision of two (or more) initial-state particles. The generated particles in that final state are propagated through the detector based on their initial momenta, and their interactions in the active and passive material of the detector are simulated.

In order to compare the recorded data to simulated events, Monte Carlo samples for signal and normalisation channels have been produced. This simulation allows to predict the ideal behaviour of the detector, providing a model to which measured data can be compared to assess detector and reconstruction performance. It also provides a prediction

of the expected detector response to specific physical processes of interest, allowing the detailed determination of physical distributions and selection efficiencies which enter in the calculation of branching fractions.

All Monte Carlo samples were produced using Pythia 6.1 [81] and EvtGen [82] to generate $b\bar{b}$ events with a $p_T > 0.1$ GeV/c and $|\eta| < 4.2$ on the parton level. CTEQ5L [83] was used as the leading-order parton density function (PDF) for the leading-order matrix element calculation in the Monte Carlo samples. On average, 0.8 (Poisson distributed) minimum-bias event have been added to each simulated event.

All simulated events are passed through a full Geant [84] simulation of the DØ detector and are then reconstructed in the same way as the data.

The most time consuming step in the simulation chain is the simulation of the interaction of all generated particles in the material of the detector. This step involves the determination of the trajectories of the particles as they enter into the detector, the simulation of decays of particles with finite lifetimes, the simulation of multiple scattering and energy loss as the particles interact with material in the detector and a realistic description of the response of the detector to the energy depositions made by the particles. In order to increase the Monte Carlo simulation efficiency cuts were applied already at the generator level to reject events, which would not have been reconstructed due to analysis cuts. This step has to be checked to ensure that it does not introduce a bias or treats signal and normalisation channels in a different way.

Monte Carlo samples of the following decays that are under study in this thesis were generated:

- $B_s^0 \to \mu^+\mu^-$
- $B_d^0 \to \mu^+\mu^-$
- $B^\pm \to J/\psi\, K^\pm$
- $B_s^0 \to J/\psi\, \phi$
- $B_s^0 \to \psi(2S)\, \phi$
- $B^\pm \to \psi(2S)\, K^\pm$

On all Monte Carlo samples, the following requirements were applied at generator level:

- At least two opposite signed muons with
  - muon $p_T > 1.5$ GeV/c
  - $|\eta| < 2.5$ for each muon

- p_T of the B_s^0 candidate > 4.0 GeV/c

where the two muons where required to come from the B_s^0 meson in the decay $B_s^0 \to \mu^+\mu^-$ and from the J/ψ or $\psi(2S)$ in the other cases.

For the decays with a Kaon in the final state ($B^\pm \to J/\psi\, K^\pm$ and $B^\pm \to \psi(2S)\, K^\pm$) this Kaon coming from the $B^\pm$ meson was required to have a minimum p_T of 0.5 GeV/c.

For the $B_s^0 \to J/\psi\, \phi$ and $B_s^0 \to \psi(2S)\, \phi$ mode, in addition one ϕ meson coming from the B_s^0 meson and decaying into two oppositely charged Kaons with a minimal Kaon p_T of 0.5 GeV/c and a pseudorapidity $|\eta| < 2.5$ for each Kaon was required.

Chapter 6

Optimisation

For the analyses presented in this thesis several different B meson decay channels are selected. The similarity among the channels allow the reconstruction of the B meson candidate in an almost identical way by imposing similar selection criteria. The pre-selection criteria will be discussed separately for each analysis. In addition to basic requirements, three different variables are being used to further discriminate against background for all analyses. These variables exploit best the distinctive properties of the B meson and are discussed in this Chapter. Also the strategy to optimise the cut values for the discriminating variables is discussed.

6.1 Discriminating Variables

The chosen three discriminating variables are the pointing angle, the isolation and the decay length significance. Figure 6.1 shows a sketch of these three variables for the decay $B_s^0 \to \mu^+\mu^-$

6.1.1 Pointing Angle

The pointing angle α is defined as the angle between the momentum vector of the B candidate and the decay length vector $\vec{l}_{Vtx}$ pointing from the primary vertex to the secondary vertex where all vectors are determined in three dimensions. It is used to check the consistency of the direction of the decay vertex with the flight direction of the B candidate. If the tracks used to build the B candidate originate from the decay of a parent particle B_s^0, the vector $\vec{l}_{Vtx}$ should point in the same direction as the momentum vector of the B candidate $\vec{p}(B)$.

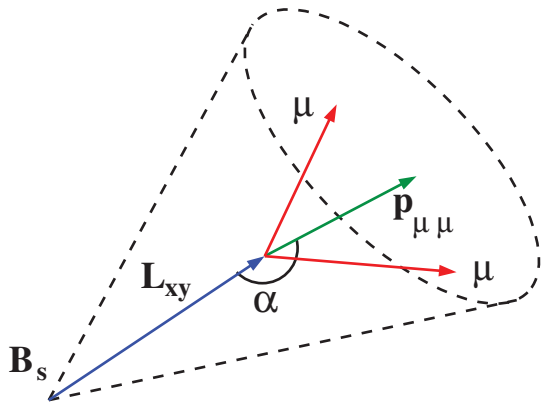

Figure 6.1: Pictogram for visualisation of the discriminating variables for the decay $B_s^0 \to \mu^+\mu^-$.

6.1.2 Isolation

The fragmentation characteristics of the b-quark are such that most of its momentum is carried by the B hadron. A large fraction of the momentum of the tracks observed close to the B meson is expected to be carried by the daughter tracks of the B, thus the number of extra tracks close to the B candidate tends to be small. The second cut parameter was therefore an isolation variable, $\mathcal{I}$, defined as:

$$\mathcal{I} = \frac{|\vec{p}(B)|}{|\vec{p}(B)| + \sum\limits_{\text{track } i \neq B} p_i(\Delta\mathcal{R} < 1)}. \tag{6.1}$$

Here, $\sum\limits_{\text{track } i \neq B} p_i$ is the scalar sum of the momenta of all tracks excluding the tracks used to build the B candidate that fall within a cone of $\Delta\mathcal{R} = 1$ around the momentum vector $\vec{p}(B)$ of the B candidate, where $\Delta\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. The z-coordinate of these tracks must be within 5 cm of the B candidate vertex to exclude tracks from other $p\overline{p}$ collisions that can occur during the same bunch crossing.

6.1.3 Transverse Decay Length Significance

The long lifetime of B mesons allows the use of the decay length as a strong rejection criteria against mostly random combinatoric background, e.g. two fake muons that happen to form a good vertex. The proper decay length $c\tau$ is defined as

$$c\tau = L_{xy} \cdot \frac{m_B}{p_T^B} \tag{6.2}$$

where the transverse decay length L_{xy} is the projection of the two dimensional decay length vector on the transverse momentum

$$L_{xy} = \frac{\vec{l}_{xy} \cdot \vec{p}_T^B}{p_T^B}. \tag{6.3}$$

As a third cut parameter, the transverse decay length significance $L_{xy}/\delta L_{xy}$ was used, since it gave a better discriminating power than the transverse decay length itself (where large values may be due to large errors).

6.2 Random Grid Search

The idea behind Random Grid Searches (RGS) [85] is to search systematically over a randomly selected grid of cut values for the discriminating variables, in order to aid in the optimisation of the signal and background classification. The density of grid points was randomly distributed following the distribution of the discriminating variables as generated in the signal Monte Carlo. A search over such a random grid is more efficient and consumes less computing time than a search over a regularly spaced grid (Figure 6.2).

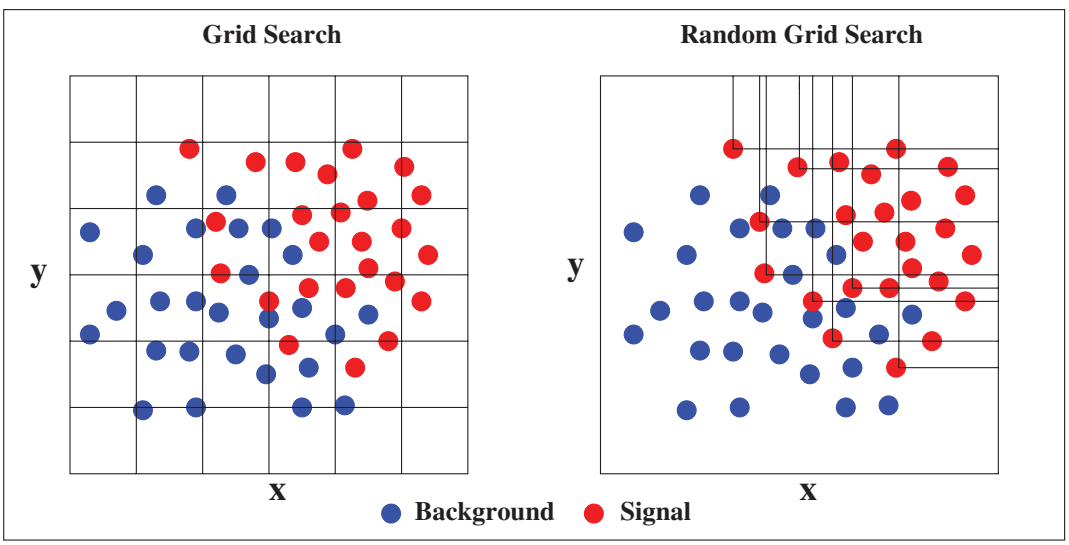

Figure 6.2: Visualisation of the advantage of a random grid (right side), over a regularly spaced grid (left side) search for two variables x and y.

For each set of cut values the Monte Carlo signal efficiency and the linear interpolation of the background from the sidebands into the signal region are calculated.

An optimisation criterion proposed by G. Punzi [86] has been used in order to find the optimal set of cuts. It consists in maximising the ratio P defined as:

$$P = \frac{\epsilon}{\frac{a}{2} + \sqrt{N_{Back}}} \tag{6.4}$$

Here, ϵ is the reconstruction efficiency of the signal Monte Carlo relative to a pre-selected sample and N_{Back} is the expected number of background events interpolated from the sidebands. The constant a is the number of standard deviations corresponding to the confidence level at which the signal hypothesis is tested[1] The value of a should be defined before the statistical test and has been set to 2, corresponding to about 95% CL.

The expression P has been used as an optimisation criterion in the random grid search, since it does not push the experimental efficiency down to very small values and does not prefer extremely low expected signal numbers over tiny background. Moreover, it can be maximised without knowing the cross section of the searched signal. The proposal by Punzi allows an unbiased optimisation that is free of priors or of limit setting algorithms.

6.3 Probability Density Ratio

For the search for the rare decay $B_s^0 \to \mu^+\mu^-$, another multi-variate approach – a so-called Likelihood Ratio – has been used. This method relies on the optimisation of the probability density ratio $\rho_s(\mathbf{x})/(\rho_s(\mathbf{x})+\rho_b(\mathbf{x}))$, based on the discriminant $\mathbf{x} = (x_1, x_2, x_3)$ with the variables x_i. The probability densities $\rho_{s,b}(\mathbf{x})$ are obtained by sampling over the available phase space for the signal Monte Carlo events in case of the signal density ρ_s and the data in case of the background density ρ_b. The isolation discriminant is largely uncorrelated with the remaining variables and can be factored out. Therefore, 1-dim probability densities $\rho_{s,b}(iso)$ for the isolation cut and 2-dim probability densities $\rho_{s,b}(\alpha, dl)$ for the discriminant pointing angle α and decay length (dl) have been determined. The overall density ratio was defined as

$$R = \frac{\rho_s(iso) \cdot \rho_s(\alpha, dl)}{\rho_s(iso) \cdot \rho_s(\alpha, dl) + \rho_b(iso) \cdot \rho_b(\alpha, dl)} \tag{6.5}$$

[1]Note, that an approximated formula was used for the optimisation criterion which is taken from reference [86]. This simplified expression should be adequate for most searches and assumes that the significance level, i.e., the probability to reject the zero-signal hypothesis H_0, when it is indeed true and the confidence level of not missing a discovery when the signal hypothesis H_1 is true, are set to the same value a.

6.4 A Multivariate Linear Classifier

In principle, any mathematical function of the cut parameters can be used as a discriminating function to separate different classes of events. In the case of Linear Discriminant Analysis, a linear function of the form

$$y_i = b_0 + b_1 \cdot x_1^i + b_2 \cdot x_2^i + \ldots + b_m \cdot x_m^i$$

is used where b_0 is a constant, b_1 through b_m are regression coefficients and the x_i are the independent variables. For two groups of data (signal and background) the parameters b_i have to be determined in such a way that the discrimination between the groups is best. The interpretation of the results of a two-group problem is that those variables with the largest (standardised) regression coefficients are the ones that contribute most to the prediction of group membership. If only two variables are used, the separator between the

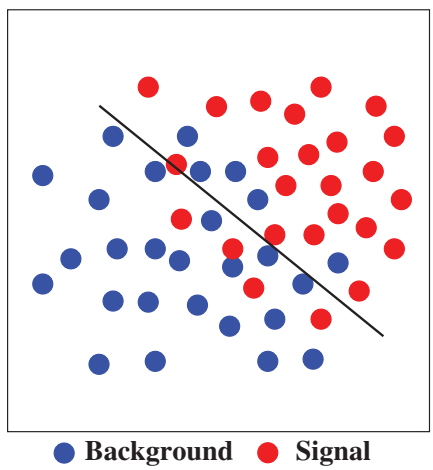

Figure 6.3: Visualisation of the the idea of separating two groups with a linear discriminant and indicating with a line the best separation of the two groups.

groups will become a line as indicated in Figure 6.3. If three independent variables are used, the separator is a plane and if more than three variables are present, the separator becomes a hyper-plane. The optimal separation of two groups is achieved, by maximising the distance to the separator using the so-called Mahalanobis metric or quadratic distance. It measures the separation of two groups of objects with means $\bar{x}^1 = (x_1^1, x_2^1, \ldots, x_m^1)^T$ and $\bar{x}^2 = (x_1^2, x_2^2, \ldots, x_m^2)^T$. The sample covariance matrix of the two groups is C and the Mahalanobis distance is then given by:

$$d^2_{\bar{x}^1,\bar{x}^2} = (\bar{x}^1 - \bar{x}^2)^T \; C \; (\bar{x}^2 - \bar{x}^1).$$

The logistic regression analysis chosen here is very similar to the standard linear discriminant method described above, but it does not rely too much on the data being normally distributed. As multivariate input, the three variables isolation, decay length significance and pointing angle as introduced above have been used.

The goal was to evaluate the background rejection performance applied to the observation of the $B_s^0 \to \psi(2S)\,\phi$ decay using a multivariate method, which is expected to give a better efficiency than the RGS using the Punzi criterion.

Chapter 7

Calculating Limits

7.1 Probabilities

The most common definition of a probability is the Frequentist approach. If an identical experiment is performed N times, and a certain outcome A occurs M times then the probability $p(A)$ for the event A to occur tends to the ratio M/N as $N \rightarrow \infty$. The set of all N cases is called a collective or ensemble and can be obtained by repeating one experiment N times or by measuring N experiments simultaneously. This is a very useful definition but also has some problems. One problem is the repeatability of an experiment. How can it be ensured that the conditions between the experiments have not changed and that the initial conditions are exactly the same? The next problem is that the limit $N \rightarrow \infty$ in a strict mathematical interpretation does not exist. How can one deduce from N experiments the outcome of the $N + 1$st experiment? Also, when does it converge to the limit? Nobody can perform an infinite number of experiments.

Another definition is that of a subjective probability, also known as the Bayesian definition of probability. This definition is based on the plausibility or credibility of the observed occurrence. It allows to test the probability of a hypothesis, which is not defined in the Frequentist approach. In the Bayesian definition, the actual knowledge about the observed event is taken into account, which makes this definition of probability subjective. The probability $p(A)$ depends on the information the observer has and transforms it into a degree of belief. If subsequent experiments are performed, the initial degree of belief can be modified, depending on the additional information obtained.

7.2 Confidence Intervals

A confidence interval [a_{min},a_{max}] is a random interval which contains an unknown parameter a with a probability $1 - \alpha$, with $1 - \alpha$ also called the confidence level (CL). In

the Frequentist approach, the confidence interval is defined by its coverage. If the true value of a parameter a is measured in a large number N of experiments, then in the limit $N \to \infty$, the fraction $1-\alpha$ of all constructed intervals will contain the true value of a. The construction of a confidence interval has to ensure this coverage for all possible true values of a.

In one dimension an interval $[x_1(a), x_2(a)]$ can be defined for each true value of a, which has to fulfil the relation

$$P(x_1(a) \leq x \leq x_2(a)|a) = \int_{x_1}^{x_2} f(x|a)\, dx = 1-\alpha \tag{7.1}$$

with $f(x|a)$ being the conditional probability density to observe x for a given true value of a. For an observed value x, the borders of the confidence interval for a are a_{min} and a_{max} and then given by:

$$\begin{aligned} a_{max} &= x_1^{-1}(a) \\ a_{min} &= x_2^{-1}(a) \end{aligned}$$

If the true value a is greater than a_{max} or smaller than a_{min} then the probability of a measurement inside this intervall [a_{min},a_{max}] is α.

In cases where the probability density function is a Gaussian with mean a and variance σ_a^2, the interval covered by $\hat{a} \pm \hat{\sigma}_a$ corresponds to the confidence interval with a confidence level of 68.3 %, where $\hat{a}$ and $\hat{\sigma}_a$ are the obtained estimator of the Gaussian probability density function.

For two sided confidence intervals, the probability $1-\alpha$ alone is not sufficient to fully determine a_{min} and a_{max} uniquely. One commonly used definition is the central interval, where the probabilities above and below the confidence interval are equal, i.e.,

$$P(x \leq x_1|a) = P(x \geq x_2|a) = \alpha/2. \tag{7.2}$$

Two other useful forms are the one-tailed limits, namely the upper and lower limits. At a given confidence level, the probability for the upper limit ($a_{min} = -\infty$) is given as

$$P(x \leq x_1|a) = \int_{-\infty}^{x_1} f(x|a)\, dx = 1-\alpha \;\text{ or }\; P(a \leq a_{max}) = 1-\alpha \tag{7.3}$$

and that for the lower limit ($a_{max} = \infty$) as

$$P(x \geq x_2|a) = \int_{x_2}^{\infty} f(x|a)\, dx = 1-\alpha \;\text{ or }\; P(a \geq a_{min}) = 1-\alpha. \tag{7.4}$$

7.3 Confidence Intervals for Poisson Distributions

In a counting experiment, the measurements can only assume integer values, namely the number of observed events (n_{obs}). The distribution of n_{obs} will be a binomial or, for a large number of experiments and small probabilities, a Poisson distribution. For a Poisson distribution with mean λ, the probability to observe n_{obs} events is given by

$$P(n_{obs};\lambda) = e^{-\lambda}\frac{\lambda^{n_{obs}}}{n_{obs}!}. \tag{7.5}$$

An upper limit λ_+ at a given CL of $1-\alpha$ with n_{obs} is given by

$$1-\alpha = \sum_{i=n_{obs}+1}^{\infty} P(i;\lambda_+) \tag{7.6}$$

$$\alpha = \sum_{i=0}^{n_{obs}} P(i;\lambda_+). \tag{7.7}$$

In an analogous way, one can get the lower limit λ_- of the mean value λ by solving the equations

$$\alpha = \sum_{i=n_{nobs}}^{\infty} P(i;\lambda_-) \tag{7.8}$$

$$1-\alpha = \sum_{i=0}^{n_{obs}-1} P(i;\lambda_-). \tag{7.9}$$

However, the above mentioned equations represent an ideal situation, as usually the number of observed events will be a sum of signal and background events, with mean values λ_S and λ_B and a given uncertainty related to the background estimation. The probability to observe n_{obs} events is then given by

$$W(n_{obs}) = \sum_{i=0}^{n_{obs}}\sum_{j=0}^{n_{obs}} P_{sig}(i;\lambda_S)Q(j;\lambda_B)\delta_{i+j,n_{obs}} = \sum_{i=0}^{n_{obs}} P(i;\lambda_S)Q(n_{obs}-i;\lambda_B) \tag{7.10}$$

with $Q(j;\lambda_B)$ the probability distribution of the background events. The sum runs over all possible combinations of background events i and signal events j which sum up to the observed number of events n_{obs}. If the background is also a Poisson distribution, Eq. 7.10 can be written as

$$W(n_{obs}) = \sum_{i=0}^{n_{obs}} P(i;\lambda_S)P(n_{obs}-i;\lambda_B) = P(n_{obs};\lambda_S+\lambda_B) = \frac{e^{-(\lambda_S+\lambda_B)}\cdot(\lambda_S+\lambda_B)^{n_{obs}}}{n_{obs}!} \tag{7.11}$$

From this it follows that the probability α to observe n_{obs} or fewer events is given by

$$\alpha = \sum_{k=0}^{n_{obs}} W(k) = \sum_{k=0}^{n_{obs}} P(k; \lambda_S + \lambda_B) \tag{7.12}$$

which has to be solved for λ_S in order to calculate an upper limit for a given CL.

This method breaks down when the number of observed events is small compared to the expected number of events, especially when no events are observed at all. In this case, one has to modify Eq. 7.12 to [87]

$$\alpha = \frac{\sum_{k=0}^{n_{obs}} P(k; \lambda_S + \lambda_B)}{\sum_{k=0}^{n_{obs}} P(k; \lambda_B)}, \tag{7.13}$$

such that the number of background events is always smaller or equal to the number of observed events.

Recently, the so-called unified approach has been suggested by Feldman and Cousins [88]. This approach introduces two significant improvements compared to classical methods. The first advantage is that it automatically switches of the calculation from a single-sided limit to a two-sided confidence interval. The other improvement was the introduction of a Likelihood ordering scheme, which avoids unphysical confidence intervals.

7.4 Confidence Intervals for the Bayesian approach

In the Bayesian approach, the parameters to extract are treated as random variables. The combined probability density $f(x; a)$ for a parameter a and a random variable x is given by Bayes' theorem,

$$f(x; a) = f(x|a)g(a) = f(a|x)h(x) \tag{7.14}$$

$$f(a|x) = \frac{f(x|a)g(a)}{h(x)} \tag{7.15}$$

where $g(a)$ and $h(x)$ are probability densities that depend only on one variable. The probability density $g(a)$ is the so-called Prior of the observation of x. This means that a probability density for the parameter a is chosen before the experiment is performed and gives the a priori probability for a before the measurement of x. For a given observation x, the conditional probability $f(x|a)$ can be identified with the Likelihood function $L(x, a)$. The probability density $h(x)$ is a multiplicative factor which is independent of a and

can be eliminated by requesting normalisation. In the Bayesian approach the confidence interval can be calculated via

$$f(a|x) = \frac{L(x,a)g(a)}{\int_{-\infty}^{+\infty} L(x,a)\, g(a)da} \tag{7.16}$$

where $f(a|x)$ is the posterior probability for a, and gives the probability density distribution for a after the data x has been obtained. The chosen Prior has to fulfil the normalisation to keep the denominator finite.

In case of a Poisson distribution $P(n_{obs};\lambda_s)$ with n_{obs} observed events and a signal rate of λ_s, one obtains in the Bayesian approach

$$P(\lambda_s|n_{obs}) = \frac{L(n_{obs}|\lambda_s)P(\lambda_s)}{\int_0^{\infty} L(n_{obs}|\lambda_s)P(\lambda_s)d\lambda_s}. \tag{7.17}$$

Equation 7.17 has to be integrated, to obtain an upper limit (or confidence interval) for the unknown parameter λ_s. Including background, λ_B, Eq. 7.17 reads as

$$P(\lambda_s|n_{obs},\lambda_B) = \frac{L(n_{obs}|\lambda_s+\lambda_B)P(\lambda_s)}{\int_0^{\infty} L(n_{obs}|\lambda_s+\lambda_B)P(\lambda_s)d\lambda_s}. \tag{7.18}$$

The values obtained from the Bayesian approach obviously depend on the chosen Prior, i.e., on $P(\lambda_s)$. If the Prior density is unknown, one can assume a uniform density over a certain interval. This flat Prior is, however, just a convention and it would be better to use a class of Priors. In some cases it is suggested to use a Prior which behaves like $a \propto s^{-p}$ or $a \propto e^{-ks}$ where s is directly related to the signal rate. If the chosen Prior for a rare process gives more weight to a region with $s = 0$, the limit will be better or more optimistic compared to a flat Prior. If on the other side, the Prior distribution gives more weight to large values of s, the limit will be worse.

Choosing a flat Prior for a Poisson distribution, the result for the Bayesian confidence interval is the same as in the classical expression of Eq. 7.13. This also holds in the presence of background.

7.5 Confidence Intervals and Systematic Uncertainties

Usually the estimation of the background λ_B has an uncertainty σ_{λ_B} with a probability density that can be assumed as a Gaussian $G(\lambda_B, \sigma_{\lambda_B})$. The upper limit can be obtained

by integrating over the uncertainty σ_{λ_B} of the background. This modifies Eq. 7.13 to

$$\alpha = \frac{\int g(\lambda_B) \sum_{k=0}^{n_{obs}} P(k; \lambda_S + \lambda_B) d\lambda_B}{\int g(\lambda_B) \sum_{k=0}^{n_{obs}} P(k; \lambda_B) d\lambda_B}. \tag{7.19}$$

There exists no analytic recipe for including systematic uncertainties in the construction of the confidence interval. The most common procedure [89] is to use a Bayesian approach, where the systematic uncertainties on the background or signal efficiency are parametrised by e.g. Gaussian distributions and convolved with probability density functions. As an example, consider a counting experiment with detector acceptance A and uncertainty σ_A that expects a background λ_B with an uncertainty σ_{λ_B} and observes n_{obs} events. The upper Poisson limit on the signal λ_S can then be defined as in Eq. 7.13 by integrating over the systematic uncertainties:

$$\alpha = \frac{\sum_{k=0}^{n_{obs}} \frac{1}{2\pi\sigma_{\lambda_S}\sigma_{\lambda_B}} \int_0^\infty \int_0^\infty P(k; \lambda'_S + \lambda'_B) e^{-\frac{(\lambda_B - \lambda'_B)^2}{2\sigma^2_{\lambda_B}}} e^{-\frac{(\lambda_S - \lambda'_S)^2}{2\sigma^2_{\lambda_S}}} d\lambda'_s d\lambda'_B}{\sum_{k=0}^{n_{obs}} P(k; \lambda'_B) e^{-\frac{(\lambda_B - \lambda'_B)^2}{2\sigma^2_{\lambda_B}}} d\lambda'_B}. \tag{7.20}$$

The uncertainty σ_{λ_S} on the signal λ_S has been expressed by $\sigma_{\lambda_S} = \lambda_S \cdot \sigma_A / A$. This integral has to be solved to find an upper limit with CL α on the signal rate λ_S. In practice this integral can be solved by Monte Carlo integration methods.

7.6 Observing fewer Events than Expected

In the case that significantly less events are observed than expected, the unified Frequentist approach [88] (aka Feldman-Cousins) and the Bayesian approach for calculating upper limits differ substantially. In particular, for zero observed events a Bayesian upper limit is independent of the number of expected background events, whereas in the unified Frequentist approach the upper limit decreases when more background is expected. The question which treatment is correct, depends on how one links confidence intervals to probabilities, to either the probability of a hypothesis given an observed set of data (Bayes) or to the probability of observing a set of data given a hypothesis (Frequentist). In case of fewer events observed than expected it is recommended to quote both, the Feldman-Cousins and the Bayesian limit. Another suggestion for such cases is to refer to the sensitivity or expected upper limit, which is a reasonable quantity to judge how large the actual observed limit fluctuated from its mean.

7.7 Expected Upper Limit

If the number of observed events remains unkown, as in the case of a "blind box" analysis, one can calculate the "expected upper limit" without knowing the number of signal events. The "expected upper limit", which is the ensemble average of all expected limits in the absence of a signal for a hypothetical repetition of the experiment, is identical to the "sensitivity"[1] defined in the unified approach of classical confidence interval construction by Feldman and Cousins [88]. Assuming that there is only background λ_B one can calculate for each possible value of observation n_{obs} a 95% C.L upper limit $\mu(n_{obs}, \lambda_B)$. The average upper limit on the number of signal events $\langle\mu(\lambda_B)\rangle$ is then obtained by weighting each limit from the hypothetical ensemble by its Poisson probability of occurrence:

$$\langle\mu(\lambda_B)\rangle = \sum_{n_{obs}=0}^{\infty} \mu(n_{obs}, \lambda_B) \cdot \frac{(\lambda_B)^{n_{obs}}}{(n_{obs})!} \exp(-\lambda_B). \tag{7.21}$$

[1] In [90] it is proposed to refer to this quantity as "exclusion potential".

Chapter 8

Search for the Decay $B_s^0 \to \mu^+\mu^-$

In this Chapter the search for the rare decay $B_s^0 \to \mu^+\mu^-$ will be presented. The measurement of a branching fraction requires the determination of the integrated luminosity or the normalisation to a similar decay with a known branching fraction. For this search the decay of $B^\pm \to J/\psi\, K^\pm$ with $J/\psi \to \mu^+\mu^-$ has been used for normalisation. The search was performed as a "blind box" analysis, with the signal region hidden during all analysis steps. The strategy of the analysis is to start with a pre-selection and then use discriminating variables in an optimisation procedure to further reduce the expected background. The estimation of the background in the signal region is obtained by interpolating the remaining background from sidebands into the signal region.

8.1 Pre-selection Requirements

The B_s^0 candidate event signature for the signal channel under study consists of two muons giving an invariant mass of the B_s^0 mass. Figure 8.1 shows the expected mass resolution from a Monte Carlo simulation for $B_s^0 \to \mu^+\mu^-$. A Gaussian fit yields a value of 88 MeV, whereas the mass difference between B_s^0 and B_d^0 is just 90 MeV [4]. Therefore, the mass resolution is not sufficient to readily separate B_s^0 from B_d^0. Considering the largely suppressed branching fraction, however, the decay $B_d^0 \to \mu^+\mu^-$ was not considered as a potential signal.

Event (pre-)selection started with requiring two muons identified by extrapolating charged tracks reconstructed in the central tracking detectors to the muon detectors, and matching them with hits in the latter. The muons had to form a common 3D-vertex with an invariant mass between 4.5 and 7.0 GeV/c^2 and a χ^2/d.o.f. of less than 10.

The transverse momentum, p_T, of each of the muons was required to be greater than 2.5 GeV/c and $|\eta|$ less than 2.0 to be well inside the fiducial volume of the detector. Tracks

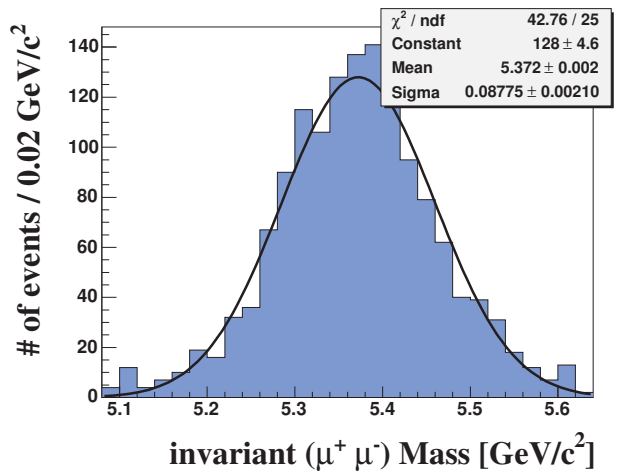

Figure 8.1: The mass resolution in the signal Monte Carlo after pre-selection.

that were matched to each muon were required to have at least three hits in the SMT and four hits in the CFT. To select well measured secondary vertices, the two-dimensional decay length in the plane transverse to the beamline (L_{xy}) was determined, and the error (δL_{xy}) on it was required to be less than 150 μm. L_{xy} was calculated as

$$L_{xy} = \frac{\vec{l}_{Vtx} \cdot \vec{p}_T^B}{p_T^B}, \tag{8.1}$$

where, $\vec{l}_{Vtx}$ represents the three-dimensional vector pointing from the primary vertex to the secondary vertex. The error on the transverse decay length, δL_{xy}, was calculated by taking into account the uncertainties in both the primary and secondary vertex positions. The primary vertex itself was found for each event using a beam spot constrained fit as described in Section 4.2. To ensure a similar p_T dependence of the $\mu^+\mu^-$-system in the signal as in the normalisation channel, the transverse momentum of the candidate B_s, p_T^B, had to be greater than 5 GeV/c. Figure 8.2 shows the $p_T(\mu^+\mu^-)$ spectra for the signal and the normalisation channel $B^\pm \to J/\psi\, K^\pm$ in Monte Carlo after the pre-selection cuts.

In data, a total of 45'696 signal candidate events survived the pre-selection requirements. The effect of the various pre-selection criteria on data events is shown in Table 8.1.

The final event selection was performed using further discriminating variables that were already introduced in Section 6.1.

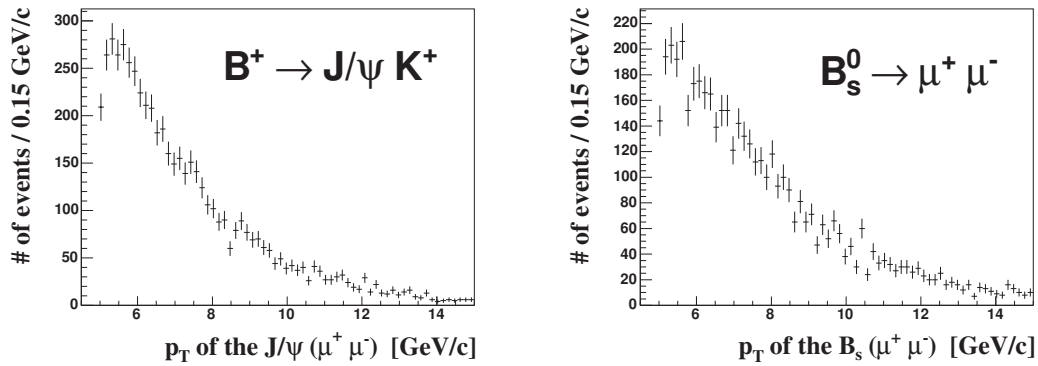

Figure 8.2: The Monte Carlo simulated $p_t(\mu^+\mu^-)$ spectra for the signal (right) and the normalisation channel (left) after the pre-selection.

Cut	Value	#candidates
Good vertex		5'826'444
Mass window (GeV/c^2)	4.5 $< m_{\mu^+\mu^-} <$7.0	615'873
Muon Quality	two medium	399'136
χ^2/d.o.f. of vertex	< 10	253'321
Muon p_T (GeV/c)	> 2.5	154'305
$\lvert\eta\rvert$	< 2.0	149'791
Tracking hits	CFT> 3, SMT > 2	110'099
δL_{xy} (mm)	< 0.15	107'925
B_s^0 Candidate p_T (GeV/c)	> 5.0	45'696

Table 8.1: Number of candidate events surviving the cuts in data used in the pre-selection analysis.

8.2 Optimisation of Discriminating Variables

Before the optimisation of the discriminating variables, the mass region of interest was restricted to 4.53 GeV/c^2 $< M_{\mu^+\mu^-} <$ 6.15 GeV/c^2, containing the signal region. The whole mass region of interest is shifted downward with respect to the world average B_s^0 mass by 30 MeV/c^2 in order to correct for the mass scale of the DØ tracker. The 30 MeV/c^2 mass shift valid at the scale of the B-meson mass was determined by a linear extrapolation of the measured mass shifts between the $J/\psi(\to \mu^+\mu^-)$ and the $\Upsilon(\to \mu^+\mu^-)$ resonances with respect to their PDG-values. This shift by 30 MeV/c^2 is smaller than the expected mass resolution at the B_s^0 mass, which is 90 MeV/c^2.

The signal region is blinded during the optimisation of cuts to avoid overtuning of the cuts into statistical fluctuations as well as biases in the optimisation procedure. Table 8.2

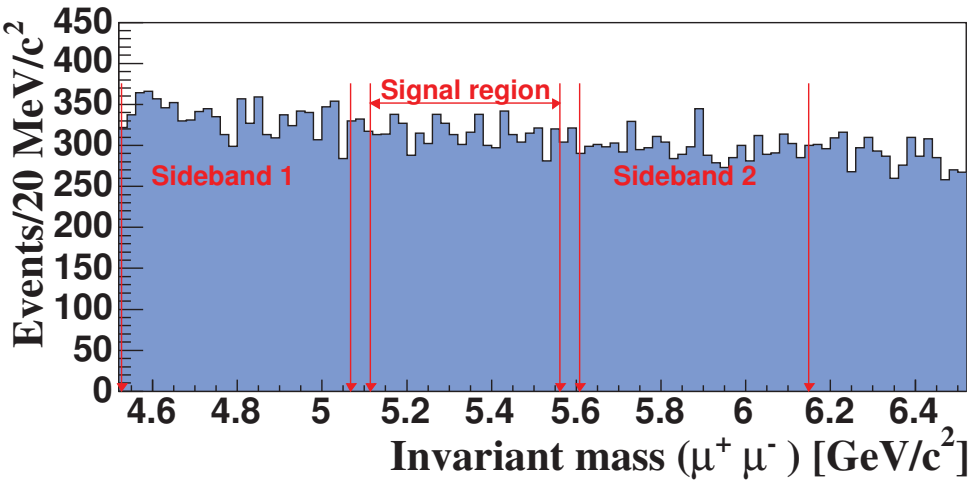

Figure 8.3: Invariant di-muon mass distribution of data events after the pre-selection. A total of 45'696 events remain.

defines the mass ranges for the sidebands and the blinded signal region that were used. The given values translate the size of the blind signal region to a window of ± 270 MeV/c^2 around the expected reconstructed B_s^0 mass. The ± 270 MeV/c^2 corresponds to approximately $\pm 3\sigma$ of the expected mass resolution for $B_s^0 \to \mu^+\mu^-$ in the Monte Carlo. After the cut optimisation the blinded signal region was shrunk to a width of $\pm 2\sigma$ for the calculation of the limit. The inner limits of the sidebands were chosen to be 3σ away from the shifted B_s^0 mass. The width of the sidebands that are used for background estimation was set to 6σ each.

Table 8.2: The different di-muon invariant mass regions for signal and sidebands used for background estimation.

Region	min Mass (GeV/c^2)	max Mass (GeV/c^2)
region of interest	4.530	6.150
blinded signal region during optimisation	5.070	5.610
final blinded signal region for search	5.160	5.520
sideband I	4.530	5.067
sideband II	5.610	6.150

8.2.1 Random Grid Search

With these definitions, the optimisation on the three discriminating variables was performed as described in Section 6.2. The sideband data sample used during this optimisation corresponded to about 90 pb^{-1}, i.e., a third of the total data sample. The result for the cut combination which maximised the figure of merit P of Eq. 6.4 is shown in Table 8.3.

Table 8.3: The optimised set of cuts and their relative Monte Carlo signal efficiencies after maximising the figure of merit P.

cut parameter	cut value	Monte Carlo efficiency (%)
Pointing angle α (rad)	< 0.2	83.4 ± 0.4
Decay length significance	> 18.5	47.5 ± 0.8
Isolation	> 0.56	97.4 ± 0.3

The distributions of the three discriminating variables for signal Monte Carlo and sideband data are shown in Figure 8.4.

The total signal efficiency of the three discriminating cuts relative to pre-selection is 38.6±0.9%. After a linear extrapolation of the sideband population for the whole data sample into the final signal region, an expected number of background events of 4.3±1.2 was obtained. This corresponds to a background rejection of 99.93%.

8.2.2 Likelihood Ratio

In an earlier stage of the analysis, an event classifier based on a probability density ratio (Likelihood Ratio (LHR), see Section 6.3) was studied. The isolation, pointing angle and decay length were used as discriminating variables. The LHR variable R is shown on the left side in Figure 8.6 and the signal efficiency versus background rejection for this Likelihood Ratio is shown on the right side. With a different data sample of ≈ 200 pb^{-1} it was found that the background rejection is worse (99.91%) than for an analogous optimisation using the RGS (99.985%). This reduced background rejection would increase the number of expected background events in the data sample quite significantly leading to almost 36 background events in the signal region, compared to three for the RGS. Although the achieved signal efficiency would have been increased from 21% in the RGS to 43%, it was therefore decided not to pursue this optimisation method further.

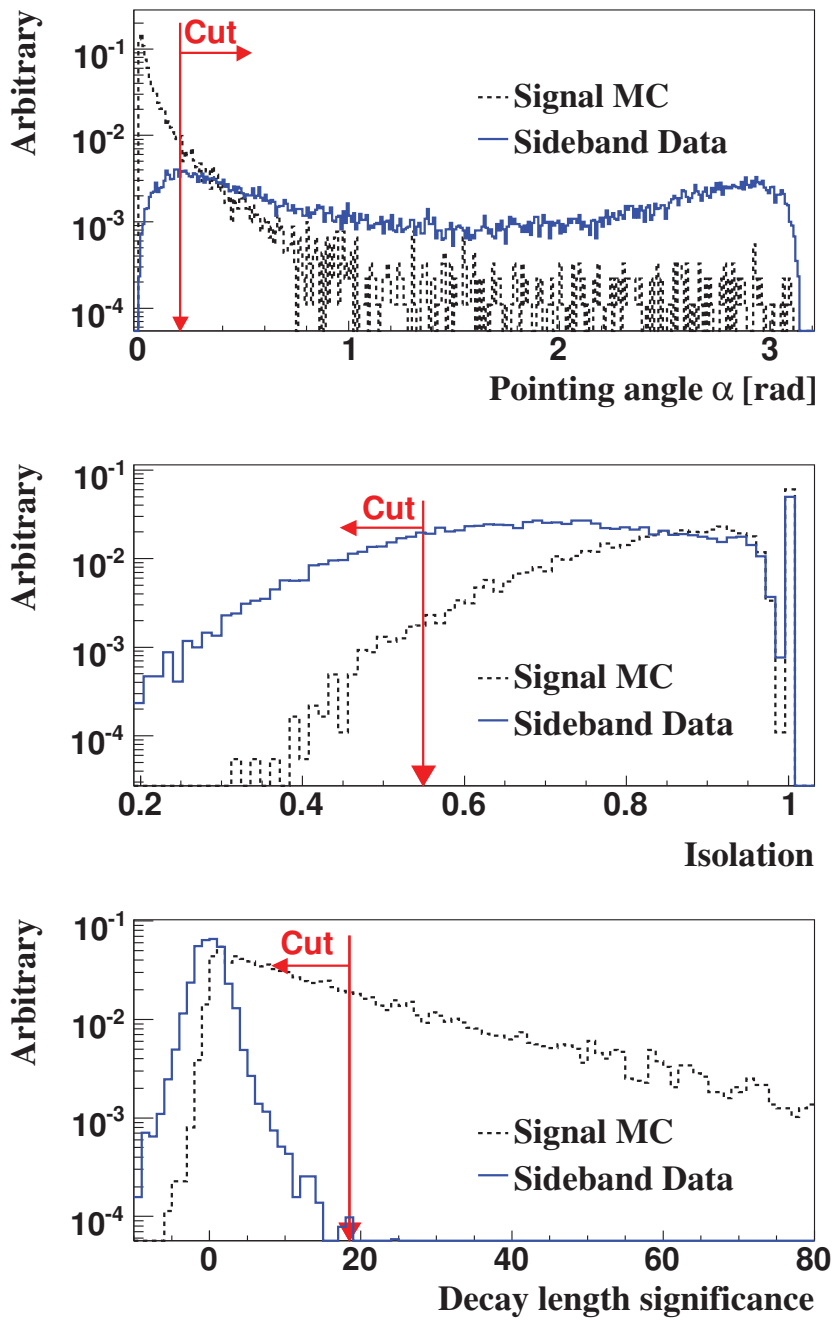

Figure 8.4: Discriminating variables after the pre-selection for signal Monte Carlo and data events from the sidebands. The arrows indicate the cut values that were obtained after optimisation. The normalisation is done on the number of signal Monte Carlo and sideband data events after pre-selection.

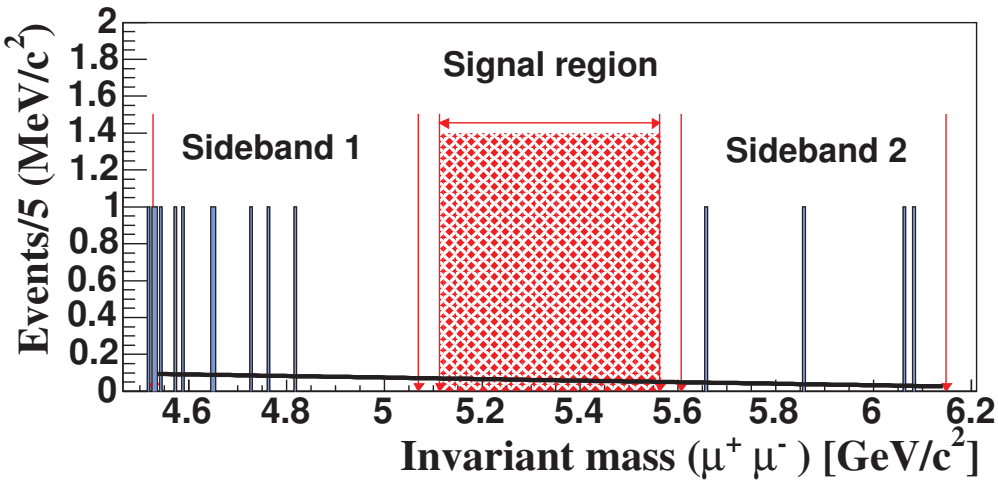

Figure 8.5: The remaining background for the full data sample with the standard discriminating variables. The line shows the linear fit to determine the expected background.

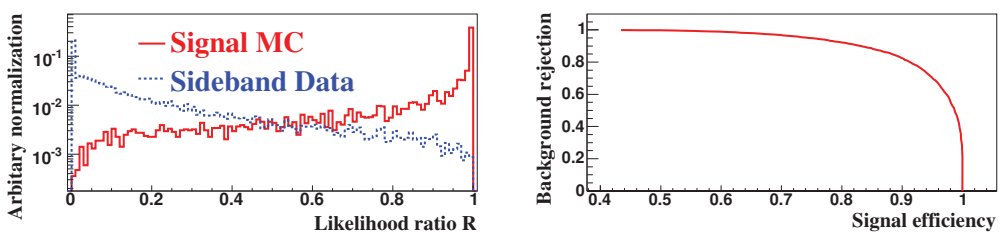

Figure 8.6: The LHR variable R on the left side and the signal efficiency versus background rejection for this Likelihood Ratio on the right side.

8.3 The Normalisation Channel $B^\pm \to J/\psi\, K^\pm$

In order to obtain a branching ratio limit for $B_s^0 \to \mu^+\mu^-$, the decay chain $B^\pm \to J/\psi\, K^\pm$ with $J/\psi \to \mu^+\mu^-$ was used as normalisation. The J/ψ decaying into $\mu^+\mu^-$ has the big advantage that the $\mu^+\mu^-$ efficiencies cancel to a large extent. Therefore, the same cuts on the discriminating variables were applied to the muons from J/ψ decays from $B^\pm$ as in the $B_s^0 \to \mu^+\mu^-$ search.

The standard selections for $B^\pm$ mesons were applied as follows: The J/ψ vertex fit of the two muons was required to have a χ^2 per d.o.f. of not more than 10, similar to the $\mu^+\mu^-$ vertex criterion in the B_s^0 search. The combined vertex fit of the J/ψ and the additional $K^\pm$ should not yield a χ^2 for three d.o.f of more than 20. The p_T of the $K^\pm$ should be larger than 0.9 GeV/c. Moreover, a collinearity between the decay length vector

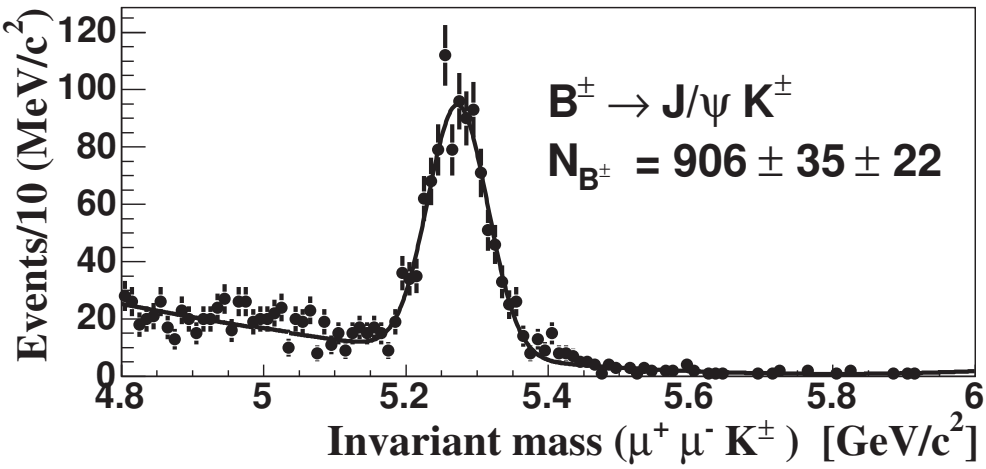

Figure 8.7: The normalisation channel $B^{\pm} \to J/\psi\, K^{\pm}$ for the full data sample after all final cuts.

of the $B^{\pm}$ meson and the combined momentum of J/ψ and $K^{\pm}$ in the transverse plane of at least 0.9 is required.

The mass spectrum of the reconstructed $B^{\pm} \to J/\psi\, K^{\pm}$ for the full data sample is shown in Figure 8.7. A fit using a Gaussian function for the signal and a second order polynomial for the background yielded 906 $\pm$ 35 $\pm$ 22 signal events, where the first uncertainty is statistical and the second due to systematics estimated by varying the fit range and background shape hypothesis. The mass resolution of the $B^{\pm}$ is 40.7 MeV/c^2, which agrees reasonably well with the resolution of 36 MeV/c^2 obtained in the Monte Carlo simulation. The $B^{\pm}$ mass for the $B^{\pm} \to J/\psi\, K^{\pm}$ selection is in the data 5278 MeV/c^2, which is just 1 MeV/c^2 below the PDG mean value of the $B^{\pm}$ meson.

The selection efficiencies for data and Monte Carlo events for the cuts on the discriminating variables are shown in Table 8.4 and show a reasonably good agreement. The data efficiencies were estimated using the fit procedure described above after various cuts and are relative efficiencies. The distributions of the variables are shown in Figure 8.8.

Instead of $B^{\pm} \to J/\psi\, K^{\pm}$, one could also use $B_s^0 \to J/\psi\, \phi$ as normalisation channel. Normalising to the decay mode $B^{\pm} \to J/\psi\, K^{\pm}$ is , however, preferable, since this mode yields higher statistics and the branching ratio and lifetime are well known from measurements at CLEO and the asymmetric B factories. In addition, understanding the efficiency to detect $B_s^0 \to J/\psi\, \phi$ events is complicated by the presence of different CP components which have different lifetimes. Finally, normalising to the mode $B_s^0 \to J/\psi\, \phi$ does not completely eliminate the systematic uncertainty due to the fragmentation ratio, since all current calculations of the branching ratio for this mode use this ratio.

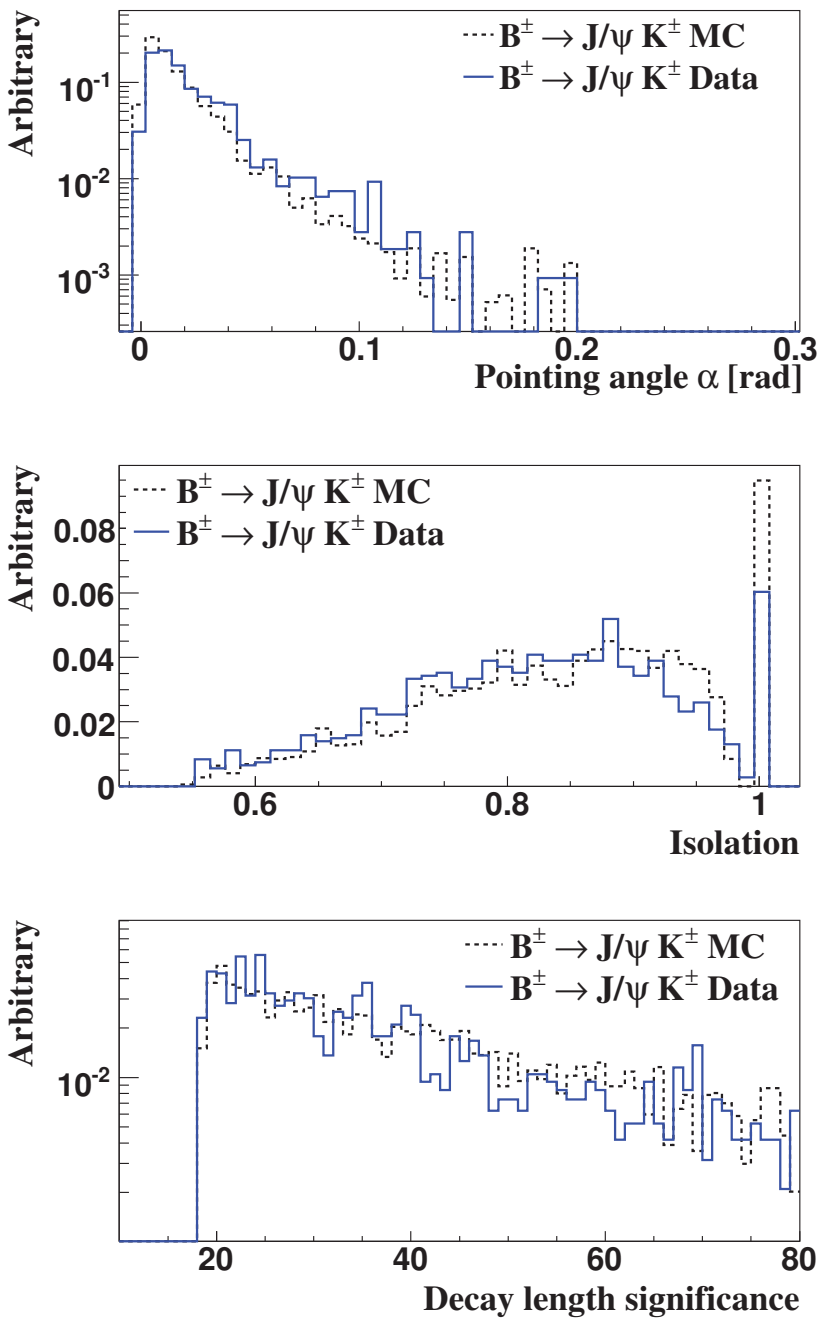

Figure 8.8: Comparison of the discriminating variables for $B^\pm \to J/\psi\, K^\pm$ Monte Carlo and data events after all final cuts. Distributions were normalised to give the same integral.

Table 8.4: Efficiencies for the discriminating variables of data and Monte Carlo events for the normalisation channel $B^{\pm} \to J/\psi\, K^{\pm}$.

cut parameter	cut value	Data efficiency (%)	Monte Carlo efficiency (%)
Pointing angle α (rad)	< 0.2	89.7 ± 0.8	94.1 ± 0.5
Decay length significance	> 18.5	56.5 ± 1.8	59.2 ± 1.4
Isolation	> 0.56	94.0 ± 0.9	96.3 ± 0.5

8.4 Deriving an Upper Limit

To calculate an upper limit on the decay, $B_s^0 \to \mu^+\mu^-$, the number $N_{B^{\pm}}$=906±35±22 of reconstructed events of the decay $B^{\pm} \to J/\psi\, K^{\pm}$ have been used as normalisation, as explained in Section 8.3. The branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ can then be calculated as (see Appendix A for the derivation):

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \cdot \left(1 + R \cdot \frac{\epsilon^{B_d^0}_{\mu^+\mu^-}}{\epsilon^{B_s^0}_{\mu^+\mu^-}} \cdot \frac{b \to B_d^0}{b \to B_s^0}\right) = \tag{8.2}$$
$$\frac{\mu(n_{obs}, n_{back})}{N_{B^{\pm}}} \cdot \frac{\epsilon^{B^{\pm}}_{\mu^+\mu^- K^{\pm}}}{\epsilon^{B_s^0}_{\mu^+\mu^-}} \cdot \frac{b \to B^{\pm}}{b \to B_s^0} \cdot \mathcal{B}(B^{\pm} \to J/\psi\, K^{\pm}) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-)$$

where

- $\mu(n_{obs}, n_{back})$ is the upper limit on the number of observed events at a chosen CL for n_{obs} observed events with n_{back} expected background events;
- $\epsilon^{B_s^0}_{\mu^+\mu^-}$ and $\epsilon^{B^{\pm}}_{\mu^+\mu^- K^{\pm}}$ are the efficiencies for the signal and normalisation channels, obtained from Monte Carlo simulations;
- $b \to B_s^0$, $b \to B^{\pm}$ and $b \to B_d^0$ are the fragmentation fractions of b or $\bar{b}$ quark to a B_s^0, a $B^{\pm}$ or a B_d^0 meson, respectively. The ratio which enters in the equation was calculated using the latest world average values [4] for the fragmentation for $B_{u,d}$ and B_s^0 mesons, respectively[1]. For the uncertainty on the fragmentation ratio, a full anti-correlation between the two was assumed and $(b \to B_{u,d})/(b \to B_s^0) = 3.70 \pm 0.47$ has been obtained. This fragmentation ratio was determined by measurements

[1]Note that the Monte Carlo uses simply $(b \to B_s^0)/(b \to B_{u,d}) = 3/10$ which need to be corrected for when calculating $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ from equation 8.2

from LEP, SLD and Tevatron. If only measurements of the fragmentation fractions from the Tevatron are used, the average is $(b \to B_{u,d})/(b \to B_s^0) = 3.32 \pm 0.59$. This is slightly different from the world average value and is dominated by the comparison of the mixing probabilities between B_s^0 and B_d^0 mesons, χ_s and χ_d. Using only Tevatron data alone would improve the upper limit by 15%;

- $\mathcal{B}(B^\pm \to J/\psi\, K^\pm) = (1.00 \pm 0.04) \times 10^{-3}$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.1)\%$ [4]; and

- $R \cdot \epsilon^{B_d^0}_{\mu^+\mu^-}/\epsilon^{B_s^0}_{\mu^+\mu^-}$ is the branching fraction ratio $\mathcal{B}(B_d^0)/\mathcal{B}(B_s^0)$ of $B_{d,s}^0$ mesons decaying into two muons multiplied with their efficiency ratio. In the Standard Model, R is given by $R = \tau(B_d^0)/\tau(B_s^0) \cdot m_{B_d^0}/m_{B_s^0} \cdot f^2_{B_d^0}/f^2_{B_s^0} \cdot |V_{td}|^2/|V_{ts}|^2$, where $\tau_{B_{d,s}^0}$ are the lifetimes of the B-mesons and $f_{B_{d,s}^0}$ the meson decay constants. The relation on the branching fraction ratio holds in fact for all new physics models that are obeying Minimal Flavour Violation. The ratio $\epsilon^{B_d^0}_{\mu^+\mu^-}/\epsilon^{B_s^0}_{\mu^+\mu^+}$ was determined from simulation to be 0.92 ± 0.04, with the uncertainty due to limited Monte Carlo statistics.

To simplify the calculation of the upper limit on the branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ it was assumed that there are no contributions from $B_d^0 \to \mu^+\mu^-$ decays ($R \approx 0$) in the search window centred around the B_s^0 mass. This assumption is acceptable since the decay is suppressed by $|V_{td}/V_{ts}|^2 \approx 0.04$ in the Standard Model. Any non-negligible contribution due to B_d^0 decays ($R > 0$) would make the obtained branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ as given in Eq. 8.2 smaller. Thus, the presented limit for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is in that context conservative.

The efficiencies $\epsilon^{B_s^0}_{\mu^+\mu^-}$ and $\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}$ are the global signal efficiencies for the search signal and normalisation channel, respectively, including the pre-selection cuts and acceptance. They were determined from Monte Carlo to be $\epsilon^{B_s^0}_{\mu^+\mu^-} = (7.02 \pm 0.12) \times 10^{-4}$ and $\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm} = (5.06 \pm 0.14) \times 10^{-4}$, where the uncertainties are due to Monte Carlo statistics. The quoted efficiencies refer to Monte Carlo events that passed the trigger simulation and hence include the trigger efficiency. The value for the raw efficiency ratio is found to be $\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}/\epsilon^{B_s^0}_{\mu^+\mu^-} = 0.71 \pm 0.02$. Note that the efficiency ratio refers to Monte Carlo generated $b\bar{b}$ events and has to be corrected for the used fragmentation ratio of 3/10 used in the Monte Carlo simulation in order to calculate the limit of Eq. 8.2.

The p_T distribution of the $B^\pm$ in data was found to have a slightly harder spectrum than that from Monte Carlo simulation. The p_T distribution of B candidate data events from the normalisation channel in data was compared with the Monte Carlo simulation prediction and a weighting function was obtained. The Monte Carlo events of the signal and normalisation channels were reweighted accordingly, to ensure good agreement between data and Monte Carlo simulation. The efficiencies were calculated using the weighted events.

Table 8.5: Fitted mass resolutions as predicted by Monte Carlo simulations and as seen in the data for J/ψ and Υ(1S).

Mode	σ_{Data} [MeV]	σ_{MC} [MeV]	Ratio $\sigma_{Data}/\sigma_{MC}$
J/ψ	72.0	55.0	1.31
J/ψ (from $B^{\pm} \to J/\psi\, K^{\pm}$)	77.3	61.2	1.26
Υ(1S)	226.2	176.3	1.28

In addition, the observed widths of known $\mu^+\mu^-$ resonances (J/ψ and $\Upsilon(1S)$) are (27±4)% larger than predicted by Monte Carlo, as shown in Table 8.5. This large discrepancy is due to an insufficient description of the passive detector material used in the Monte Carlo simulation after the upgrade from Run I to Run II. The $\pm 2\sigma_{MC}$ signal mass region using the Monte Carlo mass resolution therefore corresponds to $\pm 1.58\sigma_{Data}$ when the data mass resolution is considered, and the efficiency was corrected accordingly. The final corrected value for the efficiency ratio was then found to be $\epsilon^{B^{\pm}}_{\mu^+\mu^-K^{\pm}}/\epsilon^{B^0_s}_{\mu^+\mu^-} = 0.247 \pm 0.009\,(\mathrm{stat})$, where the uncertainty is due to Monte Carlo statistics. The systematic uncertainty associated with the weighting procedure and efficiency correction is discussed in Section 8.6.

8.5 Sensitivity

The analysis has used the technique of a "blind box" analysis, which means that the signal region was kept hidden throughout all analysis steps presented so far. Before exploring the "blinded" invariant di-muon signal mass region, one can calculate an "expected upper limit", as introduced in Section 7.7, without knowing the actual number of signal $B^0_s \to \mu^+\mu^-$ events. In Eq. 8.2, the upper limit $\mu(n_{obs}, n_{back})$ has to be replaced by $\langle\mu(n_{back})\rangle$. Using the numbers presented in Section 8.4, without their statistical uncertainties, one obtains a sensitivity of

$$\langle\mathcal{B}(B_s \to \mu^+\mu^-)\rangle = 3.6 \times 10^{-7} \text{ at a 95\% CL.} \tag{8.3}$$

After the cut optimisation and calculation of the sensitivity, its dependence on the choice of the width for the final blinded signal region was investigated. The number of expected background events from a linear interpolation from the sidebands was calculated for signal regions varying from $\pm 1.5\sigma_{MC}$ to $\pm 2.5\sigma_{MC}$ in width, and the expected sensitivity was calculated. The maximial sensitivity for the cut conditions quoted in Table 8.3 was achieved for a final signal width of about $\pm 2\sigma_{MC}$, which was just the default value. The expected branching ratios are listed in Table 8.6 and exclude the uncertainties and

corrections to the efficiency ratio, that were obtained for the different widths of the signal region. The values presented in Table 8.6 refer to a reduced data set of 240 pb^{-1}, with an expected background of 3.7 ± 1.1 events.

Table 8.6: The sensitivity and total Monte Carlo signal efficiency for signal regions of different width.

width σ_{MC}	n_{back}	$\langle\mathcal{B}\rangle$ for 95% CL	MC efficiency (%)
± 2.5	4.6	3.8×10^{-7}	$(73.2 \pm 1.2) \times 10^{-5}$
± 2.0	3.7	3.6×10^{-7}	$(70.2 \pm 1.2) \times 10^{-5}$
± 1.5	2.8	3.7×10^{-7}	$(63.5 \pm 1.4) \times 10^{-5}$

8.6 Systematic Uncertainties

All relative uncertainties that enter into the calculation of $\langle\mathcal{B}\rangle$ are listed in Table 8.7. The largest individual contribution is almost 13%, due to the fragmentation ratio $(b \to$

Table 8.7: The relative uncertainties for calculating an upper limit of $\mathcal{B}$

Source	Relative Uncertainty [%]
$\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}/\epsilon^{B_s^0}_{\mu^+\mu^-}$	7.7
# of $B^\pm \to J/\psi\, K^\pm$	5.0
$\mathcal{B}(B^\pm \to J/\psi\, K^\pm)$	4.0
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	1.7
$f_{b\to B_s^0}/f_{b\to B^\pm}$	12.7
background uncertainty	27.9

$B^\pm)/(b \to B_s^0)$. To calculate the error on the fragmentation ratio, it was conservatively assumed that the individual fragmentations $(b \to B^\pm)$ and $(b \to B_s^0)$ are fully anti-correlated. The relative statistical uncertainties on $\epsilon^{B_s^0}_{\mu^+\mu^-}$ and $\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}$ are 1.7% and 2.8% respectively. They were combined into one efficiency uncertainty assuming no correlations. An additional systematic uncertainty was assigned to this ratio: The different muon p_T distribution between $J/\psi \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ decays (2.3%), the tracking inefficiency for the additional Kaon track with respect to Monte Carlo (1%) , the uncertainty on trigger and muon identification between data and Monte Carlo (0.7%), the weighting procedure (6.4%) and the different resolutions (1.3%). Within the errors,

the Monte Carlo correctly reproduces the efficiency of the cuts on the discriminating variables when applied to the normalisation channel. This yields, with the already mentioned correction due the data resolution in data, a final value for the efficiency ratio of $\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}/\epsilon^{B^0_s}_{\mu^+\mu^-} = 0.247 \pm 0.009\,(\mathrm{stat}) \pm 0.017\,(\mathrm{sys})$.

The relative uncertainty on the number of background events is 27.9% and the $B^\pm$ normalisation channel has a relative uncertainty of 5.0%, including statistical and systematic effects.

The statistical uncertainty on the expected number of background events, as well as the uncertainties on signal and background efficiencies, can be folded into the limit/sensitivity calculation of Eq. 8.2 by integrating over probability functions which parameterise these uncertainties. A prescription [91] for constructing a Frequentist confidence interval with the Feldman and Cousins ordering scheme for the Monte Carlo integration was used. The background was modelled as a Gaussian distribution with its mean value equal to the expected number of background events and its sigma equal to the background uncertainty. The signal and background efficiency uncertainties were considered as Gaussian distributions assuming a full correlation between the two. The uncertainty on the number of $B^\pm$ events was propagated into the signal and background efficiency uncertainties. The relative errors on the fragmentation ratio and on the branching ratios were taken into account. The resulting sensitivity, including all the above mentioned statistical and systematic uncertainties at a 95% CL, is then

$$\langle \mathcal{B}(B^0_s \to \mu^+\mu^-)\rangle = 4.2 \times 10^{-7} \text{ at 95\% CL.} \tag{8.4}$$

8.7 Results

After a detailed review of the analysis by the DØ collaboration, the approval to explore the signal region was given by the DØ physics coordinator. In the spirit of a "blind box" analysis, no cut value was changed after the box was opened, only more data was added and therefore the background estimation changed.

Four events were found in the $\pm 2\sigma_{MC}$ signal window, while 4.3 ± 1.2 events were expected due to background. The properties of the four signal events are listed in Table 8.8 and the mass distribution inside the unblinded signal region is shown in Figure 8.9.

The four events were checked for any anomalous behaviour by comparing their characteristics to background events obtained from the sidebands. It was found that the four signal events do not exhibit any distinctive behaviour in any of the plotted kinematical variables and therefore they are completely consistent with background events. In the absence of an apparent signal, a limit on the branching fraction $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ (and $\mathcal{B}(B^0_d \to \mu^+\mu^-)$) was computed.

Table 8.8: The four signal events that were found after opening the box.

Run#	Event#	Mass (GeV)	Pointing angle α (rad)	$L_{xy}/\delta L_{xy}$	Isolation I
167190	1921022	5.416	0.147	37.7	0.591
176591	2307148	5.444	0.037	21.2	0.912
177820	13141437	5.341	0.009	29.2	0.682
188905	39723538	5.347	0.033	34.4	0.652

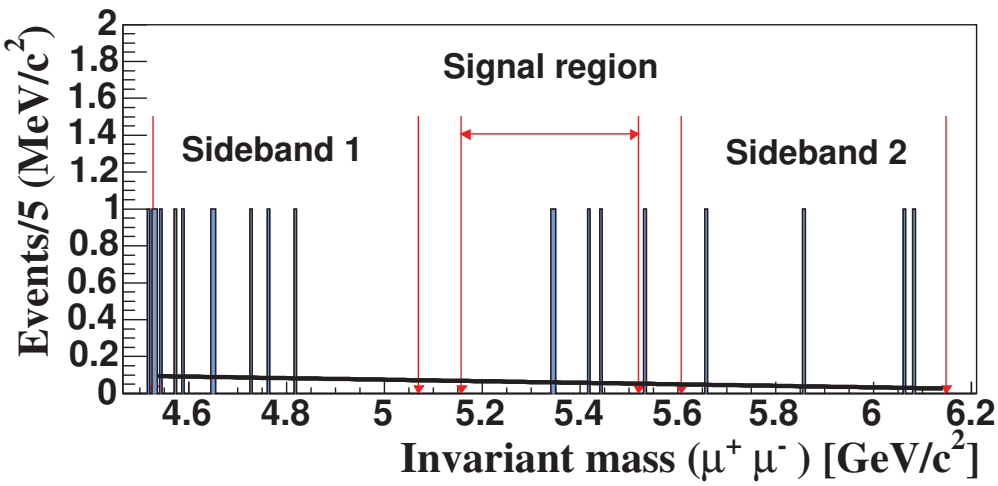

Figure 8.9: The mass distribution after opening the box for the data sample with the standard discriminating variables. Four events were found in the $\pm 2\sigma$ region.

The resulting limit on the branching fraction at a 95% (90%) CL, including all statistical and systematic uncertainties, is

$$\boldsymbol{\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 3.7 \times 10^{-7}\ (3.0 \times 10^{-7})}.$$

Using a Bayesian approach with flat prior and Gaussian (smeared) uncertainties [92] a limit of

$$\mathcal{B}(B_s^0 \to \mu^+ \mu^-) \leq 4.0 \times 10^{-7}\ (3.2 \times 10^{-7})$$

at the 95% (90%) CL was obtained. These limits represent an improvement of the previously published values by DØ at the 95% (90%) CL of

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 5.0 \times 10^{-7}\ (4.1 \times 10^{-7}).$$

obtained with a DØ Run II dataset of 240 pb^{-1} [44].

The analysis is also sensitive to $B_d^0 \to \mu^+\mu^-$. To calculate an upper limit on the branching ratio for $B_d^0 \to \mu^+\mu^-$, it was assumed that there is no contribution from the decay $B_s^0 \to \mu^+\mu^-$. The formula given in Eq. 8.2 was modified to

$$\mathcal{B}(B_d^0 \to \mu^+\mu^-) = \frac{n_{ul}}{N_{B^\pm}} \cdot \frac{\epsilon^{B^\pm}_{\mu^+\mu^- K^\pm}}{\epsilon^{B_s^0}_{\mu^+\mu^-}} \cdot \frac{\epsilon^{B_s^0}_{\mu^+\mu^-}}{\epsilon^{B_d^0}_{\mu^+\mu^-}} \cdot \frac{b \to B^\pm}{b \to B_d^0} \cdot \mathcal{B}(B^\pm \to J/\psi\, K^\pm) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-) \tag{8.5}$$

with the same notations as above. The factor $b \to B^\pm / b \to B_d^0$ was taken to be unity. The event upper limit, n_{ul}, was calculated including the additional uncertainty on the efficiency ratio $\epsilon^{B_s^0}_{\mu^+\mu^-}/\epsilon^{B_d^0}_{\mu^+\mu^-}$. This yields and upper limit on the branching fraction $\mathcal{B}(B_d^0 \to \mu^+\mu^-)$ at a 95% CL of

$$\mathcal{B}(B_d^0 \to \mu^+\mu^-) \leq 11.1 \times 10^{-8}.$$

Chapter 9

Search for the Rare Decay $B_s^0 \to \phi\,\mu^+\mu^-$

In this Chapter, the search for the rare decay $B_s^0 \to \phi\,\mu^+\mu^-$ will be presented. For this search, the decay of $B_s^0 \to J/\psi\,\phi$, with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$, was used for normalisation. The search was performed again as a "blind box" analysis, with the signal region hidden during all analysis steps. The strategy of the analysis was to start with a pre-selection and then use discriminating variables in an optimisation procedure to further reduce the expected background. The estimation of the background in the signal region was obtained by interpolating the remaining background from sidebands into the signal region.

9.1 Introduction

As described in Section 2.4.2, the Standard Model prediction for the decay rate of the exclusive FCNC decay depends on the assumed form factor model. Also the shape of the invariant di-muon mass depends on the form factor model used. For the signal Monte Carlo events, a decay model in EvtGen [82] was used which includes next-to-next-to-leading logarithmic (NNLO) improved Wilson coefficients [21] for the short-distance part. The form factors were obtained from QCD light cone sum rules taken from Ali *et al.* [22]. Just recently, Ball *et al.* presented an improved form factor calculation [93]. A comparison of the invariant mass spectra for the two muons in these two models is shown in Figure 9.1. The difference reaches about 20% for $m_{\mu^+\mu^-} < 1$ GeV/c^2, while elsewhere it remains well below 10%. The Standard Model decay rate of $B_s^0 \to \phi\,\mu^+\mu^-$ is predicted to be of the order of 1.6×10^{-6} [94].

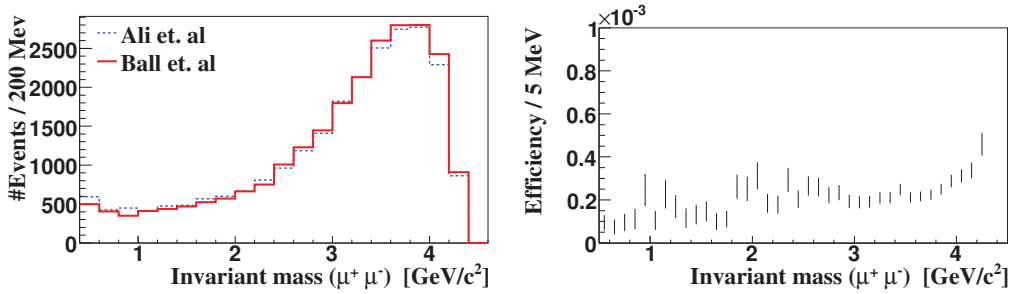

Figure 9.1: Invariant mass spectra generated with the two decay models [21, 93] before (left) all selection cuts and the efficiency as function of the di-muon mass (right) after all cuts for the decay model of Ali *et al.* [22].

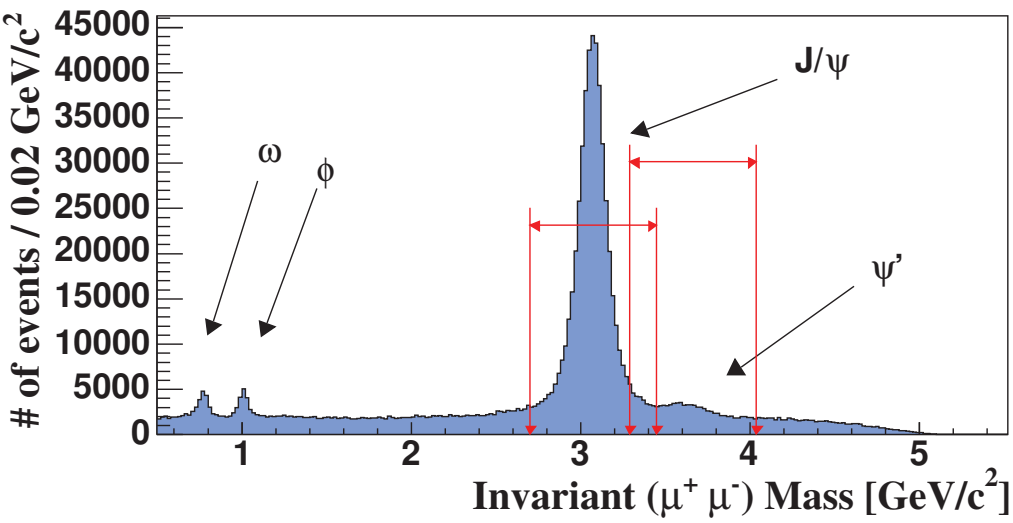

Figure 9.2: Di-muon invariant mass distribution of the loose B_s^0 candidate selection.

9.2 Pre-selection Requirements

The pre-selection starts with a loose selection of B_s^0 candidates consisting of two muons and two oppositely charged tracks forming a good vertex. In this first selection step, the mass of the two kaon candidate tracks was required to be between 0.980 GeV/c^2 $< m_\phi <$ 1.080 GeV/c^2 and the mass of the B_s^0 candidate was required to be within 4.4 GeV/c^2 $< m_B <$ 6.2 GeV/c^2. The di-muon invariant mass distributions for the pre-selected events is shown in Figure 9.2

Next, it was required that the invariant mass of the two muons is 0.5 GeV/c^2 $< m_{\mu^+\mu^-} <$ 4.4 GeV/c^2. In this mass region, the $J/\psi(\rightarrow \mu^+\mu^-)$ and $\psi(2S)(\rightarrow \mu^+\mu^-)$ resonances are excluded with cut-out regions that cover $\pm 5\sigma$ wide windows around the observed

resonance masses as indicated in Figure 9.2. The J/ψ mass resolution obtained in data is $\sigma = 75$ MeV/c^2 while the Monte Carlo resolution is about 25% better than that.

The two muons had to be of medium quality, requiring two of the three muon layers to have fired and a central track to be matched. The χ^2/d.o.f. of the two-muon vertex was required to be χ^2/d.o.f.<10. The tracks that are matched to each muon leg had to have at least three hits in the SMT and four hits in the CFT. In addition, the transverse momentum of each of the muons was required to be greater than 2.5 GeV/c and their pseudorapidity η had to be $|\eta| < 2.0$ in order to be well inside the fiducial volume of the tracking and muon detectors. The error on the transverse decay length, δL_{xy}, was required to be smaller than 150 μm.

In the following, the number of B_s^0 candidates was further restricted by requiring $p_T^B > 5$ GeV/c and asking the B_s^0 candidate vertex to fulfil $\chi^2 < 36$ for five d.o.f. The two tracks that are combined with the two muons to the B_s^0 candidate should each have $p_T > 0.7$ GeV/c and each need at least one hit in the SMT. The two tracks forming the ϕ candidate were further restricted in their invariant mass to be within 1.008 GeV/c^2 $< m_\phi <$ 1.032 GeV/c^2. The successive cuts and the number of candidate events surviving each cut are shown in Table 9.1.

Table 9.1: Number of candidate events in data surviving the cuts used in the pre-selection analysis.

Cut	Value	#candidates
Good vertex		1555320
Mass region (GeV/c^2)	$0.5 < m_{\mu^+\mu^-} <$4.4 excl. J/ψ,$\psi(2S)$	530892
Muon Quality	two medium	276875
χ^2/d.o.f. of vertex	< 10	127509
Muon p_T (GeV/c)	> 2.5	73555
Muon $\|\eta\|$	< 2.0	72350
Tracking hits	CFT> 3, SMT > 2	58012
δL_{xy} (mm)	< 0.15	54752
B_s^0 Candidate p_T (GeV/c)	> 5.0	54399
B_s^0 χ^2 vertex	< 36	53195
Kaon p_T (GeV/c)	> 0.7	9639
ϕ mass (GeV/c^2)	$1.008 < m_\phi <$1.032	2602

9.3 Optimisation of Discriminating Variables

Before the optimisation of the discriminating variables, the mass region of interest was restricted to 4.51 GeV/c^2 $< M_{\mu^+\mu^-} <$ 6.13 GeV/c^2, containing the signal region around the PDG [4] world average value of the B_s^0 mass of $m_{B_s^0} = 5369.6 \pm 2.4$ MeV/c^2. The mass region of interest was shifted downward with respect to the world average B_s^0 mass by 44 MeV/c^2 in order to correct the mass scale of the DØ tracker. The 44 MeV/c^2 mass shift was taken from the mean B_s^0 mass obtained from the fit to the $B_s^0 \to J/\psi\,\phi$ mass spectra without constraining the di-muon pair to the J/ψ mass, as will be discussed in Section 9.4.

In this analysis, the signal region was again blinded and sufficiently well separated from the sidebands. Table 9.2 defines the regions for the sidebands and the blinded signal region that were used.

The selected values translate into a width of the blinded signal region of $\pm$ 270 MeV/c^2 around the (shifted) world average mass value of the B_s^0. The expected mass resolution for $B_s^0 \to \phi\,\mu^+\mu^-$ from Monte Carlo is $\approx$ 75 MeV/c^2, the 270 MeV/c^2 window corresponds therefore to $\pm 3.6\sigma$ [95].

After the optimisation of the discriminating cuts the blinded signal region was shrunk to a $\pm$ 225 MeV/c^2 ($\pm 3\sigma$) window for the calculation of the sensitivity. The inner limits of the sidebands was chosen such that they are 270 MeV/c^2 away from the expected B_s^0 mass. The width of each of the sidebands that were used for the background estimation was set to 540 MeV/c^2.

Table 9.2: The different four track invariant mass regions for signal and sidebands used for background estimation.

Region	min Mass (GeV/c^2)	max Mass (GeV/c^2)
region of interest	4.5156	6.1356
blinded signal region during optimisation	5.0556	5.5956
final blinded signal region for limit	5.1381	5.5131
sideband I	4.5156	5.0556
sideband II	5.5956	6.1356

With these definitions, the optimisation on the three discriminating variables, isolation, pointing angle and decay length significance, was performed as described in Section 6.2. The signal Monte Carlo for this optimisation included the charmonium resonances to increase the number of cut combinations. This can be justified, since the discriminating variables do not depend on the invariant di-muon mass. The resulting cut

Table 9.3: The optimised cuts and their relative Monte Carlo signal efficiencies including their statistical uncertainties after maximising P.

cut parameter	value	Monte Carlo eff. (%)
Opening angle (rad)	< 0.1	81±4
Decay length significance	> 10.3	73±4
Isolation	> 0.72	92±5

values that were obtained from optimising the figure of merit P of Eq. 6.4 are listed in Table 9.3.

The distributions of the three discriminating variables is shown in Figure 9.3 for signal Monte Carlo and sideband data.

The total signal efficiency relative to pre-selection of the three discriminating cuts was found to be $(54 \pm 3)\%$ with the uncertainty due to Monte Carlo statistics. After a linear interpolation of the sideband population into the final signal region an expected number of background events of 1.6±0.4 for the whole data sample was obtained. Figure 9.4 shows the invariant mass distribution for the remaining background events populating the lower sideband.

9.4 The Normalisation Channel $B_s^0 \to J/\psi\,\phi$

In order to obtain a branching ratio limit for $B_s^0 \to \phi\,\mu^+\mu^-$, events from the decay $B_s^0 \to J/\psi\,\phi$ with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ were used as normalisation. As mentioned above, the same cuts were applied to the $B_s^0 \to J/\psi\,\phi$ candidates. In addition, the two muons were constrained to have an invariant mass equal to the J/ψ mass [4]. This approach assumes that the non-resonant signal will be negligible compared to the resonant, since the non-resonant invariant mass could fall inside the J/ψ region and therefore appear in the $B_s^0 \to J/\psi\,\phi$ signal region. In Figure 9.5 and Figure 9.6, the invariant masses for the ϕ and J/ψ candidates are shown for data and Monte Carlo, respectively. The indicated fit in all plots is the sum of a Gaussian for the signal plus a linear function for the background description. The obtained resolution for the ϕ candidate in data (4.2 MeV/c^2) compares well to the value of the Monte Carlo simulation (4.3 MeV/c^2) .

The reconstructed $B_s^0 \to J/\psi\,\phi$ mass spectrum for the full data sample is shown in Figure 9.7. A fit using a Gaussian function for the signal and a second order polynomial for the background yields $73 \pm 10 \pm 4$ B_s^0 candidates, where the first uncertainty is due to statistics and the second the systematic uncertainty which was estimated by varying the fit range, background and signal shape hypotheses. The mass resolution on the B_s^0 is

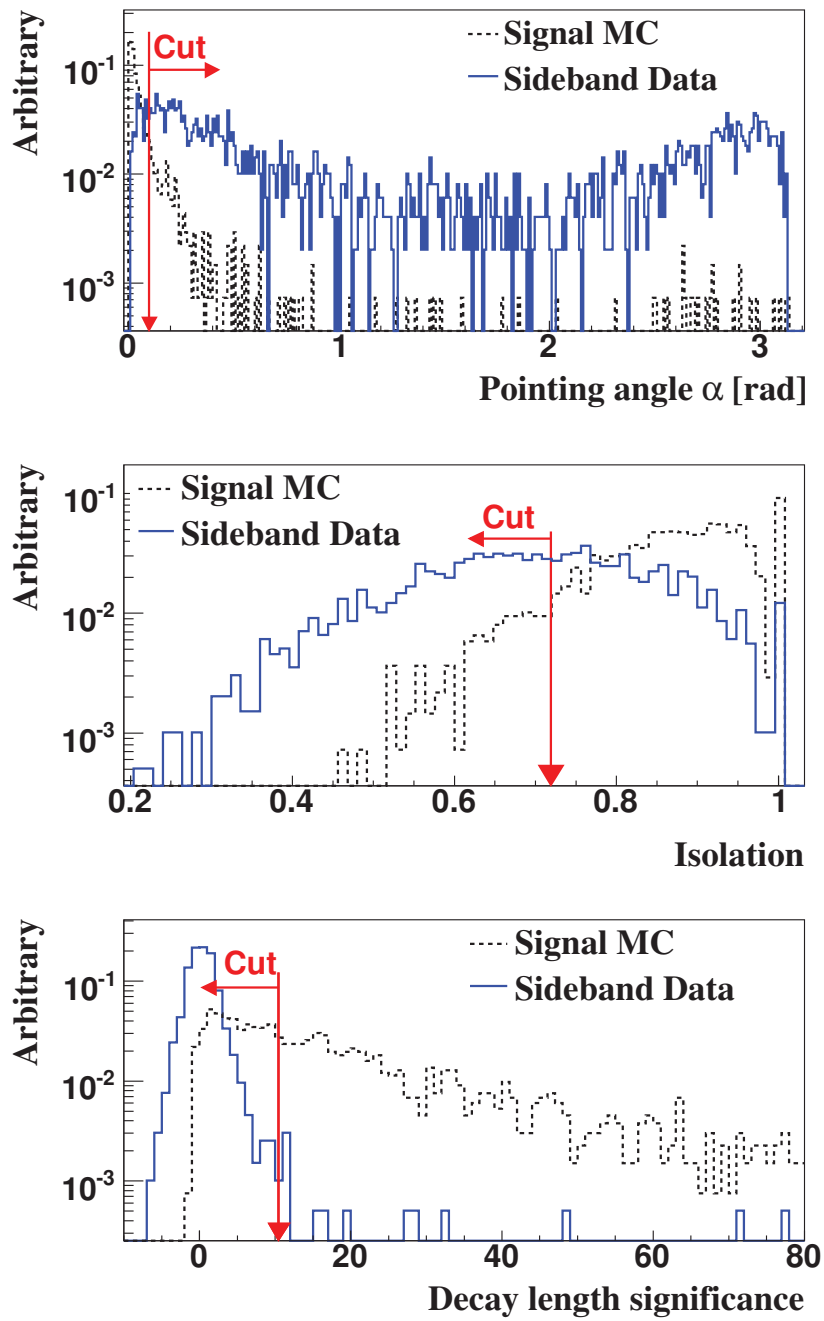

Figure 9.3: Discriminating variables after the pre-selection for signal Monte Carlo and data events from the sidebands for the $B_s^0 \to \phi\,\mu^+\mu^-$ analysis. The arrows indicate the cut values that were obtained after optimisation. The normalisation is done on the number of signal Monte Carlo and sideband data events after pre-selection.

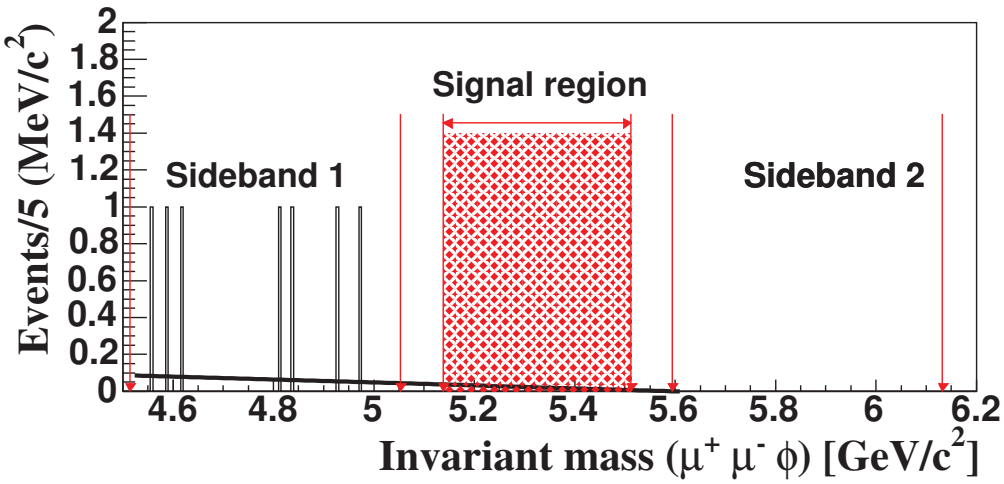

Figure 9.4: The remaining background for the full data sample with the standard discriminating variables.

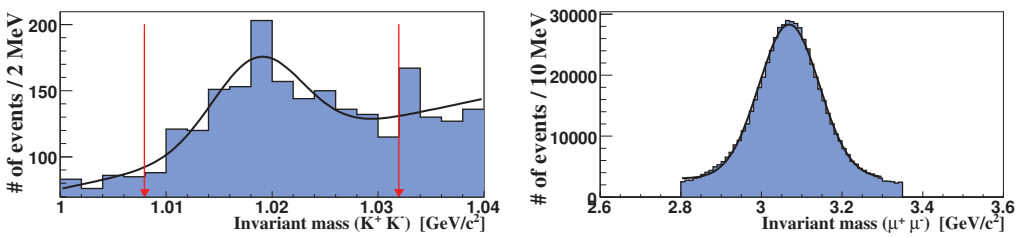

Figure 9.5: Invariant masses obtained from data for the ϕ candidates (left) and J/ψ candidates (right).

27.3 MeV, which compares very well to the resolution of 26.2 MeV obtained in the Monte Carlo simulation. The B_s^0 mass in the data from the $B_s^0 \to J/\psi\,\phi$ selection turns out to be 15 MeV/c^2 below the PDG mean value for the B_s^0. Using the four track invariant mass without constraining the di-muon to the J/ψ mass, a 44 MeV/c^2 shifted B_s^0 mass was obtained. This is exactly what is expected for the rare signal $B_s^0 \to \phi\,\mu^+\mu^-$. Therefore, all mass windows were shifted 44 MeV/c^2 to lower invariant masses with respect to the PDG mean value of the B_s^0 mass. Table 9.4 summarises the various resolutions obtained from data and Monte Carlo for the two different mass reconstructions fits (constraining and unconstraining the J/ψ mass). The $B_s^0 \to J/\psi\,\phi$ mass spectra are shown in Figure 9.8 for data and Monte Carlo events.

A comparison of data and Monte Carlo $B_s^0 \to J/\psi\,\phi$ events for various kinematical variables is shown in Appendix B.

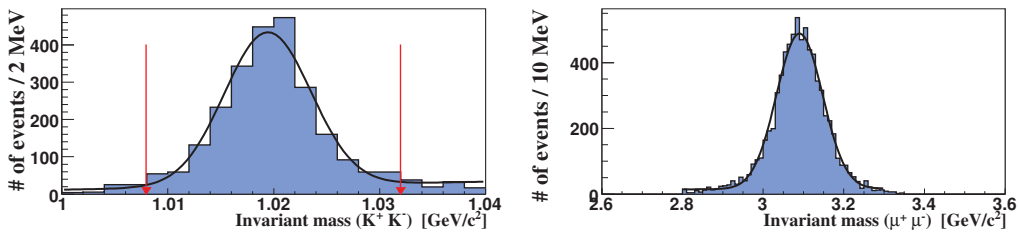

Figure 9.6: Invariant masses obtained from Monte Carlo for the ϕ candidates (left) and J/ψ candidates (right).

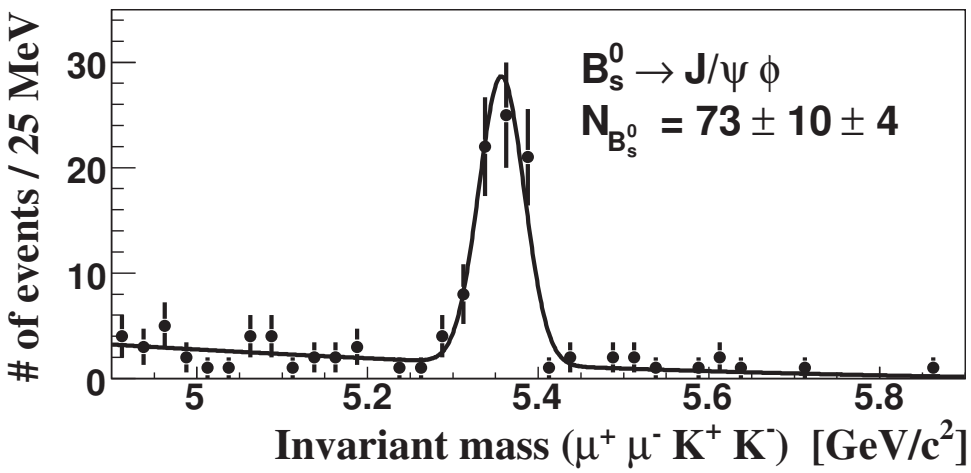

Figure 9.7: The normalisation channel $B_s^0 \to J/\psi\,\phi$ for the full data sample.

9.5 Deriving an Upper Limit

To calculate an upper limit on the branching ratio for the decay $B_s^0 \to \phi\,\mu^+\mu^-$, the number of reconstructed events $73 \pm 10 \pm 4$ of B_s^0 decaying into $J/\psi(\mu^+\mu^-)\,\phi$ were used as normalisation. Thus, $\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)$ was calculated as:

$$\frac{\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = \frac{\mu(n_{obs}, n_{back})}{N_{B_s^0 \to J/\psi\,\phi}} \cdot \frac{\epsilon^{B_s^0}_{J/\psi\phi}}{\epsilon^{B_s^0}_{\mu^+\mu^-\phi}} \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-) \tag{9.1}$$

where

- $\mu(n_{obs}, n_{back})$ is the upper limit on the number of observed events;

	Data	Monte Carlo
no J/ψ mass constraint		
Mean MeV/c^2	5321.98	5365.55
σ MeV/c^2	78.66	73.34
J/ψ mass constraint		
Mean MeV/c^2	5354.95	5368.74
σ MeV/c^2	27.32	26.23

Table 9.4: Data and Monte Carlo resolution comparison for the $B_s^0 \to J/\psi\,\phi$ candidates.

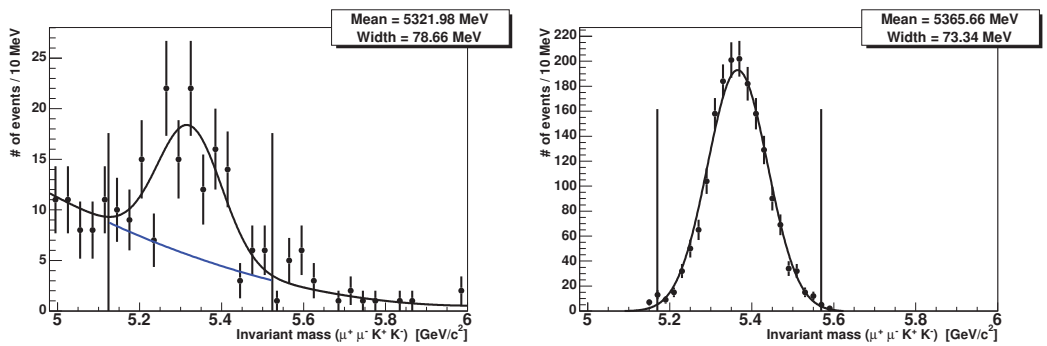

Figure 9.8: Mass spectra for Data (left) and Monte Carlo (right) of the unconstraint $B_s^0 \to J/\psi\,\phi$ candidates.

- $\epsilon^{B_s^0}_{\phi\mu^+\mu^-}$ and $\epsilon^{B_s^0}_{J/\psi\phi}$ are the efficiencies of the signal and normalisation channels, obtained from Monte Carlo simulations and

- the branching ratio $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.1)\%$ was taken from [4].

The efficiencies $\epsilon^{B_s^0}_{\phi\mu^+\mu^-}$ and $\epsilon^{B_s^0}_{J/\psi\phi}$ are the global signal efficiencies for the search signal and normalisation channel, respectively, including pre-selection cuts, the acceptance and trigger efficiency. They were determined from Monte Carlo to be $\epsilon^{B_s^0}_{\phi\mu^+\mu^-} = (3.6 \pm 0.2) \times 10^{-5}$ and $\epsilon^{B_s^0}_{J/\psi\phi} = (9.9 \pm 0.6) \times 10^{-5}$, where the uncertainties are due to Monte Carlo statistics. The quoted efficiencies refer to triggered Monte Carlo events using the trigger simulation only. Finally, the Monte Carlo events were weighted according to the data/MC difference for the $p_T(B)$ distribution and a final efficiency ratio of 2.80 ± 0.21 was determined. This efficiency ratio was used to calculate the limit and an additional uncertainty on the efficiency ratio was assigned.

9.6 Sensitivity

The analysis used the technique of a "blind box" analysis, which means that the signal region was blinded throughout all analysis steps presented so far. Before exploring the "blinded" invariant di-muon signal mass region one can calculate an "expected upper limit", as introduced in Section 7.7, without knowing the number of signal $B_s^0 \to \phi\,\mu^+\mu^-$ events. For this, the upper limit $\mu(n_{obs}, n_{back})$ in Eq. 9.1 is replaced by $\langle\mu(n_{back})\rangle$. Using the numbers presented in Section 9.5, without their statistical uncertainties, the result for the statistical sensitivity is

$$\frac{\langle\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)\rangle}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 1.0\ \times 10^{-2} \tag{9.2}$$

at a 95% CL.

Using the central value of the world average branching ratio [4] of $B_s^0 \to J/\psi\,\phi = 9.3\ \times 10^{-4}$, a sensitivity for the rare decay of

$$\langle\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)\rangle = 9.2\ \times 10^{-6}$$

at a 95% CL was obtained.

9.7 Systematic Uncertainties

The different sources of relative uncertainties that enter the calculation on the limit of the branching ratio are listed in Table 9.5. In principle, the branching ratio for $B_s^0 \to J/\psi\,\phi$ has the largest uncertainty, but its uncertainty cancels out due to the normalization.

The second largest uncertainty is 25% from the background interpolation into the signal region and is due to the statistical uncertainty on the integral of the fit. The total uncertainty on the number of observed $B_s^0 \to J/\psi\,\phi$ events is 14.8%.

The efficiency ratio $\epsilon^{B_s^0}_{J/\psi\phi}/\epsilon^{B_s^0}_{\phi\mu^+\mu^-}$ was determined in this analysis as a single number from Monte Carlo and hence correlations are taken into account correctely. The statistical uncertainty on the ratio is 7.4%.

An additional uncertainty of 1.1% is applied to take into account Monte Carlo weighting. This weighting is due to the different p_T distributions of the B mesons in data and Monte Carlo. The same weighting procedure as used in the search for the decay $B_s^0 \to \mu^+\mu^-$ and described in Section 8.4 was used.

Within the framework of the Standard Model, the B_s^0 mesons are expected to mix in such a way that the mass and decay width differences between the heavy and light

Table 9.5: The relative uncertainties for calculating an upper limit of $\mathcal{B}$. Note that the first uncertainty is not taken into account since the limit has been normalised to $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$.

Source	Relative Uncertainty [%]
$\mathcal{B}(B_s^0 \to J/\psi\,\phi)$	35.5
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	1.7
$\epsilon_{J/\psi\phi}/\epsilon_{\phi\mu^+\mu^-}$	7.4
# of $B_s^0 \to J/\psi\,\phi$	14.8
Monte Carlo weighting	1.1
CP-odd/even lifetime differences	8.0
pre-Geant weighting	3.5
Total	18.8
background uncertainty	25.0

eigenstates, $\Delta M \equiv M_H - M_L$ and $\Delta\Gamma \equiv \Gamma_H - \Gamma_L$ are sizeable. The mixing phase $\delta\phi$ is small and to a good approximation the two mass eigenstates correspond to the two CP eigenstates. The light mass eigenstate is expected to be CP-even and has a larger decay width, and thus a shorter lifetime, than the heavy mass eigenstate that is expected to be CP-odd. The Monte Carlo samples for the normalization channel contained only CP-even states. The efficiency difference between the CP-even and CP-odd state for the normalization channel was estimated using Monte Carlo to be 6%, with an uncertainty of 8% on the ratio. Therefore, another uncertainty of 8% was conservatively assigned to take into account contributions from the CP-odd/even eigenstates.

An additional uncertainty was introduced to take into account pre-Geant weighting. As described in Section 5.4, kinematical cuts for particles at generator level were applied in order to save computing time. One has to check for possible biases introduced at the pre-Geant (detector simulation) stage. To study those effects, Pythia samples for each relevant decay channel were produced without any cuts on kinematic variables at the event generator level. After that, each B meson candidate was weighted according to the obtained weighting function and an efficiency was calculated. The efficiency ratio with and without kinematic cuts for weighted events was estimated. A difference of 3.5% was observed and taken as a systematic uncertainty.

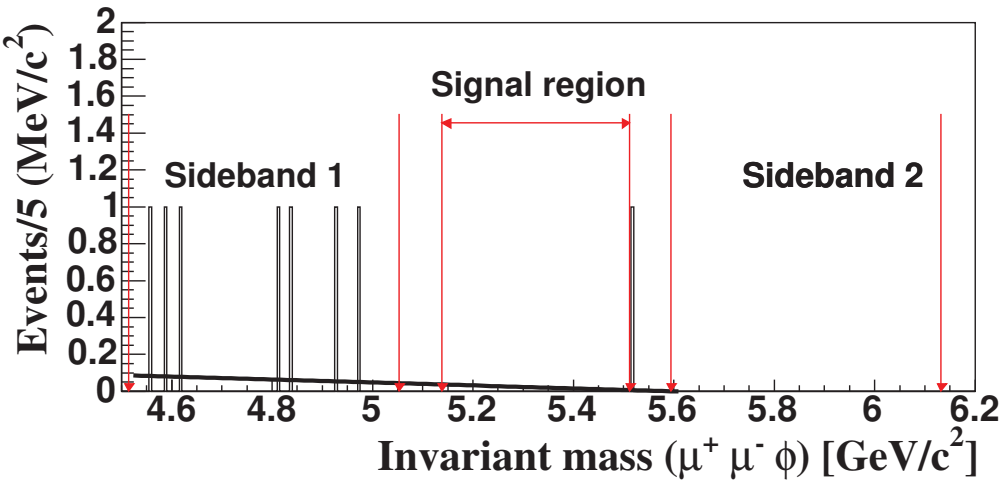

Figure 9.9: The invariant mass distribution for the full data sample with the standard discriminating variables.

9.8 Results

After review and approval by the DØ collaboration, and after finalising the cuts, the signal box was opened. No signal event was found for the cuts used in this analysis. The invariant mass distribution is shown in Figure 9.9. The Poisson probability for observing zero events for an expected background of 1.6 ± 0.4 is $p = 0.22$. Since fewer events than expected were observed, the background prediction and the number of observed events were checked with different combinations of looser cuts. Table 9.6 gives an overview of some of the tested combinations and shows a reasonable agreement between the expected background and the observed numbers of events.

Upper limits depend strongly on the algorithm that is used to derive the limit in case that fewer events are observed than expected. Therefore, two limits, derived using the unified and the Bayesian approach, are quoted here. Including the statistical and systematic uncertainties, the limit using the Frequentist approach is:

$$\frac{\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} < 4.4\,(3.5)\ \times 10^{-3} \tag{9.3}$$

at a 95% (90%) CL. Using the central value of the world average for the branching ratio [4] of $B_s^0 \to J/\psi\,\phi = 9.3 \pm 3.3\ \times 10^{-4}$, this corresponds to a limit of

$$\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-) < 4.1\,(3.2)\ \times 10^{-6} \tag{9.4}$$

at a 95% (90%) CL.

Table 9.6: Comparison of the expected background and observed events after various cuts.

cut parameter	cut values	expected	observed
no discriminating cuts		534 ± 14	517
Pointing angle (rad)	< 0.11	37 ± 4	39
Pointing angle (rad) &	< 0.11		
Decay length significance	> 4	7.8 ± 1.8	3
Pointing angle (rad) &	< 0.11		
Decay length significance	> 2	18.6 ± 2	19
Pointing angle (rad) &	< 0.11		
Decay length significance	> 11.1	3.4 ± 0.8	0
Pointing angle (rad) &	< 0.11		
Decay length significance &	> 11.1		
Isolation	> 0.72	1.6 ± 0.4	0

Taking a Bayesian approach ([92]) with flat prior, and uncertainties treated as Gaussians in the integration, an upper limit at a 95% (90%) CL of

$$\frac{\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} < 7.4\,(5.6)\ \times 10^{-3} \tag{9.5}$$

was found.

A good statistical practice is also to quote the sensitivity of the search, if fewer events are observed than expected. Here, the sensitivity was determined as the average expected upper limit from an ensemble of hypothetical experiments including the statistical and systematic uncertainties it was found to be

$$\frac{\langle\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-)\rangle}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 1.1\ \times 10^{-2} \tag{9.6}$$

at a 95% CL.

Chapter 10

Observation of the Decay $B_s^0 \to \psi(2S)\,\phi$

In this Chapter, the observation of the decay $B_s^0 \to \psi(2S)\,\phi$ and a measurement of the branching ratio $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)$ relative to $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ is reported. In addition, the branching ratio of the decay $B^\pm \to \psi(2S)\,K^\pm$ relative to $B^\pm \to J/\psi\,K^\pm$ has been measured and used as control channel.

10.1 Introduction

Studies of the decays of B mesons to $\psi(2S)$ final states have contributed to knowledge about hadronic B-meson decays and provide insight into the interplay between weak and strong interactions. The decay $B^\pm \to \psi(2S)\,K^\pm$ was first observed [96] at ARGUS, $B^0 \to \psi(2S)\,K^{*0}$ was frist observed [97] at CDF (Run I) , and CLEO frist observed [98] $B^0 \to \psi(2S)\,K_S$ and $B^\pm \to \psi(2S)\,K^{\star\pm}$. Subsequently, all these decay modes have been studied with more statistics by many experiments. The measurements show that the rates of $B^\pm$ and B^0 meson decays to the $\psi(2S)$ final states are approximately 60% of the rates of the corresponding decays to the J/ψ final state. For the B_s^0 meson, up to now only the decay $B_s^0 \to \psi(2S)\,\phi$ has been observed by the ALEPH collaboration, which reported [99] one candidate event in 1993 in their B_s^0 mass measurement analysis. A branching ratio for $B_s^0 \to \psi(2S)\,\phi$ relative to $B_s^0 \to J/\psi\,\phi$ has not yet been measured. Main properties of the J/ψ and $\psi(2S)$ are summarised in Tab. 10.1[4].

10.2 Event Selection

The $B_s^0 \to \psi(2S)\,\phi$ candidate event signature contains two muons from the $\psi(2S)$ and two kaon candidate tracks which form a ϕ candidate. This signature is the same for the normalisation channel $B_s^0 \to J/\psi\,\phi$.

Table 10.1: Comparison of properties of J/ψ and $\psi(2S)$ mesons (taken from [4]).

	J/ψ	$\psi(2S)$
Mass [MeV/c^2]	3096.916 $\pm$ 0.011	3686.093 $\pm$ 0.034
Full width Γ [keV]	91.0 $\pm$ 3.2	281 $\pm$ 17
J^{PC}	1^{--}	1^{--}
Decay to hadrons [%]	87.7 $\pm$ 0.5	97.85 $\pm$ 0.13
Decay to e^+e^- [%]	5.93 $\pm$ 0.10	0.755 $\pm$0.031
Decay to $\mu^+\mu^-$ [%]	5.88 $\pm$ 0.10	0.73 $\pm$ 0.08
Decay to $\tau^+\tau^-$ [%]	-	0.28 $\pm$ 0.07

10.2.1 Pre-selection

The pre-selection started with a loose selection of $B^\pm$ (B_s^0) candidates, consisting of two identified muons and one (two oppositely) track(s) of charged particle(s) forming a good vertex. In the selection step for the B_s^0, the invariant mass m_ϕ of the two kaon candidate tracks should be between 0.980 GeV/c^2 $< m_\phi <$ 1.080 GeV/c^2. The invariant mass of the loose B candidate was required to be 4.4 GeV/c^2 $< m_B <$ 6.2 GeV/c^2. For the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\,\phi$ candidates, the invariant masses of the muon pair was required to be within 250 MeV/c^2 of the J/ψ and $\psi(2S)$ masses [4], respectively.

The χ^2/d.o.f. of the two-muon vertex was required to be smaller than 16. The transverse momentum of each of the muons was required to be greater than 2.0 GeV/c and their pseudorapidity had to be $|\eta| < 2.0$ to be well inside the fiducial volume of tracking and muon detectors. The tracks that were matched to each muon leg had to have at least one hit in the SMT and one hit in the CFT.

For events that passed these pre-selection cuts, the two-dimensional decay length L_{xy} of the B candidate in the plane transverse to the beamline was calculated. The error on this transverse decay length δL_{xy} was required to be smaller than 150 μm. The transverse momentum of the $\mu^+\mu^-$ pair had to be greater than 4 GeV/c.

The selection of the decays $B^\pm \to J/\psi\,K^\pm$ and $B_s^0 \to J/\psi\,\phi$ described here differs slightly from the reconstruction described in the searches for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B_s^0 \to \phi\,\mu^+\mu^-$ (Section 8.3 and 9.4). The selection criteria were modified in order to adapt for the searched signals $B^\pm \to \psi(2S)\,K^\pm$ and $B_s^0 \to \psi(2S)\,\phi$ and are described in the following.

10.3 Reconstruction of the Decays $B^{\pm} \to J/\psi\, K^{\pm}$ and $B^{\pm} \to \psi(2S)\, K^{\pm}$

To reconstruct the decays $B^{\pm} \to (J/\psi, \psi(2S))\, K^{\pm}$ with $(J/\psi, \psi(2S)) \to \mu^+\mu^-$, the two muons were in addition required to have at least the medium quality criteria fulfilled. The formed candidates were then constrained to the invariant mass of the J/ψ or $\psi(2S)$ [4], respectively. The combined vertex fit of the J/ψ and the additional $K^{\pm}$ should not yield a χ^2 of more than 20 for 3 d.o.f. The p_T of the $K^{\pm}$ had to be larger than 0.9 GeV/c. Moreover, a collinearity between the decay length vector of the $B^{\pm}$ and the combined momentum of J/ψ and $K^{\pm}$ of at least 0.9 in the transverse plane was required. To remove prompt background, the decay length significance $L_{xy}/\delta L_{xy}$ of the $B^{\pm}$ candidate had to be larger than four. The results of a fit to the obtained invariant mass spectrum, using a Gaussian and a second order polynomial, are summarised in Table 10.2 for Monte Carlo and in Table 10.3 for the full data sample. The signal event yields for the $B^{\pm} \to J/\psi\, K^{\pm}$ decay were obtained by leaving all fit parameters floating. For the $B^{\pm} \to \psi(2S)\, K^{\pm}$ decay however, the values for the mean and the width of the Gaussian were fixed. The mean was set to the fit result obtained for the $B^{\pm} \to J/\psi\, K^{\pm}$ data events. The width was obtained by scaling the width obtained for the Monte Carlo sample with a scale factor obtained from the data/MC difference[1] for $B^{\pm} \to \psi(2S)\, K^{\pm}$ events. The mean for both channels ($B^{\pm} \to J/\psi\, K^{\pm}$, $B^{\pm} \to \psi(2S)\, K^{\pm}$) are lower than expected but the widths are consistent with expectations from Monte Carlo. The change in the yield that is expected for the case that all values are floating and not fixed to the Monte Carlo expectation was taken as a systematic uncertainty, as discussed in Section 10.7.2. In Figure 10.1, the resulting invariant mass distribution for the $B^{\pm} \to J/\psi\, K^{\pm}$ is shown and in Figure 10.2 the corresponding distribution is shown for $B^{\pm} \to \psi(2S)\, K^{\pm}$.

Table 10.2: Summary of fitting results of $B^{\pm} \to (J/\psi, \psi(2S))\, K^{\pm}$ final states for Monte Carlo events.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	ϵ_{MC}
$B^{\pm} \to J/\psi K^{+}$	5278.8 $\pm$ 0.8	37.6 $\pm$ 0.8	(1.14$\pm$0.02) $\times 10^{-3}$
$B^{\pm} \to \psi(2S) K^{+}$	5278.3 $\pm$ 1.0	29.0 $\pm$ 0.6	(1.07$\pm$0.04) $\times 10^{-3}$

[1] $\sigma^{Data}_{B^{\pm} \to \psi(2S)\, K^{\pm}} = \frac{\sigma^{Data}_{B^{\pm} \to J/\psi\, K^{\pm}}}{\sigma^{MC}_{B^{\pm} \to \psi(2S)\, K^{\pm}}} \cdot \sigma^{MC}_{B^{\pm} \to \psi(2S)\, K^{\pm}}$

Table 10.3: Summary of fitting results of $B^\pm \to (J/\psi, \psi(2S))\,K^\pm$ final states for the data.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	Yield
$B^\pm \to J/\psi\,K^\pm$	5273.7 $\pm$ 1.2	40.1 $\pm$ 1.2	1970$\pm$62
$B^\pm \to \psi(2S)\,K^\pm$	5273.7 (fixed)	30.9 (fixed)	157 $\pm$18
$B^\pm \to \psi(2S)\,K^\pm$	5269.9 $\pm$ 3.2	27.1 $\pm$ 2.9	149$\pm$18

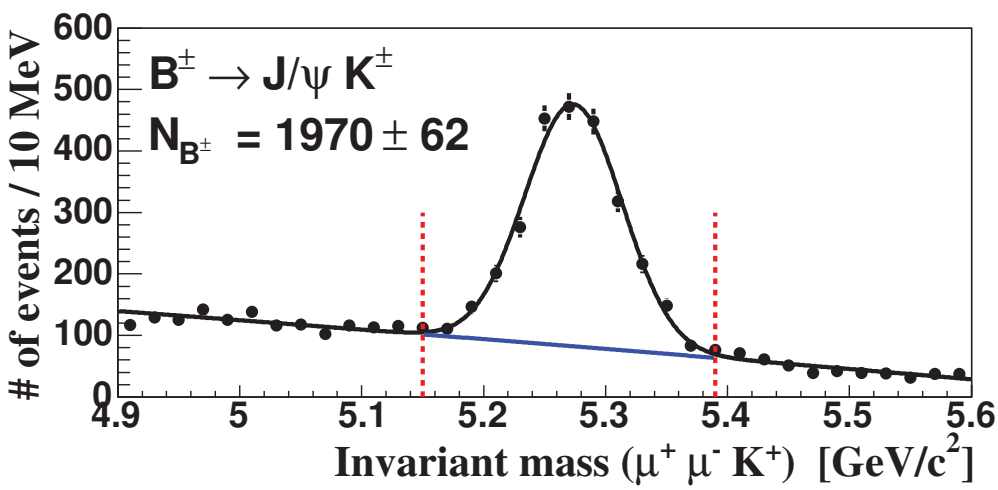

Figure 10.1: Invariant mass distribution for the $B^\pm \to J/\psi\,K^\pm$ decay for the data sample.

10.4 Reconstruction of the Decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\,\phi$

To reconstruct the decays $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ with $(J/\psi, \psi(2S)) \to \mu^+\mu^-$), the two muons were in addition required to have at least a hit in the first layer of the muon system and to be matched to a central track. They were constrained to the invariant mass of the J/ψ or $\psi(2S)$ [4]. The muon pair was then combined with another pair of tracks of oppositely charged particles (ϕ candidate), each with $p_T > 0.9$ GeV/c, to a B_s^0 candidate vertex with a $\chi^2 < 36$ for five d.o.f. Each of the Kaon candidates needed at least one hit in the SMT and the ϕ candidate was required to have an invariant mass between 1.008 GeV/c^2 and 1.032 GeV/c^2. To remove prompt background, the decay length significance $L_{xy}/\delta L_{xy}$ of the B_s^0 candidate had to be larger than four. The results of a fit to

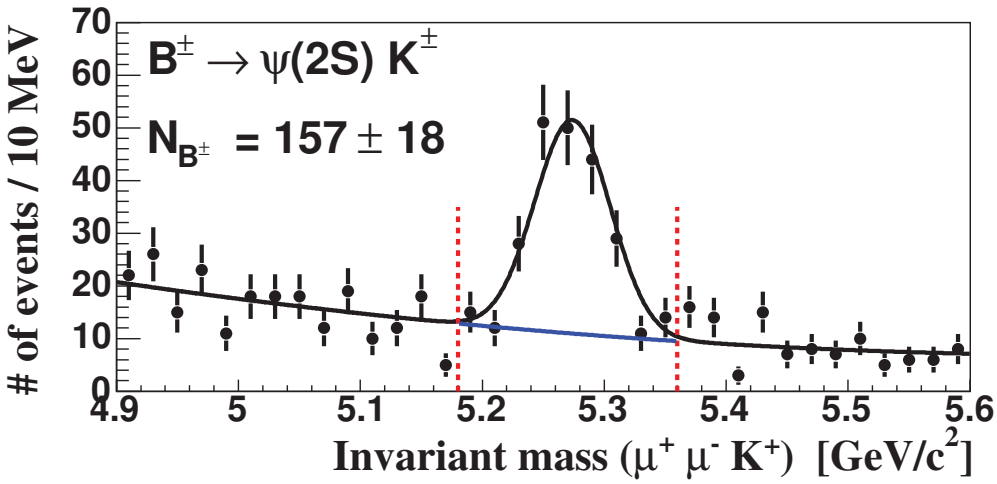

Figure 10.2: Invariant mass distribution for the $B^{\pm} \to \psi(2S)\, K^{\pm}$ decay for the data sample.

the invariant mass spectrum, using a Gaussian and a second order polynomial, are summarised in Table 10.4 for Monte Carlo and in Table 10.5 for the full data sample. The yields for the decay $B_s^0 \to J/\psi\, \phi$ were obtained by leaving all parameters floating. Similar as in the case of $B^{\pm} \to \psi(2S)\, K^{\pm}$, the values for the mean and width of the Gaussian were fixed for the $B_s^0 \to \psi(2S)\, \phi$ decay. The mean was set to the value obtained for $B_s^0 \to J/\psi\, \phi$ and the width was obtained by scaling the width from Monte Carlo with a factor obtained from the data/MC difference[2] for the $B_s^0 \to J/\psi\, \phi$ decay.

In Figure 10.3, the resulting invariant mass distribution for the $B_s^0 \to J/\psi\, \phi$ is shown and in Figure 10.4 the corresponding distribution is shown for $B_s^0 \to \psi(2S)\, \phi$.

To further enhance the signal to background ratio, discriminating variables described in Section 6.1 and the optimisation procedure described in Section 6.2 were used.

10.5 Optimisation Procedure

Before the optimisation of the discriminating variables, the mass region of interest was restricted to 4.8 GeV/c^2 $< M_{\phi\mu^+\mu^-} <$ 6.0 GeV/c^2, containing the signal region around the world average value of the B_s^0 mass. Table 10.6 defines the mass regions that were used for the sidebands and the signal region.

[2] $\sigma^{Data}_{B_s^0 \to \psi(2S)\,\phi} = \frac{\sigma^{Data}_{B_s^0 \to J/\psi\,\phi}}{\sigma^{MC}_{B_s^0 \to J/\psi\,\phi}} \cdot \sigma^{MC}_{B_s^0 \to \psi(2S)\,\phi}$ and using the CP-even Monte Carlo events for $B_s^0 \to J/\psi\, \phi$.

Table 10.4: Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for Monte Carlo events.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	ϵ_{MC}
$B_s^0 \to J/\psi\,\phi$ (CP-even)	5369.6 $\pm$ 1.3	26.6 $\pm$ 1.2	(1.94$\pm$0.08) $\times 10^{-4}$
$B_s^0 \to J/\psi\,\phi$ (CP-odd)	5370.2 $\pm$ 1.3	28.0 $\pm$ 1.2	(2.05$\pm$0.09) $\times 10^{-4}$
$B_s^0 \to \psi(2S)\,\phi$ (CP-even)	5370.1 $\pm$ 0.8	22.5 $\pm$ 0.7	(2.08$\pm$0.07) $\times 10^{-4}$

Table 10.5: Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for the data.

Decay	Mean [MeV/c^2]	Width [MeV/c^2]	Yield
$B_s^0 \to J/\psi\,\phi$	5356.7 $\pm$ 2.7	28.9 $\pm$ 2.3	200$\pm$18
$B_s^0 \to \psi(2S)\,\phi$	5356.7 (fixed)	24.4 (fixed)	13$\pm$8

The signal region corresponds to a window of ± 150 MeV/c^2 around the B_s^0 mass. The expected mass resolution for $B_s^0 \to \psi(2S)\,\phi$ from Monte Carlo simulation is $\approx$ 25 MeV/c^2, the chosen mass window therefore covers a $\pm 6\sigma$ window. During optimisation the signal region was hidden.

Table 10.6: The different four track invariant mass regions for signal and sidebands used for background estimation.

Region	min Mass (GeV/c^2)	max Mass (GeV/c^2)
region of interest	4.80	6.00
hidden signal region during optimisation	5.22	5.52
sideband I	4.80	5.22
sideband II	5.52	6.00

Two different methods were used for the optimisation of the three discriminating variables. The first of those methods is the standard cut-based Random Grid Search and the second method employs a linear combination of the three variables to a so-called “linear classifier”. The results of these two methods are described in the following.

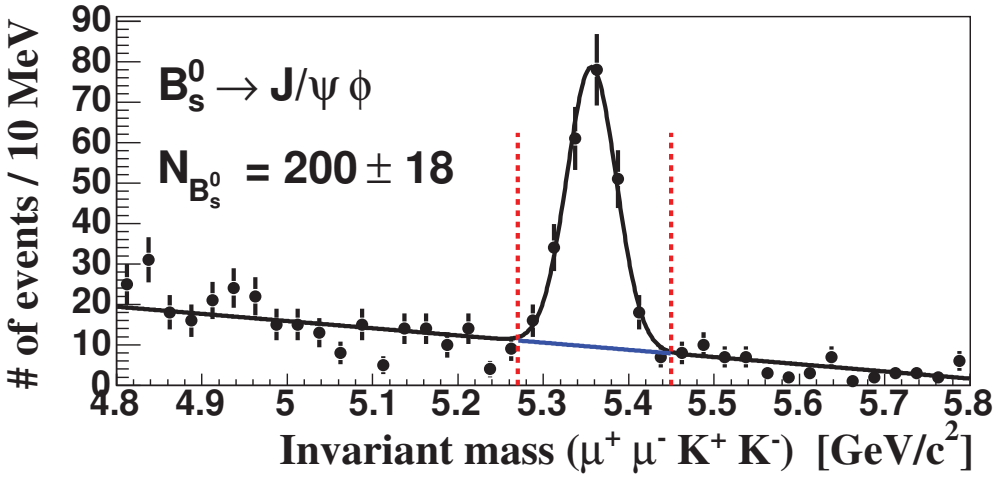

Figure 10.3: Invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ decay after the pre-selection for the data sample using a loose decay length significance cut of greater than four.

10.5.1 Random Grid Search

A Random Grid Search, as described in Section. 6.2, was used in order to find the optimal set of cuts. As optimisation criterion, the variable P defined as:

$$P = \frac{\epsilon_{\psi(2S)\,\phi}}{\frac{a}{2} + \sqrt{N_{Back}}} \tag{10.1}$$

was used. Instead of the definition given in Section 6.2, the constant a was set to a value of five here. The constant a is the number of standard deviations that corresponds to the confidence level at which the signal hypothesis is tested. Therefore, the value of five corresponds to a maximised sensitivity to observe a 5σ signal. The reconstruction efficiency from signal Monte Carlo after pre-selection is given by $\epsilon_{\psi(2S)\,\phi}$, and N_{Back} is the expected number of background events interpolated from the sidebands as defined in Table 10.6. The resulting cut values that were obtained from maximising P are listed in Table 10.7 as well as a comparison of the signal efficiency with the Monte Carlo efficiency. As expected, they are very similar to the $B_s^0 \to \phi\,\mu^+\mu^-$ analysis due to the same final states and almost identical background. Table 10.7 and Table 10.8 compare the efficiencies for Monte Carlo and signal yield for the discriminating cuts. The values show a reasonably good agreement between data and Monte Carlo simulation.

The total signal efficiency for the decay $B_s^0 \to \psi(2S)\,\phi$ for the three discriminating variables relative to pre-selection was found to be $(60 \pm 5)\%$. After a linear interpolation of the sideband population into the final signal region, an expected number of background

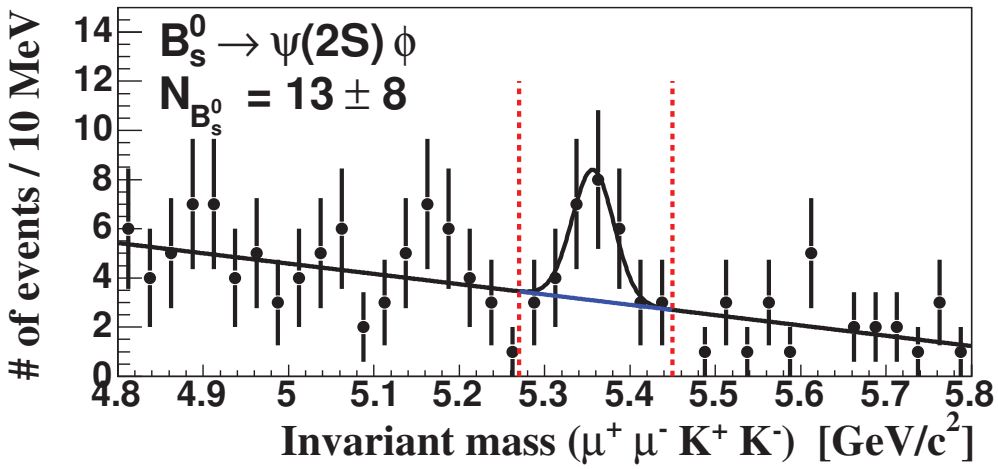

Figure 10.4: Invariant mass distribution for the $B_s^0 \to \psi(2S)\,\phi$ decay for the data sample using a loose decay length significance cut of greater than four.

Table 10.7: Comparison of the signal efficiency with the Monte Carlo efficiency for $B_s^0 \to \psi(2S)\,\phi$ events.

cut parameter	cut value	MC eff. (%)	Data eff. (%)
Pointing angle (rad)	< 0.11	88.6 ± 4.5	77±33
Decay length significance	> 11.1	74.9 ± 4.3	90±41
Isolation	> 0.74	90.9 ± 5.6	89±29

events of 0.8±0.4 was obtained for the whole data sample. Figure 10.5 shows the resulting invariant mass distribution after all cuts. In the signal region, 11 events were found. The Poisson probability that the expected background fluctuates to give the observed 11 events or more is given by $p = 1.04\ \times 10^{-9}$. The significance S of this enhancement, quoted in number of standard deviations, is 5.99σ. The significance S is defined as

$$1 - p = \int_{-\infty}^{S} N(0,1)\,dx \tag{10.2}$$

with $N(0,1)$ denoting a standardised Gaussian probability function with expectation value and variance equal to 0 and 1, respectively.

A fit to the obtained mass spectrum assuming a Gaussian signal over a linear background, yielded 8.6±3.3 signal events with 1.8±1.3 background events. The probability

Table 10.8: Comparison of the signal efficiency with the Monte Carlo simulation efficiency for $B_s^0 \to J/\psi\,\phi$ events.

cut parameter	cut value	MC eff. (%)	Data eff. (%)
Pointing angle (rad)	< 0.11	89.5±5.8	83.0±8.8
Decay length significance	> 11.1	80.8±5.7	75.9±8.9
Isolation	> 0.74	90.2±6.8	87.3±11.4

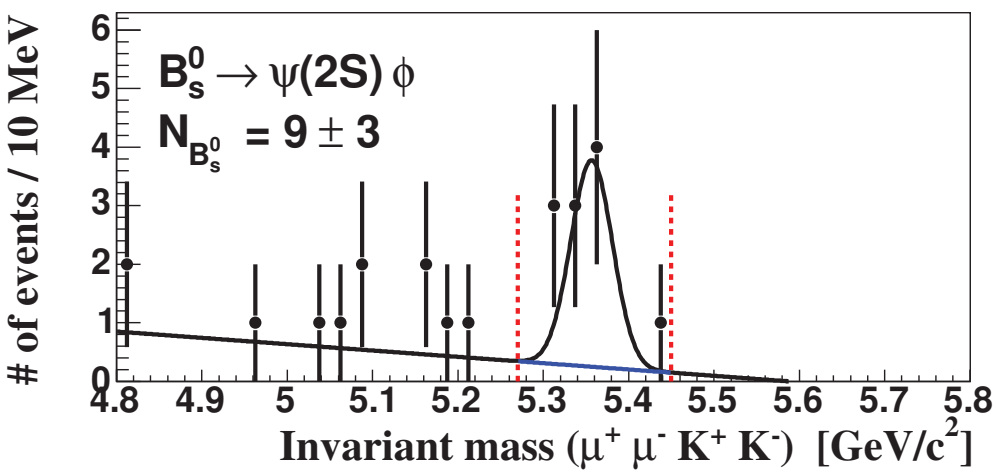

Figure 10.5: Invariant mass distribution for the $B_s^0 \to \psi(2S)\,\phi$ decay for the data sample and the optimised discriminating cuts with the RGS.

for a background fluctuation is $p = 1.1 \times 10^{-4}$. The significance S of the fitted signal peak can also be evaluated as $S = \sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{max})}$ = 3.89, with $\mathcal{L}_{max}$ being the Maximum-Likelihood of the best fit including both signal and background and $\mathcal{L}_0$ being the best fit if the signal yield is set to zero. For the log-likelihood fit the mean and width of the Gaussian were fixed as described in Section 10.4.

10.5.2 A Multivariate Linear Classifier

A logistic regression model as described in Section 6.4 was employed in a multivariate approach. As multivariate input, the three variables isolation, decay length significance and pointing angle were used. Figure 10.6 shows the distribution of the resulting discriminating variable ξ for signal Monte Carlo events and data. The optimal cut value on ξ

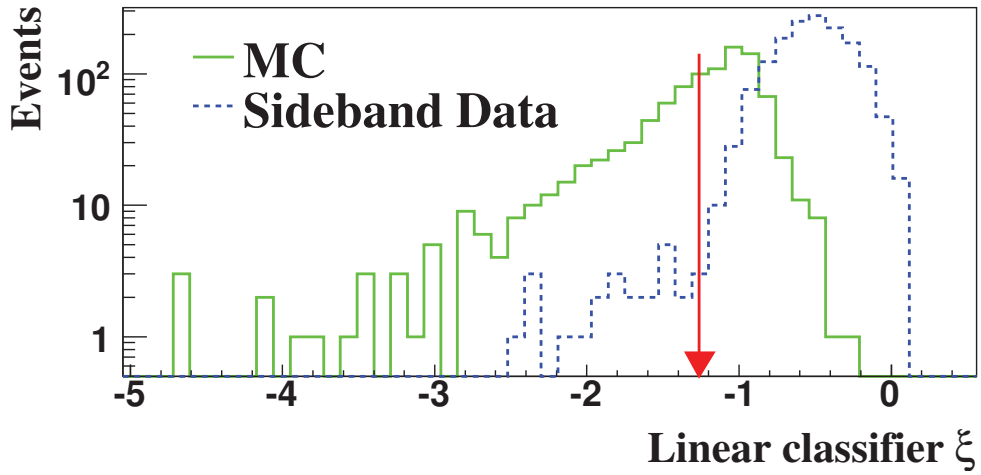

Figure 10.6: Output of the logistic regression for signal Monte Carlo and sideband data.

was again found using the variable P as maximisation criterion and is indicated in Figure 10.6. The optimised linear classifier model yielded a background in data of 0.6 ± 0.2 events after interpolation, while 10 candidate events were observed. A Gaussian fit as shown in Figure 10.7 yielded 9.2±3.3 signal events and 0.8±0.8 background events. The signal efficiencies and background rejection values are comparable to the ones obtained with the RGS. For reasons of transparency, however, it was decided not to use the linear classifier, but the RGS approach.

10.6 Relative Branching Ratio Measurements

10.6.1 Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

Since a significant excess of signal events was observed, a relative branching ratio was determined using the decay $B_s^0 \to \psi(2S)\,\phi$. The ratio of the branching fractions is[3]:

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = \frac{N_{B_s^0 \to \psi(2S)\,\phi}}{N_{B_s^0 \to J/\psi\,\phi}} \cdot \frac{\epsilon_{J/\psi\,\phi}}{\epsilon_{\psi(2S)\,\phi}} \cdot \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi(2S) \to \mu^+\mu^-)} \tag{10.3}$$

where

- $\epsilon_{\psi(2S)\,\phi}$ and $\epsilon_{J/\psi\,\phi}$ are the efficiencies of the signal and normalisation channels, obtained from Monte Carlo simulations, and

[3]The same formula is valid to calculate the ratio for the $B^\pm$ decay, with ϕ interchanged with $K^\pm$

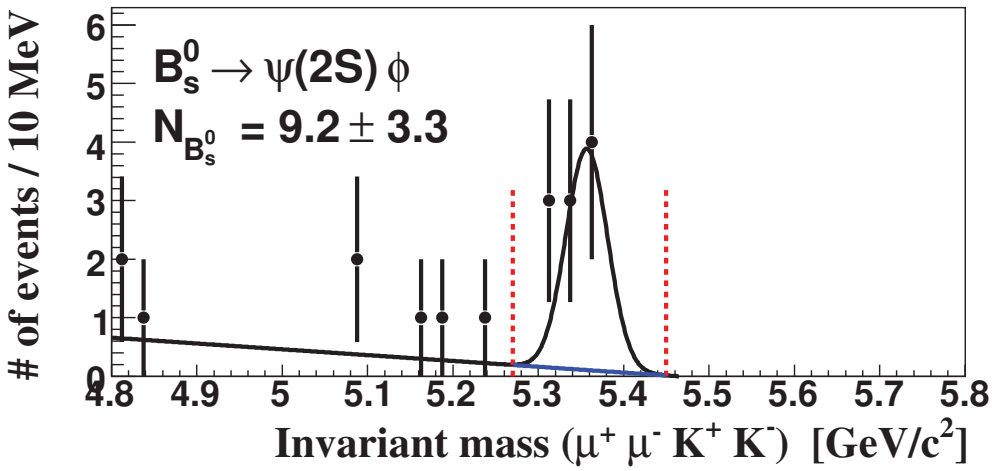

Figure 10.7: The invariant mass distribution for the full data sample using a linear discriminant.

- $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.1)\%$ and $\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = (7.3 \pm 0.8) \times 10^{-3}$ [4] are the measured branching fractions.

The efficiencies $\epsilon_{\psi(2S)\,\phi}$ and $\epsilon_{J/\psi\,\phi}$ are the global signal efficiencies for the signal and normalisation channel, respectively, including the pre-selection cuts and the acceptance. In Table 10.9, the various Monte Carlo efficiencies are given that were needed to calculate the ratio of the branching ratios. The quoted values refer to triggered Monte Carlo events in the trigger simulator and the uncertainties are due to statistics only.

Table 10.9: Efficiencies for the two decay channels (both channels CP-even only) after trigger and reconstruction with respect to the generated $b\bar{b}$-pairs.

	$\epsilon_{\psi(2S)\,\phi}$	$\epsilon_{J/\psi\,\phi}$	$\epsilon_{J/\psi\,\phi}/\epsilon_{\psi(2S)\,\phi}$
RGS (CP-even)	$(12.5 \pm 0.6) \times 10^{-5}$	$(11.5 \pm 0.7) \times 10^{-5}$	$(91.7 \pm 7.2)\%$

The invariant mass plot for $B_s^0 \to J/\psi\,\phi$ events, after applying the same cuts on the discriminating variables as quoted in Table 10.8 for the $B_s^0 \to \psi(2S)\,\phi$ events, is shown in Figure 10.8. The observed number of $B_s^0 \to J/\psi\,\phi$ signal events obtained from a gaussian fit is 110 ± 11 events, including statistical uncertainty only.

This gives for the RGS-optimization

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24.\ \text{(stat)}$$

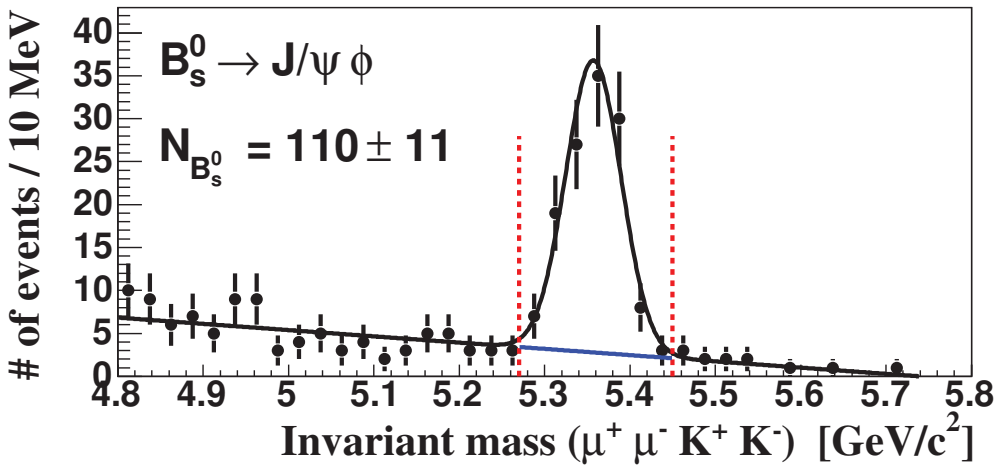

Figure 10.8: Invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ decay for the data sample and the optimized discriminating cuts with the RGS.

Using the signal event yields quoted in Table 10.5 (not optimised discriminating cuts) one obtains:

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.48 \pm 0.30 \text{ (stat)}$$

10.6.2 Measurement of $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$

The ratio of the branching ratios was calculated using Eq. 10.3 replacing the ϕ by the $K^\pm$. The signal event yields were taken from Table 10.3 and the efficencies were taken from Table 10.2. Putting all these numbers together one obtains:

$$\frac{\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)}{\mathcal{B}(B^\pm \to J/\psi\,K^\pm)} = 0.60 \pm 0.07 \text{ (stat)}$$

10.7 Systematic Uncertainties

Different types of systematics are involved in the determination of the relative branching ratio:

- Systematics due to the branching ratio $\mathcal{B}(J/\psi \to \mu^+\mu^-)\,/\,\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = 8.05 \pm 0.89$. This uncertainty was calculated from the world average values [4] assuming no correlations among the two single decays.

- Systematics due to the determination of the signal yield.
- Systematics due to the determination of the efficiencies $\epsilon_{\psi(2S)\,\phi}$ and $\epsilon_{J/\psi\,\phi}$. One expects that most effects cancel out in the ratio. This is due to the fact that both decay modes have very similar topologies. However, for the B_s^0 decays the polarization could be different. The signal Monte Carlo was generated as a pure CP-even state, while for the normalization channel a pure CP-odd state was generated.

10.7.1 Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

All relative uncertainties that enter the calculation of the relative branching ratio are listed in Table 10.10. The dominating uncertainty arises from the determination of the branching fraction $\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$. The relative statistical uncertainties on $\epsilon_{\psi(2S)\,\phi}$ and $\epsilon_{J/\psi\,\phi}$

Table 10.10: The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

Source	Relative Uncertainty [%]
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	1.7
$\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$	11.0
Total ($\mathcal{B}$)	11.1
$\epsilon_{J/\psi\,\phi}/\epsilon_{\psi(2s)\,\phi}$	7.2
CP-odd/even ($J/\psi\phi$)	7.2
Total (sys)	10.2

are 4.5% and 5.6%, respectively. They were combined into a single efficiency uncertainty assuming no correlations. The uncertainty related due to the polarization of the normalization channel $B_s^0 \to J/\psi\,\phi$ was determined from Monte Carlo. Two event samples, one for each polarization state (even and odd), were generated and the efficiencies were calculated. The difference of 7.2% was taken as systematic uncertainty.

The ratio of the relative branching ratios is then

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24\ \text{(stat)} \pm 0.06\ \text{(sys)} \pm 0.07\ (\mathcal{B}).$$

Confidence Interval for $N_{B_s^0 \to \psi(2S)\,\phi}$

Using the number of observed events, the expected number of background events, and their uncertainties, a confidence interval can be constructed. The method proposed by

Conrad *et al.* [91] was used, which employs the Feldman & Cousins (FC) ordering scheme and allows the inclusion of systematic uncertainties (listed in Table 10.10) into the calculation of the confidence belt by integrating over these uncertainties. For 11 observed signal events and 0.8 expected background events, a confidence intervall of $[4.3, 18.8]$ was obtained at 95% CL. Using the result of the Gaussian fit with 9 signal events and 1.7 expected background events, the obtained confidence intervall is $[2.1, 15.3]$.

10.7.2 Measurement of $\mathcal{B}(B^\pm \to J/\psi\, K^\pm)/\mathcal{B}(B^\pm \to \psi(2S)\, K^\pm)$

The relative uncertainties for the calculation of the relative branching ratio $\mathcal{B}(B^\pm \to \psi(2S)\, K^\pm)/\mathcal{B}(B^\pm \to J/\psi\, K^\pm)$ are listed in Table 10.11. The relative statistical uncertainties on $\epsilon_{\psi(2S)\,K^\pm}$ and $\epsilon_{J/\psi\,K^\pm}$ are 3.7% and 1.9%, respectively, and were combined assuming no correlations. As mentioned in Section 10.3, the fitted mean and width of the signal Gaussian for the $\mathcal{B}(B^\pm \to \psi(2S)\, K^\pm)$ candidates are smaller than expected, when all fit parameters are floating. A systematic uncertainty due to this effect was assigned. The difference in the obtained signal event yield for fixed and floating fit parameters is 5%, as shown in Table 10.3.

Table 10.11: The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B^\pm \to \psi(2S)\, K^\pm)/\mathcal{B}(B^\pm \to J/\psi\, K^\pm)$

Source	Relative Uncertainty [%]
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	1.7
$\mathcal{B}(\psi(2S) \to \mu^+\mu^-)$	11.0
Total ($\mathcal{B}$)	11.1
$\epsilon_{J/\psi\,K^\pm}/\epsilon_{\psi(2S)\,K^\pm}$	4.1
Signal yield	5
Total (sys)	4.1

Including all uncertainties the ratio of the branching ratios is then:

$$\frac{\mathcal{B}(B^\pm \to \psi(2S)\, K^\pm)}{\mathcal{B}(B^\pm \to J/\psi\, K^\pm)} = 0.60 \pm 0.07\ \text{(stat)} \pm 0.04\ \text{(sys)} \pm 0.06\ (\mathcal{B}).$$

10.8 Conclusions

The observation of the decay channel $B_s^0 \to \psi(2S)\,\phi$ with a significance of 5.99σ based on counting statistics has been presented. For the first time, also the ratio of the branching

ratios

$$R(B_s^0 \to \psi(2S)\,\phi) = \frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24\ \text{(stat)} \pm 0.06\ \text{(sys)} \pm 0.07\ (\mathcal{B})$$

has been measured. In addition, a measurement of the relative branching ratio

$$R(B^\pm \to \psi(2S)\,K^\pm) = \frac{\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)}{\mathcal{B}(B^\pm \to J/\psi\,K^\pm)} = 0.60 \pm 0.07\ \text{(stat)} \pm 0.04\ \text{(sys)} \pm 0.06\ (\mathcal{B})$$

has been presented. A comparison between the values obtained in this analysis and published values is given in Table 10.12, where the additional short notations

$$R(B_d^0 \to \psi(2S)\,K^0) = \frac{\mathcal{B}(B_d^0 \to \psi(2S)\,K^0)}{\mathcal{B}(B_d^0 \to J/\psi K^0)}$$

and

$$R(B_d^0 \to \psi(2S) K^*) = \frac{\mathcal{B}(B_d^0 \to \psi(2S) K^*)}{\mathcal{B}(B_d^0 \to J/\psi K^*)}$$

are used. The result obtained for $R(B^\pm \to J/\psi\,K^\pm)$ shows a good agreement with previous measurements and is of comparable statistical uncertainty.

Table 10.12: The measured relative branching ratios from this analysis and published results. The first uncertainty for the published values is due to statistics and the second due to systematics.

Decay	This analysis	Published result
$R(B_d^0 \to \psi(2S)\,K^0)$	-	$0.82 \pm 0.13 \pm 0.12$ (PDG 04,[4])
$R(B_d^0 \to \psi(2S)\,K^*)$	-	$0.61 \pm 0.19 \pm 0.06$ (PDG 04,[4])
$R(B^\pm \to \psi(2S)\,K^\pm)$	$0.60 \pm 0.07 \pm 0.04 \pm 0.06$	$0.64 \pm 0.06 \pm 0.06$ (BaBar 02,[100])
$R(B_s^0 \to \psi(2S)\,\phi)$	$0.58 \pm 0.24 \pm 0.06 \pm 0.07$	

Chapter 11

Outlook and Summary

The Tevatron is expected to deliver luminosity to the experiments until the end of the year 2009. This should yield a total integrated luminosity of 4 fb^{-1} to 8 fb^{-1}, depending on the performance of the Tevatron accelerator and foreseen upgrades of the accelerator. Presently, different scenarios about a possible extension of the Tevatron running period are discussed.

11.1 Outlook for $B_s^0 \to \mu^+\mu^-$

In Figure 11.1, the expected average upper limit for the decay $B_s^0 \to \mu^+\mu^-$ is shown as a function of integrated luminosity for the DØ experiment. The horizontal band indicates the Standard Model prediction with its uncertainty. To ensure a reasonably low expected background even at high integrated luminosities, the expected limits for 1 fb^{-1} and 2 fb^{-1} have been re-optimised by re-sampling the background of the available data at 300 pb^{-1}. The other points have been scaled according to the luminosity. The shown band indicates a $\pm$10% variation of the number of normalisation events. The plot makes it clear that even with twice the statistics expected at the end of 2009, the Standard Model prediction for $B_s^0 \to \mu^+\mu^-$ can not be reached. However, improved limits on this decay mode will further reduce possible models beyond the Standard Model.

One way to improve the sensitivity is by combining results with the other Tevatron experiment, CDF. This was done for this analysis and combined limits for both decay modes, $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ [37], have been set. A Bayesian integration method was used to calculate the combined limits [92]. The method takes into account correlated and uncorrelated systematic uncertainties between the two experiments. The combined results exclude branching ratios of $B_s^0 \to \mu^+\mu^- > 1.5 \times 10^{-7}$ and $B_d^0 \to \mu^+\mu^- > 4.0 \times 10^{-8}$ at a 95% CL. These are presently the most stringent limits on these decays. A similar combination was done for the expected average limits of CDF and DØ.

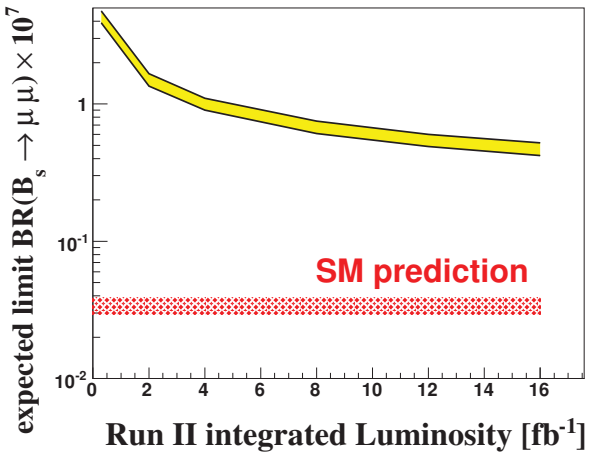

Figure 11.1: Expected average upper limit at a 95% CL for the decay $B_s^0 \to \mu^+\mu^-$ as a function of integrated luminosity.

In Figure 11.2, the combined expected average upper limit from the Tevatron experiments is shown as a function of the integrated luminosity collected by each experiment. The two additional lines represent the 3σ and 5σ discovery potential if a signal is present.

11.2 Outlook for $B_s^0 \to \phi\,\mu^+\mu^-$

Figure 11.3 shows the expected average upper limit as a function of integrated luminosity for the decay $B_s^0 \to \phi\,\mu^+\mu^-$. The predicted Standard Model branching fraction is indicated as a horizontal band [94] with a 20% uncertainty. The extrapolation was obtained by scaling the number of background and normalisation events according to the total integrated luminosity. A re-optimisation for each point was not performed since the expected number of background events is still reasonably small even at higher integrated luminosities. The limit band indicates a $\pm 10\%$ variation of the number of normalisation events. The Standard Model branching fraction should be observable with an integrated luminosity of 4 fb^{-1}.

11.3 Summary

This thesis presented searches for rare B_s^0 decays with the DØ detector at $\sqrt{s} = 1.96$ TeV, using a data sample of 300 pb^{-1} integrated luminosity collected between 2002 and 2004.

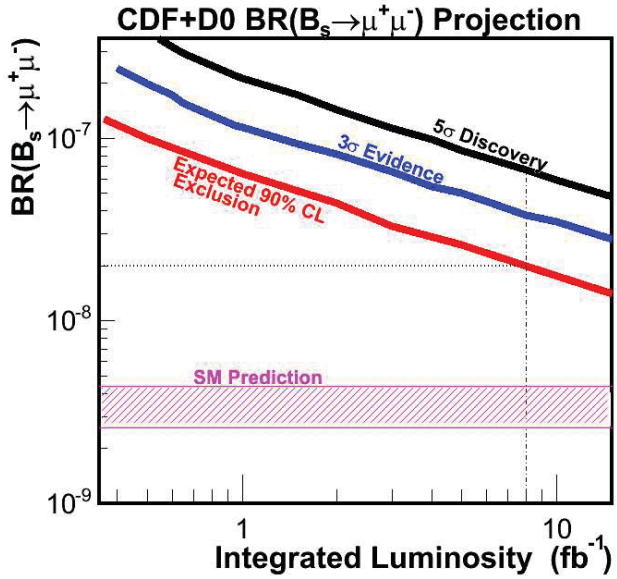

Figure 11.2: Expected average upper limit from the Tevatron for the decay $B_s^0 \to \mu^+\mu^-$ as a function of integrated luminosity.

The rare decays $B_s^0 \to \mu^+\mu^-$ and $B_s^0 \to \phi\,\mu^+\mu^-$ are flavour-changing neutral current decays and are forbidden in the Standard Model at tree level. However, they proceed at a very low rate through higher order processes. Within the Standard Model the low predicted rates place the observation of these modes outside the reach of this study, but several extensions to the Standard Model predict a substantial enhancement of the branching fractions.

In the search for the decay $B_{s,d}^0 \to \mu^+\mu^-$, four candidate events were observed while 4.3±1.2 events were expected as background. In the absence of a signal, a limit on the branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ was set. The resulting limit on the branching fraction limit at a 95% (90%) CL, including all the statistical and systematic uncertainties, is given by

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 3.7 \times 10^{-7}\ (3.0 \times 10^{-7}).$$

The result was combined with the CDF experiment to obtain the currently most stringent limits on the decay modes $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ [37]. The obtained limits can be used to constrain models of new physics beyond the Standard Model. The decay $B_s^0 \to \mu^+\mu^-$ gives information not only on SUSY breaking mediation mechanisms but also places a severe limit on the neutralino dark matter scattering cross section within a large class of supergravity models in the large $\tan\beta$ region.

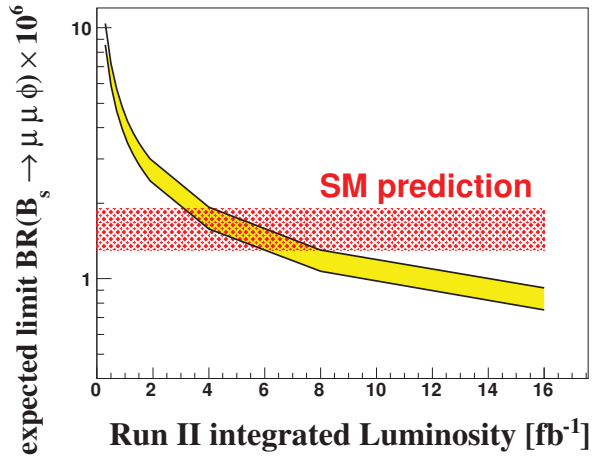

Figure 11.3: Expected average upper limit at a 95% CL for the decay $B_s^0 \to \phi\,\mu^+\mu^-$ as a function of integrated luminosity.

In the search for the decay $B_s^0 \to \phi\,\mu^+\mu^-$, no candidate event was observed in the signal region, while the background interpolation from the sidebands predicted 1.6±0.4 events. Finding no evidence for a signal, an upper limit on the branching fraction was set. Including statistical and systematic uncertainties, and using the central value of the world average branching ratio for $B_s^0 \to J/\psi\,\phi$, a limit of

$$\mathcal{B}(B_s^0 \to \phi\,\mu^+\mu^-) < 4.1\,(3.2)\ \times 10^{-6} \tag{11.1}$$

at a 95% (90%) CL was found following a Frequentist approach. This improves the currently best published limit from CDF [30] by one order of magnitude.

In addition to these searches for rare decays, also the decay mode $B_s^0 \to \psi(2S)\,\phi$ was studied. An observation with a significance of 5.99σ based on counting statistics was obtained. For the first time, the relative branching ratio

$$\frac{\mathcal{B}(B_s^0 \to \psi(2S)\,\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = 0.58 \pm 0.24\ \text{(stat)} \pm 0.09\ \text{(sys)}$$

was measured. In addition, a measurement of the relative branching

$$\frac{\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)}{\mathcal{B}(B^\pm \to J/\psi\,K^\pm)} = 0.60 \pm 0.07\ \text{(stat)} \pm 0.07\ \text{(sys)}$$

has been presented. The obtained value for the $B^\pm$ meson decay agrees well with previous measurements. The relative fraction of $B_s^0 \to \psi(2S)\,\phi$ decays to $B_s^0 \to J/\psi\,\phi$ decays is consistent with the corresponding fraction for decays of the $B^\pm$ meson measured by B factories.

Appendix A

Deriving the Master Formula

When observing the decay of a B meson to a specific final state f, the expected number of observed events is given by

$$N_{obs}(f) = \int \mathcal{L}\, dt \cdot \sigma(B) \cdot \mathcal{B}(B \to f) \cdot \epsilon \tag{A.1}$$

where $\int \mathcal{L}\, dt$ is the total integrated luminosity of the collected data sample, $\sigma(B)$ the B meson production cross section, $\mathcal{B}(B \to f)$ the branching fraction of the specific decay and ϵ the corresponding efficiency and acceptance. The inclusive $b\bar{b}$ quark cross section has been measured, the cross section $\sigma_{b\bar{b}}$ for a specific B meson however, depends on the fragmentation into that B meson type. In the search for the rare decay $B_s^0 \to \mu^+\mu^-$, the following cross sections are involved

$$\begin{aligned} \sigma_{B\pm} &= \sigma_{b\bar{b}} \cdot \bar{b} \to B^{\pm} \\ \sigma_{B_d} &= \sigma_{b\bar{b}} \cdot \bar{b} \to B_d^0 \\ \sigma_{B_s} &= \sigma_{b\bar{b}} \cdot \bar{b} \to B_s^0 \end{aligned} \tag{A.2}$$

where $\sigma_{b\bar{b}}$ is the total $b\bar{b}$ production cross section, and $\bar{b} \to B^{\pm}$, $\bar{b} \to B_d^0$ and $\bar{b} \to B_s^0$ are the fragmentation fractions of a b or $\bar{b}$ quark producing a B_s^0, a $B^{\pm}$ or a B_d^0 meson, respectively. Using Eq. A.1, the numbers of observed events are given by

$$\begin{aligned} N_{B_s^0 \to \mu^+\mu^-} &= \int \mathcal{L}\, dt \cdot \sigma_{B_s^0} \cdot \mathcal{B}(B_s^0 \to \mu^+\mu^-) \cdot \epsilon_{\mu^+\mu^-}^{B_s^0} \\ N_{B_d^0 \to \mu^+\mu^-} &= \int \mathcal{L}\, dt \cdot \sigma_{B_d^0} \cdot \mathcal{B}(B_d^0 \to \mu^+\mu^-) \cdot \epsilon_{\mu^+\mu^-}^{B_d^0} \\ N_{B^{\pm} \to J/\psi\, K^{\pm}} &= \int \mathcal{L}\, dt \cdot \sigma_{B^{\pm}} \cdot \mathcal{B}(B^{\pm} \to J/\psi\, K^{\pm}) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-) \cdot \epsilon_{\mu^+\mu^- K}^{B^{\pm}}. \end{aligned} \tag{A.3}$$

By performing a measurement relative to a decay with a similar final state, common reconstruction efficiencies as well as the total luminosity will cancel, if the same dataset

is used. The number of observed events in our $B_s^0 \to \mu^+\mu^-$ signal region, N, is the sum of $N_{B_s^0\to\mu^+\mu^-}$ and $N_{B_d^0\to\mu^+\mu^-}$ since the mass resolution is not good enough to separate the two decay modes. When calculating the ratio with respect to the normalisation channel $B^\pm \to J/\psi\, K^\pm$, this yields

$$\frac{N}{N_{B^\pm\to J/\psi\, K^\pm}} = \frac{\sigma_{B_s^0}\cdot\mathcal{B}(B_s^0\to\mu^+\mu^-)\cdot\epsilon^{B_s^0}_{\mu^+\mu^-} + \sigma_{B_d^0}\cdot\mathcal{B}(B_d^0\to\mu^+\mu^-)\cdot\epsilon^{B_d^0}_{\mu^+\mu^-}}{\sigma_{B^\pm}\cdot\mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)\cdot\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}. \tag{A.4}$$

Defining R as the ratio of the branching ratios $R = \mathcal{B}(B_d^0\to\mu^+\mu^-)/\mathcal{B}(B_s^0\to\mu^+\mu^-)$ one can write this as

$$\begin{aligned}\frac{N}{N_{B^\pm\to J/\psi\, K^\pm}} = \frac{\sigma_{B_s^0}}{\sigma_{B^\pm}}\cdot\frac{\epsilon^{B_s^0}_{\mu^+\mu^-}}{\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}\cdot\frac{\mathcal{B}(B_s^0\to\mu^+\mu^-)}{\mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)} + \\ \frac{\sigma_{B_d^0}}{\sigma_{B^\pm}}\cdot\frac{\epsilon^{B_d^0}_{\mu^+\mu^-}}{\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}\cdot\frac{\mathcal{B}(B_s^0\to\mu^+\mu^-)\cdot R}{\mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)}\end{aligned} \tag{A.5}$$

and extracting common factors it can be written as

$$\begin{aligned}\frac{N}{N_{B^\pm\to J/\psi\, K^\pm}} = \frac{\mathcal{B}(B_s^0\to\mu^+\mu^-)}{\mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)}\times \\ \frac{1}{\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}\cdot\left(\epsilon^{B_s^0}_{\mu^+\mu^-}\cdot\frac{\sigma_{B_s^0}}{\sigma_{B^\pm}} + \epsilon^{B_d^0}_{\mu^+\mu^-}\cdot\frac{\sigma_{B_d^0}}{\sigma_{B^\pm}}\cdot R\right).\end{aligned} \tag{A.6}$$

This yields, with the usage of the fragmentation defined in Equation A.2, the relation

$$\begin{aligned}\mathcal{B}(B_s^0\to\mu^+\mu^-) = \frac{N}{N_{B^\pm\to J/\psi\, K^\pm}}\times \\ \mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)\cdot\frac{\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}{\epsilon^{B_s^0}_{\mu^+\mu^-}}\cdot\frac{1}{\frac{\bar{b}\to B_s^0}{\bar{b}\to B^\pm} + R\cdot\frac{\bar{b}\to B_d^0}{\bar{b}\to B^\pm}\cdot\frac{\epsilon^{B_d^0}_{\mu^+\mu^-}}{\epsilon^{B_s^0}_{\mu^+\mu^-}}}\end{aligned} \tag{A.7}$$

from which the formula

$$\begin{aligned}\mathcal{B}(B_s^0\to\mu^+\mu^-)\cdot\left(1 + R\cdot\frac{\epsilon^{B_d^0}_{\mu^+\mu^-}}{\epsilon^{B_s^0}_{\mu^+\mu^-}}\cdot\frac{\bar{b}\to B_d^0}{\bar{b}\to B_s^0}\right) = \\ \frac{\mu(n_{obs}, n_{back})}{N_{B^\pm\to J/\psi\, K^\pm}}\cdot\frac{\epsilon^{B^\pm}_{\mu^+\mu^-K^\pm}}{\epsilon^{B_s^0}_{\mu^+\mu^-}}\cdot\frac{\bar{b}\to B^\pm}{\bar{b}\to B_s^0}\cdot\mathcal{B}(B^\pm\to J/\psi\, K^\pm)\cdot\mathcal{B}(J/\psi\to\mu^+\mu^-)\end{aligned} \tag{A.8}$$

is obtained, where, for calculating an upper limit, the number of observed events N was replaced by the event upper limit $\mu(n_{obs}, n_{back})$.

Appendix B

Signal Monte Carlo Data Matching for $B_s^0 \to J/\psi\,\phi$ Events

A sample obtained using less restrictive cuts for the normalisation channel $B_s^0 \to J/\psi\,\phi$ was employed to perform a comparison between data and Monte Carlo simulation. The main differences with respect to the sampel used in the analysis are that no trigger was required, the cut on bad runs was omitted and no cuts on the muon quality and the muon transverse momentum were applied. This yielded a signal of 340 candidate events. The distributions in the signal region were corrected for the background interpolated from the sidebands. Monte Carlo and data distributions for the kinematic variables are shown in Figure B.1 to B.3, and distributions for the discriminating variables are shown in Figure B.4 for the $B_s^0 \to J/\psi\,\phi$ candidate events. The Monte Carlo events were weighted according to the p_T of the B. The weighting procedure was applied as described in Section 8.4.

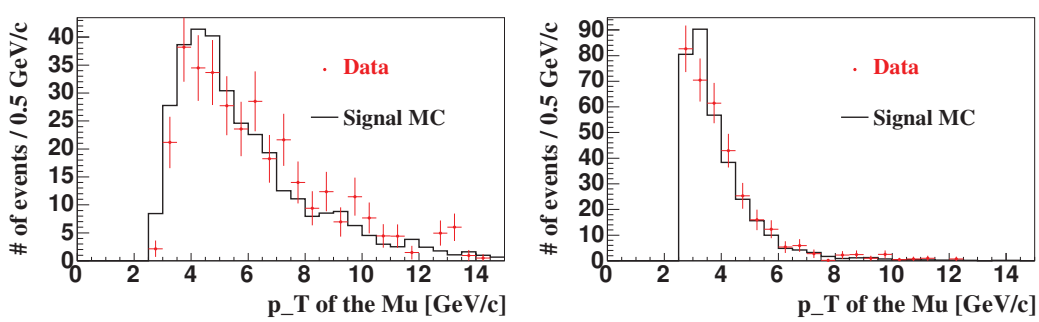

Figure B.1: Data and Monte Carlo events comparison for the leading Muon (left) and trailing Muon (right) for $B_s^0 \to J/\psi\,\phi$ candidates.

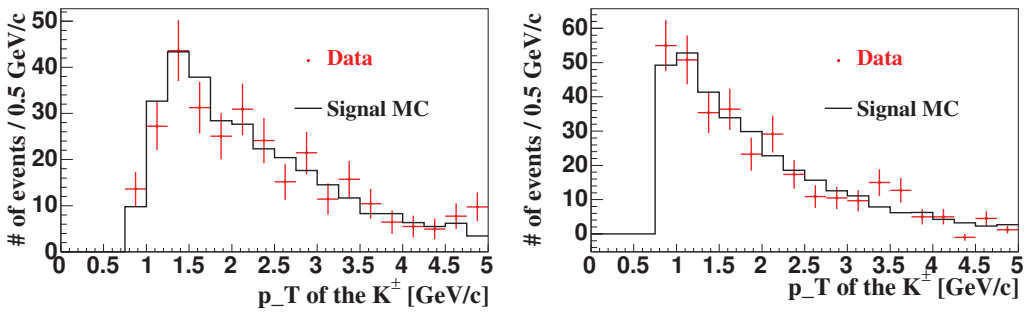

Figure B.2: Data and Monte Carlo events comparison for the leading Kaon (left) and trailing Kaon (right) for $B_s^0 \to J/\psi\,\phi$ candidates.

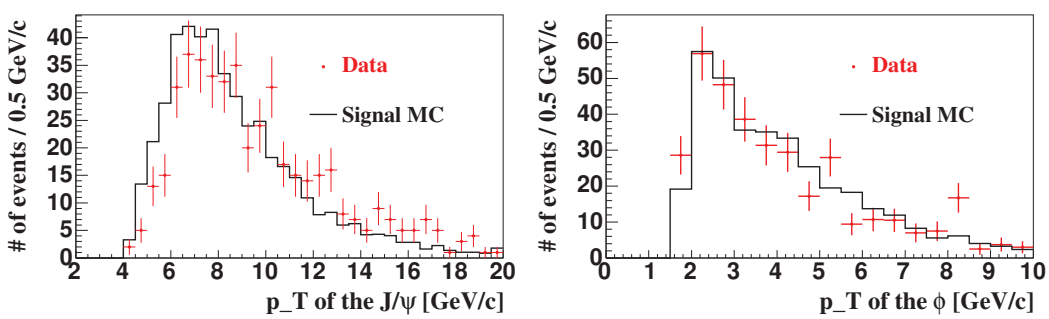

Figure B.3: Data and Monte Carlo events comparison for the J/ψ candidate (left) and ϕ candidate (right) for $B_s^0 \to J/\psi\,\phi$ candidate events.

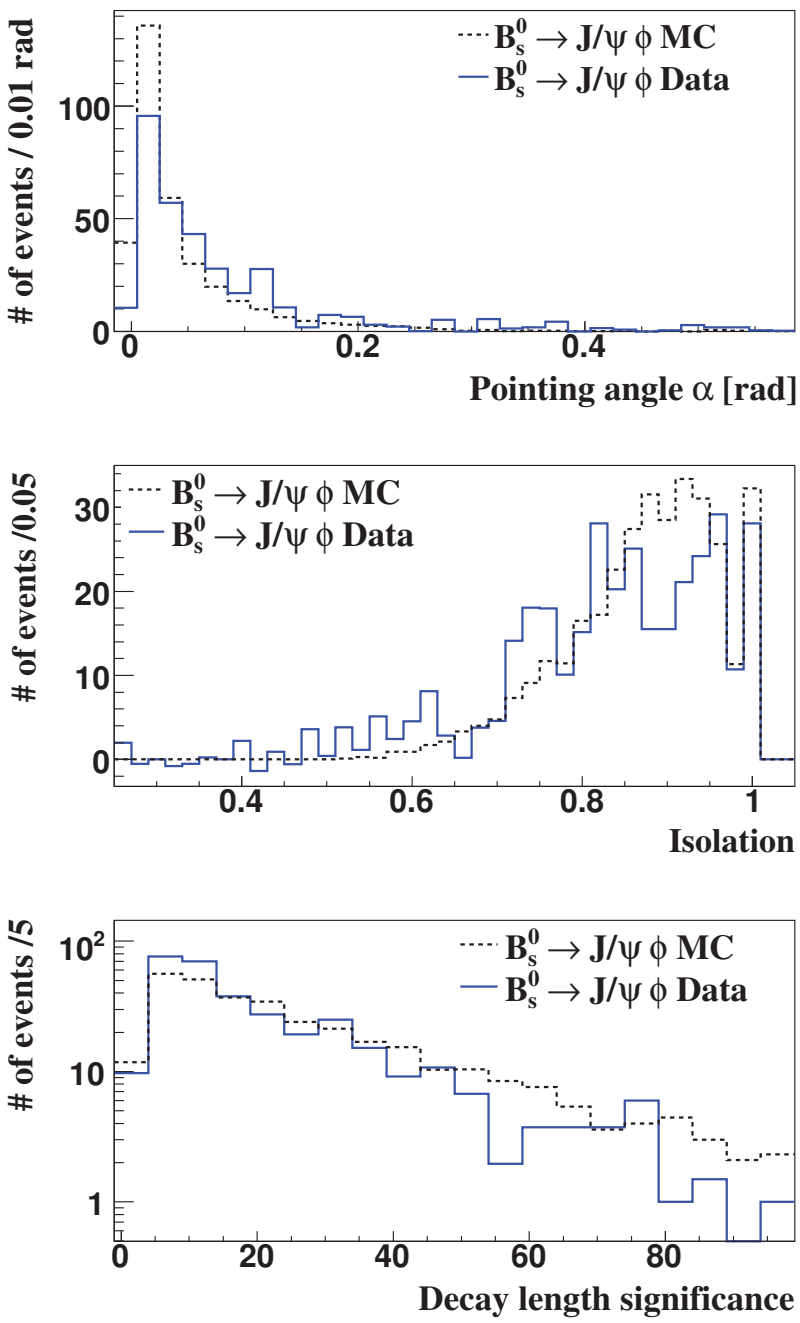

Figure B.4: Comparison of the discriminating variables for $B_s^0 \to J/\psi\,\phi$ Monte Carlo and background subtracted data events.

Appendix C

Depletion Voltage Studies

The DØ Silicon Microstrip Tracker (SMT) was designed for use in Run IIa with an expected integrated luminosity of 2 fb^{-1}. It was predicted that the SMT could withstand about 4 fb^{-1} before succumbing due to effects of radiation damage. Special runs to study the depletion voltage have been taken for three different accumulated luminosities. A method has been set up to fit the charge collection curve and extract the depletion voltage of the sensors. It is expected that at the currently accumulated integrated luminosity the effect of radiation damage should be visible, although type inversion has not yet taken place.

The model of charge collection in irradiated sensors proposed in [101] was used to determine the depletion voltage from measured charge collection curves. In the following Sections, the model and the obtained results are described.

C.1 Radiation Damage

An estimate of the SMT lifetime based on calculations and measurements of radiation damage can be found in [102]. The expectation is that the lifetime of the SMT will be limited by micro-discharge breakdown of the junction in the double-sided silicon detectors in the inner four barrels. Micro-discharges will begin to occur at bias voltages of $\approx$ 150 V and all channels will fail at $\approx$ 200 V. In addition, for double-sided detectors, breakdown of the coupling capacitors will start soon after bias voltages exceed 100 V per side. Figure C.1 shows the voltage required to deplete the sensors as a function of the fluence of 1 MeV eq. n/cm^2 [103]. The line corresponds to the so-called Hamburg model [104] for the innermost layer. The vertical lines indicate the accumulated luminosities where bias voltage scans were performed. From this figure and the above limits for the bias voltage one finds an operation limit of $\approx$ 1.5x10^{13} 1 MeV eq. n/cm^2 (corresponding to $\approx$ 4 fb^{-1} or 2 MRad). Errors in estimating the micro-discharge formation and dose accumulation

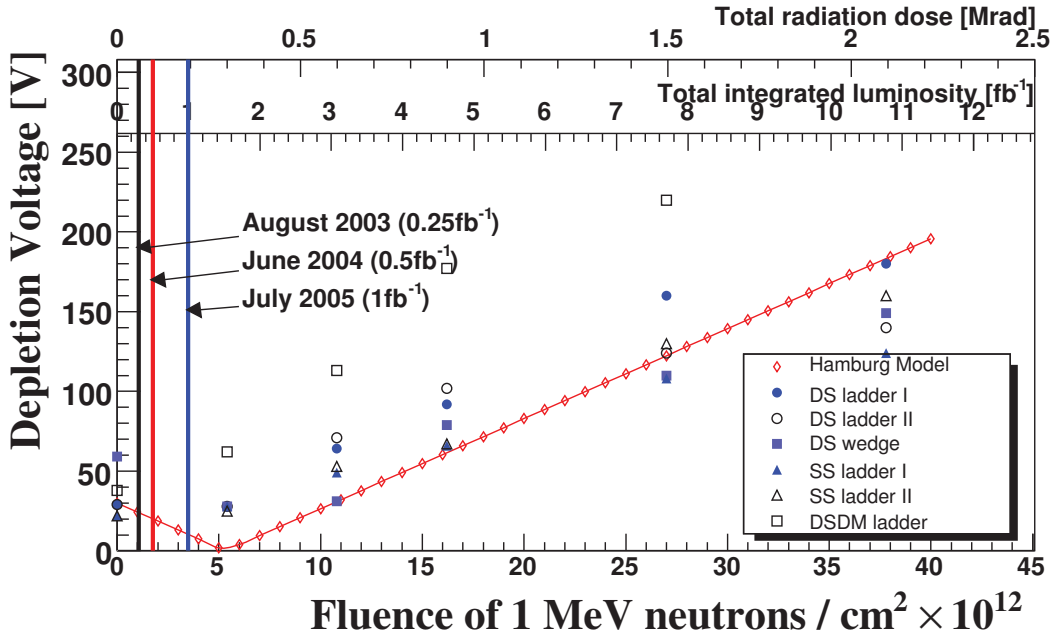

Figure C.1: Depletion voltage from the Hamburg Model predictions and from laser measurements for various detectors.

lead to an uncertainty of about 50% on this result. Different varieties and locations of detectors are listed in Table C.1.

C.2 Method for Measuring the Depletion Voltage

The depletion voltage for one superlayer at a time was determined using recorded data for tracks from $p\overline{p}$ collisions. Data were taken for three different integrated luminosities (0.25 fb^{-1}, 0.5 fb^{-1} and 1 fb^{-1}) for all layers. The bias voltage for one layer was scanned from 10% to 100% in steps of 10% while the rest of the detectors were kept at 100% of the current operating bias voltage. Clusters associated to tracks that were identified by the CFT and the other ladders were used to measure the cluster charge as a function of the applied bias voltage. As long as the detectors are underdepleted, the charge collection increases linear and saturates around the full depletion voltage. For these measurements the SMT was fully read out (no sparsification was performed in the SVXII chips).

Table C.1: SMT detector geometries and types: DS-double sided, SS-single sided, DSDM double sided double metal. The central barrels are Barrel 2 to 5, the outer barrels are 1 and 6.

Module	Type	Layer	Inner radius	Outer radius	Manufacturer
Central barrels	DSDM	1	2.715	3.645	Micron
Central barrels	DS	2	4.55	5.554	Micron
Central barrels	DSDM	3	6.768	7.582	Micron
Central barrels	DS	4	9.101	10.51	Micron
Outer barrels	SS	1	2.715	3.645	Micron
Outer barrels	DS	2	4.55	5.554	Micron
Outer barrels	DS	3	6.768	7.582	Micron
Outer barrels	SS	4	9.101	10.51	Micron

C.3 The Model

After hadron irradiation, radiation-induced defects produced in the crystal lattice act as trapping centres. A fraction of the charge generated in the active volume may be trapped for longer than the integration time of the readout amplifier. Consequently, charge trapping can cause a signal deficit. In the model used here, it is assumed that trapping depends on the carrier velocity since slowly moving carriers will spend longer time in the vicinity of the trapping centres. This means that above full depletion the signal size in irradiated detectors continues to grow until the saturation velocity of the charge carriers is reached. Beyond this point, trapping remains constant and the signal reaches a plateau.

As a consequence of trapping, only a fraction of the carriers produced at a point x_0 will reach another point located at x_1. That fraction will be proportional to $exp[-(x_1 - x_0)/\lambda]$, with λ being the signal attenuation length. A linear relationship between λ and v, the carrier velocity, is assumed:

$$\lambda(v) = \lambda_0 + \lambda_1 \cdot \frac{v}{v_s}. \tag{C.1}$$

Where, v_s is the saturation velocity. The collected signal for an applied bias voltage V is then:

$$q(V) = \frac{Q_0}{\omega_0} \int_0^{\omega(V)} exp\Big(-\int_x^{\omega_0} \frac{dx'}{\lambda(x')}\Big) dx, \tag{C.2}$$

where Q_0 is the total charge deposited inside the detector volume and ω_0 is the thickness of the detector. The outer integral is performed over the position of the charge generation inside the detector's depleted bulk (x) whilst the inner integral is performed from that

point to the readout plane of the sensor. The different parameters in Equation C.2 are:

$$\omega(V) = \begin{cases} \omega_0 \cdot \sqrt{V/V_{dep}} & \text{if } V < V_{dep} \\ \omega_0 & \text{if } V \geq V_{dep} \end{cases} \tag{C.3}$$

$$\lambda(x) = \lambda_0 + \lambda_1 \cdot \frac{v(x)}{v_s} \tag{C.4}$$

$$v(x) = \mu(x) \cdot \epsilon(x) \tag{C.5}$$

$$\mu(x) = \frac{\mu_0}{1 + \mu_0 \cdot \epsilon(x)/v_s} \tag{C.6}$$

$$\epsilon(x) = \frac{2 \cdot V_{dep}}{\omega_0^2} \cdot \big(\omega(V) - x\big) \tag{C.7}$$

where $\epsilon(x)$ is the electric field in the silicon detector. The five free parameters of the model are the total the charge (Q_0), the depletion voltage (V_{dep}), the saturation velocity (v_s), and λ_0 and λ_1 which describe the effect of charge trapping. Since v_s is expected to be ≈ 100 μm/ns [105] its value was fixed whereas the remaining four parameters were fitted. The value of the charge mobility parameter μ_0 was calculated for the temperature at which the measurements were performed using the values given in [105]. It should be noted that in the absence of trapping ($\lambda \gg \omega_0$) Eq. C.2 becomes just $q(V) = Q_0 \cdot \omega(V)/\omega_0$. That is to say, the signal amplitude is proportional to the thickness of the depleted region.

C.4 Results

In order to apply the model, one needs to know the actual voltage applied to the sensor. In the case of the DØ detector, there are close to 1000 modules and their operation voltages are not all the same. For this reason, approximately 20 modules that were operated at the same operation voltage were selected for this study. In addition, only modules from the central barrels where selected, which ensured that all modules had the same type of sensor. As an example, the obtained signals for different applied operation voltages for the data accumulated at 1fb^{-1} are shown in Fig. C.2. The plotted signal amplitudes were corrected for the path length of the particle within the silicon. For each setting, a Landau function convolved with a Gaussian was fit to the signal shape. The width of the Landau function was fixed to its most probable value (MPV) according to Bichsel [106] for the actual thickness of the silicon sensor. One exception was the data taken at only 10% of the nominal operation voltage. The Landau fit failed and only a Gaussian function was fit to extract the mean value of the signal.

The signal heights versus the applied bias voltages for the three different data sets are shown in Fig. C.3. The curves were fitted with the described model and the four parameters were extracted. The results of the fits are summarised in Table C.2. As expected, the

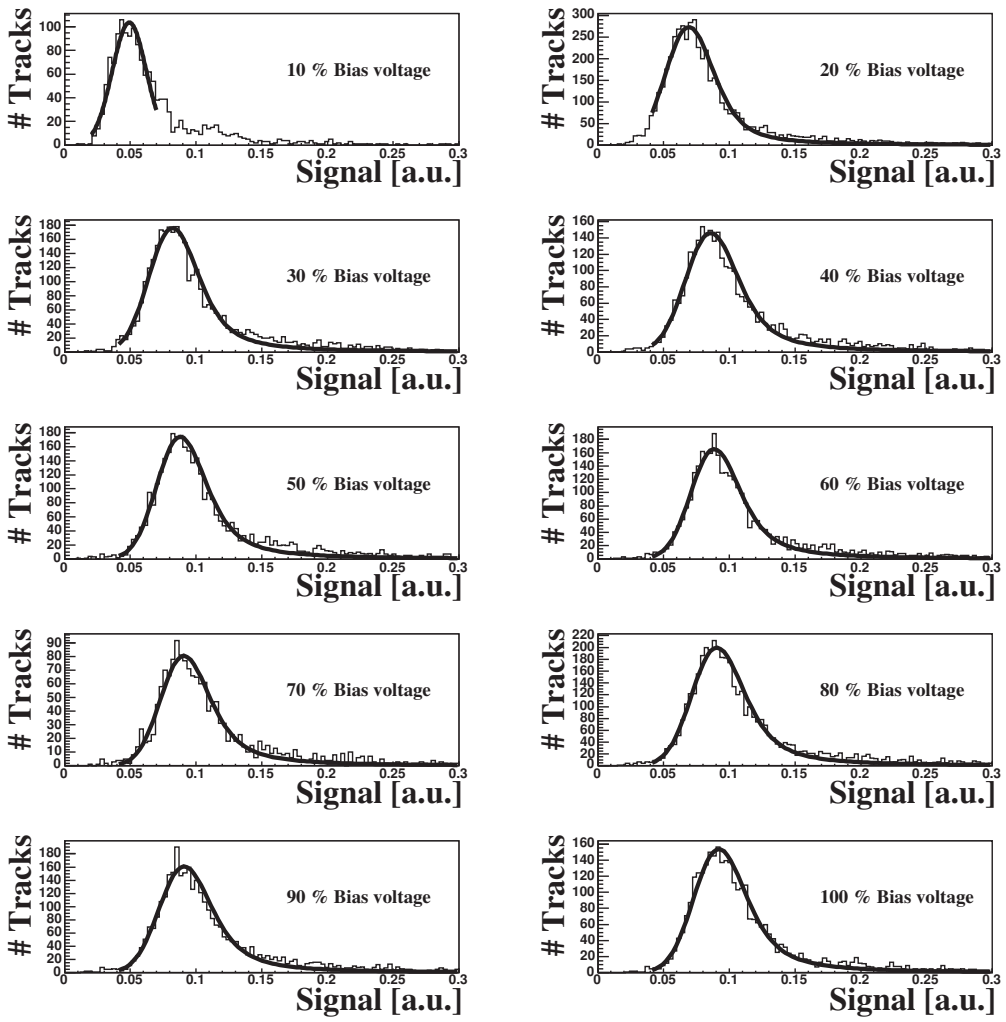

Figure C.2: Signal distributions.

extracted depletion voltage V_{dep} decreases with increasing irradiation dose, or integrated luminosity.

The results for the other fit parameters are not in agreement with the trend found in [101]. It seems from the fit that the trapping does not increase with higher radiation dose. Since no clear correlation between λ_1 and the radiation dose is seen. These results suggest that the model, though providing a good fit to the data is not applicable in our case. It should also be mentioned, that the radiation doses investigated in [101] were a factor 100 higher than those accumulated in the DØ silicon detector.

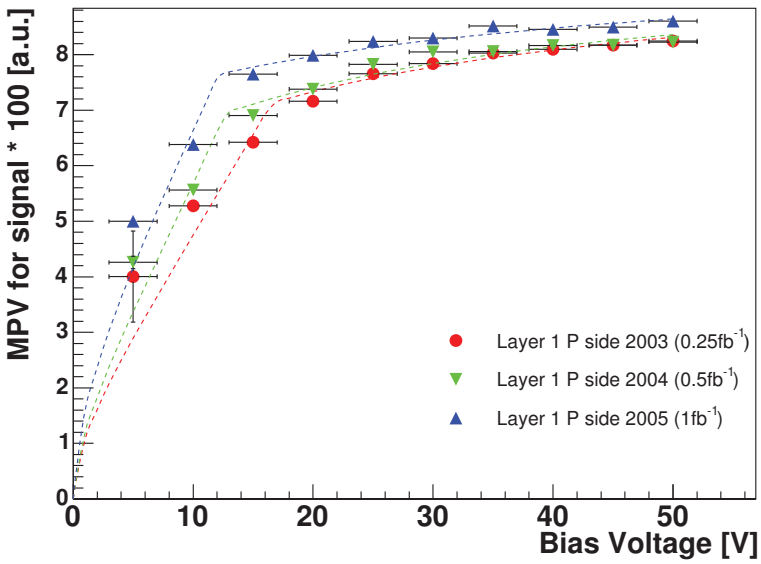

Figure C.3: Charge collection as a function of the applied bias voltage.

Table C.2: Fit results

Luminosity	V_{dep} [V]	Q_0 [a.u.]	λ_0	λ_1
0.25fb^{-1}	16.5	10.0	334	7107
0.5fb^{-1}	12.7	10.0	321	7429
1.0fb^{-1}	12.2	10.1	459	7340

List of Figures

1.1 The unitarity triangle. 9

2.1 Example of penguin and box diagrams 12

2.2 Typical diagrams for FCNC decays for the six different classes as described in the text. 14

2.3 Main diagrams contributing to the decay $B_s^0 \to \phi\,\ell^+\ell^-$. The diagrams (a) and (b) are short distance processes and (c) represents the long distance processes leading to $c\bar{c}\,(J/\psi,\,\psi(2S),\,\ldots,\,\psi^{(n)})$ intermediate states. 17

2.4 Decay processes contributing to $B_{s,d} \to \ell^+\ell^-$ in the Standard Model. . . 22

2.5 The branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ in the 2HDM as a function of the charged Higgs mass M_{H^+} and various values of $\tan\beta$. 27

2.6 $\sigma_{\tilde{\chi}p}$ and $B(B_s^0 \to \mu^+\mu^-$ within mSUGRA with universal and non-universal Higgs mass parameters. 28

2.7 Contours of constant χ^2 for m_{16} = 3 TeV/c^2 and m_A = 500 GeV/c^2 (left side) and 700 GeV/c^2 (right side), indicating the allowed region for $B_s^0 \to \mu^+\mu^-$ [55]. 30

3.1 The accelerator facilities at FERMILAB. 33

3.2 The integrated luminosity per week and total integrated luminosity for Run II from May 2001 until August 2005 in pb^{-1}. 34

3.3 The initial luminosity for each fill for the Tevatron Run II from May 2001 until August 2005. 35

3.4 Cross-sectional view of the DØ Run II detector. 36

3.5 Luminosity Monitor layout. The $r-\phi$ view is shown on the left, the $r-z$ view of the two arrays is shown on the right. 37

3.6 The sketch on the left shows the differentiation between inelastic collisions and beam halo. Expected z vertex distribution for inelastic collisions, centred at $z = 0$ cm, p halo centred at $z = -140$ cm and $\bar{p}$ halo centred at $z = 140$ cm (right). 38

3.7 Cross-sectional view of the DØ tracking and preshower system. 39

3.8 Isometric view of the DØ silicon vertex detector. 40

3.9 Cross section of the silicon vertex detector barrel structure. 41

3.10 Isometric view of the central and two end calorimeters. 43

3.11 Schematic view of a quarter of the DØ calorimeter. 44

3.12 Schematic view of a calorimeter cell. 45

3.13 A cut-away view of the muon system. 46

3.14 Side view of the DØ muon system. 47

3.15 The design of the DØ Level 1 and Level 2 trigger system. 50

4.1 Invariant mass distributions for the processes $K_s^0 \to \pi^+\pi^-$ (left) and $\Lambda^0 \to p^+\pi^-$ (right). Also shown is a Gaussian distribution plus a linear fit to the signal and background. 52

4.2 Impact parameter resolution measured in data and in simulated single muon events. The lines show fits to data and Monte Carlo events. 54

4.3 Invariant mass distributions for the processes $\Xi^\pm \to \Lambda^0\pi^\pm$ (top left) and $\Omega^\pm \to \Lambda^0 K^\pm$ (bottom left). The reconstruction of the processes are examples of events with multiple tracks with impact parameter in the order of centimetres. A sketch of the decay $\Xi^- \to \Lambda^0\pi^- \to p^+\pi^-\pi^-$ is also shown (right). 55

5.1 Invariant di-muon mass spectrum of various meson resonances: ω, ϕ, J/ψ, ψ' and Υ. 58

6.1 Pictogram for visualisation of the discriminating variables for the decay $B_s^0 \to \mu^+\mu^-$. 62

6.2 Visualisation of the advantage of a random grid (right side), over a regularly spaced grid (left side) search for two variables x and y. 63

6.3 Visualisation of the the idea of separating two groups with a linear discriminant and indicating with a line the best separation of the two groups. 65

8.1 The mass resolution in the signal Monte Carlo after pre-selection. 75

8.2 The Monte Carlo simulated $p_t(\mu^+\mu^-)$ spectra for the signal (right) and the normalisation channel (left) after the pre-selection. 76

8.3 Invariant di-muon mass distribution of data events after the pre-selection. A total of 45'696 events remain. 77

8.4 Discriminating variables after the pre-selection for signal Monte Carlo and data events from the sidebands. The arrows indicate the cut values that were obtained after optimisation. 79

8.5 The remaining background for the full data sample with the standard discriminating variables. The line shows the linear fit to determine the expected background. 80

8.6 The LHR variable R on the left side and the signal efficiency versus background rejection for this Likelihood Ratio on the right side. 80

8.7 The normalisation channel $B^\pm \to J/\psi\, K^\pm$ for the full data sample after all final cuts. 81

8.8 Comparison of the discriminating variables for $B^\pm \to J/\psi\, K^\pm$ Monte Carlo and data events after all final cuts. 82

8.9 The mass distribution after opening the box for the data sample with the standard discriminating variables. Four events were found in the $\pm 2\sigma$ region. 88

9.1 Invariant mass spectra generated with the two decay models [21, 93] before (left) all selection cuts and the efficiency as function of the di-muon mass (right) after all cuts for the decay model of Ali *et al.* [22]. . . 91

9.2 Di-muon invariant mass distribution of the loose B_s^0 candidate selection. 91

9.3 Discriminating variables after the pre-selection for signal Monte Carlo and data events from the sidebands for the $B_s^0 \to \phi\, \mu^+\mu^-$ analysis. The arrows indicate the cut values that were obtained after optimisation. The normalisation is done on the number of signal Monte Carlo and sideband data events after pre-selection. 95

9.4 The remaining background for the full data sample with the standard discriminating variables. 96

9.5 Invariant masses obtained from data for the ϕ candidates (left) and J/ψ candidates (right). 96

9.6 Invariant masses obtained from Monte Carlo for the ϕ candidates (left) and J/ψ candidates (right). 97

9.7 The normalisation channel $B_s^0 \to J/\psi\, \phi$ for the full data sample. 97

9.8 Mass spectra for Data (left) and Monte Carlo (right) of the unconstraint $B_s^0 \to J/\psi\,\phi$ candidates. 98

9.9 The invariant mass distribution for the full data sample with the standard discriminating variables. 101

10.1 Invariant mass distribution for the $B^\pm \to J/\psi\,K^\pm$ decay for the data sample. 106

10.2 Invariant mass distribution for the $B^\pm \to \psi(2S)\,K^\pm$ decay for the data sample. 107

10.3 Invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ decay after the pre-selection for the data sample using a loose decay length significance cut of greater than four. 109

10.4 Invariant mass distribution for the $B_s^0 \to \psi(2S)\,\phi$ decay for the data sample using a loose decay length significance cut of greater than four. . 110

10.5 Invariant mass distribution for the $B_s^0 \to \psi(2S)\,\phi$ decay for the data sample and the optimised discriminating cuts with the RGS. 111

10.6 Output of the logistic regression for signal Monte Carlo and sideband data. 112

10.7 The invariant mass distribution for the full data sample using a linear discriminant. 113

10.8 Invariant mass distribution for the $B_s^0 \to J/\psi\,\phi$ decay for the data sample and the optimized discriminating cuts with the RGS. 114

11.1 Expected average upper limit at a 95% CL for the decay $B_s^0 \to \mu^+\mu^-$ as a function of integrated luminosity. 119

11.2 Expected average upper limit from the Tevatron for the decay $B_s^0 \to \mu^+\mu^-$ as a function of integrated luminosity. 120

11.3 Expected average upper limit at a 95% CL for the decay $B_s^0 \to \phi\,\mu^+\mu^-$ as a function of integrated luminosity. 121

B.1 Data and Monte Carlo events comparison for the leading Muon (left) and trailing Muon (right) for $B_s^0 \to J/\psi\,\phi$ candidates. 124

B.2 Data and Monte Carlo events comparison for the leading Kaon (left) and trailing Kaon (right) for $B_s^0 \to J/\psi\,\phi$ candidates. 125

B.3 Data and Monte Carlo events comparison for the J/ψ candidate (left) and ϕ candidate (right) for $B_s^0 \to J/\psi\,\phi$ candidate events. 125

B.4 Comparison of the discriminating variables for $B_s^0 \to J/\psi\,\phi$ Monte Carlo and background subtracted data events. 126

C.1 Depletion voltage from the Hamburg Model predictions and from laser measurements for various detectors. 128

C.2 Signal distributions. 131

C.3 Charge collection as a function of the applied bias voltage. 132

List of Tables

1.1 The spin-1/2 particles (fermions) of the Standard Model of particle physics. 4

1.2 The integral-spin particles (bosons) of the Standard Model of particle physics. 4

2.1 Standard Model values of the Wilson coefficients at NLO. 18

2.2 Standard Model branching fraction predictions for the decays $B \to X_s \ell^+\ell^-$ and experimental measurements for $m_{\ell^+\ell^-} > 0.2$ GeV$/c^2$. . . . 20

2.3 Branching fractions for $B \to K^{(*)} \ell^+\ell^-$ predicted in the framework of the Standard Model. 21

2.4 Measurements of the branching fractions for $B \to K^{(*)} \ell^+\ell^-$. 21

2.5 Standard Model branching fraction predictions for the decays $B^0_{s,d} \to \ell^+\ell^-$. 24

2.6 Experimental upper limits for the decays $B^0_{s,d} \to \ell^+\ell^-$ at a 90% CL. . . 24

3.1 Specifications of the silicon vertex detector, where "s,d" means single, double sided. 40

5.1 Run quality requirements. 57

8.1 Number of candidate events surviving the cuts in data used in the pre-selection analysis. 76

8.2 The different di-muon invariant mass regions for signal and sidebands used for background estimation. 77

8.3 The optimised set of cuts and their relative Monte Carlo signal efficiencies after maximising the figure of merit P. 78

8.4 Efficiencies for the discriminating variables of data and Monte Carlo events for the normalisation channel $B^\pm \to J/\psi\, K^\pm$. 83

8.5 Fitted mass resolutions as predicted by Monte Carlo simulations and as seen in the data for J/ψ and Υ(1S). 85

8.6 The sensitivity and total Monte Carlo signal efficiency for signal regions of different width. 86

8.7 The relative uncertainties for calculating an upper limit of $\mathcal{B}$ 86

8.8 The four signal events that were found after opening the box. 88

9.1 Number of candidate events in data surviving the cuts used in the preselection analysis. 92

9.2 The different four track invariant mass regions for signal and sidebands used for background estimation. 93

9.3 The optimised cuts and their relative Monte Carlo signal efficiencies including their statistical uncertainties after maximising P. 94

9.4 Data and Monte Carlo resolution comparison for the $B_s^0 \to J/\psi\,\phi$ candidates. 98

9.5 The relative uncertainties for calculating an upper limit of $\mathcal{B}$. Note that the first uncertainty is not taken into account since the limit has been normalised to $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$. 100

9.6 Comparison of the expected background and observed events after various cuts. 102

10.1 Comparison of properties of J/ψ and $\psi(2S)$ mesons (taken from [4]). . 104

10.2 Summary of fitting results of $B^\pm \to (J/\psi, \psi(2S))\,K^\pm$ final states for Monte Carlo events. 105

10.3 Summary of fitting results of $B^\pm \to (J/\psi, \psi(2S))\,K^\pm$ final states for the data. 106

10.4 Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for Monte Carlo events. 108

10.5 Summary of fitting results of $B_s^0 \to (J/\psi, \psi(2S))\,\phi$ final states for the data. 108

10.6 The different four track invariant mass regions for signal and sidebands used for background estimation. 108

10.7 Comparison of the signal efficiency with the Monte Carlo efficiency for $B_s^0 \to \psi(2S)\,\phi$ events. 110

10.8 Comparison of the signal efficiency with the Monte Carlo simulation efficiency for $B_s^0 \to J/\psi\,\phi$ events. 111

10.9 Efficiencies for the two decay channels (both channels CP-even only) after trigger and reconstruction with respect to the generated $b\bar{b}$-pairs. . . 113

10.10 The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B_s^0 \to \psi(2S)\,\phi/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ 115

10.11 The relative uncertainties for the measurement of the relative branching ratio $\mathcal{B}(B^\pm \to \psi(2S)\,K^\pm)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$ 116

10.12 The measured relative branching ratios from this analysis and published results. 117

C.1 SMT detector geometries and types: DS-double sided, SS-single sided, DSDM double sided double metal. The central barrels are Barrel 2 to 5, the outer barrels are 1 and 6. 129

C.2 Fit results 132

Bibliography

[1] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley Publishing Company, New York, 1995.

[2] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[3] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[4] S. Eidelman *et al.* (Particle Data Group Collaboration). Review of particle physics. *Phys. Lett.*, B592:1, 2004.

[5] L. Wolfenstein. Final-state interactions and CP violation in weak decays. *Phys. Rev.*, D43:151, 1991.

[6] H. Georgi and S. L. Glashow. Unity of all Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974.

[7] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton - hadron symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[8] A. J. Buras and R. Fleischer. Quark mixing, CP violation and rare decays after the top quark discovery. *Adv. Ser. Direct. High Energy Phys.*, 15:65–238, 1998.

[9] P. Gambino and M. Misiak. Quark mass effects in $\bar{B} \to X_s \gamma$. *Nucl. Phys.*, B611:338–366, 2001.

[10] T. E. Coan *et al.* (CLEO Collaboration). Study of exclusive radiative B meson decays. *Phys. Rev. Lett.*, 84:5283–5287, 2000.

[11] S. Chen *et al.* (CLEO Collaboration). Branching fraction and photon energy spectrum for $b \to s\gamma$. *Phys. Rev. Lett.*, 87:251807, 2001.

[12] R. Barate *et al.* (ALEPH Collaboration). A measurement of the inclusive $b \to s\gamma$ branching ratio. *Phys. Lett.*, B429:169–187, 1998.

[13] C. Jessop. A world average for $b \to s\,\gamma$. SLAC-PUB-9610, 2002.

[14] K. Abe (Belle Collaboration). Observation of $b \to d\,\gamma$ and determination of $|V(td)/V(ts)|$. *hep-ex/0506079*, 2005.

[15] A. J. Buras, M. Misiak, M. Munz, and S. Pokorski. Theoretical uncertainties and phenomenological aspects of $B \to X_s\,\gamma$ decay. *Nucl. Phys.*, B424:374–398, 1994.

[16] A. Ali, G. Hiller, L. T. Handoko, and T. Morozumi. Power corrections in the decay rate and distributions in $B \to X_s\,\ell^+\ell^-$ in the standard model. *Phys. Rev.*, D55:4105–4128, 1997.

[17] J. Kaneko *et al.* (Belle Collaboration). Measurement of the electroweak penguin process $B \to X_s\,\ell^+\ell^-$. *Phys. Rev. Lett.*, 90:021801, 2003.

[18] B. Aubert *et al.* (BaBar Collaboration). Measurement of the $B \to X_s\,\ell^+\ell^-$ branching fraction with a sum over exclusive modes. *Phys. Rev. Lett.*, 93:081802, 2004.

[19] M. Iwasaki *et al.* (Belle Collaboration). Improved measurement of the electroweak penguin process $B \to X_s\,\ell^+\ell^-$. *hep-ex/0503044*, 2005.

[20] A. Ghinculov, T. Hurth, G. Isidori, and Y. P. Yao. The rare decay $B \to X_s\,\ell^+\ell^-$ to NNLL precision for arbitrary dilepton invariant mass. *Nucl. Phys.*, B685:351–392, 2004.

[21] A. Ali, E. Lunghi, C. Greub, and G. Hiller. Improved model-independent analysis of semileptonic and radiative rare B decays. *Phys. Rev.*, D66:034002, 2002.

[22] A. Ali, P. Ball, L. T. Handoko, and G. Hiller. A comparative study of the decays $B \to (K, K^*)(l^+\,l^-)$ in standard model and supersymmetric theories. *Phys. Rev.*, D61:074024, 2000.

[23] D. Melikhov, N. Nikitin, and S. Simula. Rare decays $B \to (K, K^*)(l^+\,l^-,\,\nu\,\bar{\nu})$ in the quark model. *Phys. Lett.*, B410:290–298, 1997.

[24] K. Abe *et al.* (Belle Collaboration). Observation of the decay $B \to K\,\mu^+\mu^-$. *Phys. Rev. Lett.*, 88:021801, 2002.

[25] A. Ishikawa *et al.* (Belle Collaboration). Observation of the electroweak penguin decay $B \to K^*\,\ell^+\ell^-$. *Phys. Rev. Lett.*, 91:261601, 2003.

[26] B. Aubert *et al.* (BaBar Collaboration). Evidence for the rare decay $B \to K^*\,\ell^+\ell^-$ and measurement of the $B \to K\,\ell^+\ell^-$ branching fraction. *Phys. Rev. Lett.*, 91:221802, 2003.

[27] K. Abe *et al.* (Belle Collaboration). Measurement of the differential q^2 spectrum and forward–backward asymmetry for $B \to K(^*)l^+l^-$. *hep-ex/0410006*, 2004.

[28] B. Aubert *et al.* (BaBar Collaboration). Measurements of the rare decays $B \to K\,\ell^+\ell^-$ and $B \to K^*\,\ell^+\ell^-$. *hep-ex/0507005*, 2005.

[29] A. Ishikawa (Belle Collaboration). Measurement of forward-backward asymmetry and Wilson coefficients in $B \to K(^*)l^+l^-$. *hep-ex/0508009*, 2005.

[30] D. Acosta *et al.* (CDF Collaboration). Search for the decay $B_s^0 \to \phi\,\mu^+\mu^-$ in p anti-p collisions at $\sqrt{s}$ = 1.8 TeV. *Phys. Rev.*, D65:111101, 2002.

[31] B. Grzadkowski and P. Krawczyk. Higgs particle effects in flavor changing transitions. *Z. Phys.*, C18:43–45, 1983.

[32] P. Krawczyk. Flavor changing Yukawa coupling of the standard Higgs boson: Effects of the external quark masses. *Z. Phys.*, C44:509, 1989.

[33] G. Buchalla and A. J. Buras. QCD corrections to rare K and B decays for arbitrary top quark mass. *Nucl. Phys.*, B400:225–239, 1993.

[34] A. J. Buras. Relations between $\Delta(M_{s,d})$ and $B_{s,d} \to \mu\bar{\mu}$ in Models with Minimal Flavour Violation. *Phys. Lett.*, B566:115–119, 2003.

[35] B. Aubert *et al.* (BaBar Collaboration). Search for decays of $B^0 \to e^+e^-, B^0 \to \mu^+\mu^-, B^0 \to e^{+-}\mu^{-+}$. *Phys. Rev. Lett.*, 94:221803, 2005.

[36] M. Acciarri *et al.* (L3 Collaboration). Search for neutral b meson decays to two charged leptons. *Phys. Lett.*, B391:474–480, 1997.

[37] R. Bernhard *et al.* A Combination of CDF and DØ Limits on the Branching Ratio $B_{s,d}^0 \to \mu^+\mu^-$ Decays. *hep-ex/0508058*, 2005.

[38] C. Potter. A Search for the Rare Decay $B_d^0 \to \tau^+\tau^-$ at the BaBar Experiment. PhD thesis, University of Oregon, 2005.

[39] Y. Grossman, Z. Ligeti, and E. Nardi. $B \to \tau^+\tau^-(X)$ decays: First constraints and phenomenological implications. *Phys. Rev.*, D55:2768–2773, 1997.

[40] A. Dedes, H. K. Dreiner, U. Nierste, and P. Richardson. Trilepton events and $B_s^0 \to \mu^+\mu^-$: No-lose for mSUGRA at the Tevatron? FERMILAB-Pub-02/129-T, *hep-ph/0207026*, 2002.

[41] K. S. Babu and C. F. Kolda. Higgs-mediated $B_s^0 \to \mu^+\mu^-$ in minimal supersymmetry. *Phys. Rev. Lett.*, 84:228–231, 2000.

[42] A. Dedes and B. T. Huffman. Bounding the MSSM Higgs sector from above with the Tevatron's $B_s^0 \to \mu^+\mu^-$. *Phys. Lett.*, B600:261–269, 2004.

[43] H. E. Logan and U. Nierste. $B_{s,d}^0 \to \ell^+\ell^-$ in a two-Higgs-doublet model. *Nucl. Phys.*, B586:39–55, 2000.

[44] V. M. Abazov *et al.* (DØ Collaboration). A search for the flavor-changing neutral current decay $B_s^0 \to \mu^+\mu^-$ in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV with the DØ detector. *Phys. Rev. Lett.*, 94:071802, 2005.

[45] A. Abulencia et al. (CDF Collaboration). Search for $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ decays in $p\bar{p}$ collisions with CDF II. *Phys. Rev. Lett.*, 95:221805, 2005.

[46] D. N. Spergel *et al.* (WMAP Collaboration). First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *Astrophys. J. Suppl.*, 148:175, 2003.

[47] P. Belli, R. Cerulli, N. Fornengo, and S. Scopel. Effect of the galactic halo modeling on the DAMA/NaI annual modulation result: An extended analysis of the data for WIMPs with a purely spin-independent coupling. *Phys. Rev.*, D66:043503, 2002.

[48] D. S. Akerib *et al.* (CDMS Collaboration). First results from the cryogenic dark matter search in the soudan underground lab. *Phys. Rev. Lett.*, 93:211301, 2004.

[49] S. Baek, Y. G. Kim, and P. Ko. Neutralino dark matter scattering and $B_s^0 \to \mu^+\mu^-$ in SUSY models. *JHEP*, 02:067, 2005.

[50] A. Dedes, H. K. Dreiner, and U. Nierste. Correlation of $B_s^0 \to \mu^+\mu^-$ and (g-2)$_\mu$ in minimal supergravity. *Phys. Rev. Lett.*, 87:251804, 2001.

[51] G. W. Bennett *et al.* (Muon g-2 Collaboration). Measurement of the negative muon anomalous magnetic moment to 0.7-ppm. *Phys. Rev. Lett.*, 92:161802, 2004.

[52] A. Vainshtein. Hadronic effects in the muon anomalous magnetic moment. *Prog. Part. Nucl. Phys.*, 55:451–459, 2005.

[53] T. Blazek, R. Dermisek, and S. Raby. Predictions for Higgs and SUSY spectra from SO(10) Yukawa unification with $\mu > 0$. *Phys. Rev. Lett.*, 88:111804, 2002.

[54] T. Blazek, R. Dermisek, and S. Raby. Yukawa unification in SO(10). *Phys. Rev.*, D65:115004, 2002.

[55] R. Dermisek, S. Raby, L. Roszkowski, and R. Ruiz de Austri. Dark matter and $B_s^0 \to \mu^+\mu^-$ with minimal SO(10) soft SUSY breaking II. *JHEP*, 09:029, 2005.

[56] R. Arnowitt, B. Dutta, T. Kamon, and M. Tanaka. Detection of $B_s^0 \to \mu^+\mu^-$ at the Tevatron Run II and constraints on the SUSY parameter space. *Phys. Lett.*, B538:121–129, 2002.

[57] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali. The hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998.

[58] P. Dey and G. Bhattacharyya. Ultraviolet sensitivity of rare decays in nonuniversal extra dimensional models. *Phys. Rev.*, D69:076009, 2004.

[59] Z. Xiong and J. M. Yang. $B_{s,d}^0 \to \mu^+\mu^-$ in technicolor model with scalars. *Phys. Lett.*, B546:221–227, 2002.

[60] S. Abachi *et al.* (DØ Collaboration). The DØ detector. *NIM*, A338:185, 1994.

[61] S. Abachi *et al.* (DØ Collaboration). The DØ upgrade: The detector and its physics. Technical report, FERMILAB-PUB-96/357-E, 1996.

[62] V. M. Abazov *et al.* (DØ Collaboration). The Upgraded DØ detector. *submited to NIM, Fermilab-Pub-05-341-E, physics/0507191*, 2005.

[63] C. Miao *et al.* The DØ Run II luminosity monitor. *Nucl. Phys. Proc. Suppl.*, 78:342–347, 1999.

[64] T. Edwards *et al.* Determination of the effective inelastic $p\bar{p}$ cross-section for the DØ Run II luminosity measurement. 2004. FERMILAB-TM-2278-E.

[65] B. Asman *et al.* The DØ Silicon Microstrip Vertex Detector, to be submitted to Nucl. Instrum. Meth.

[66] E. Kajfasz *et al.* The DØ silicon microstrip tracker for Run IIa. *Nucl. Instrum. Meth.*, A511:16–19, 2003.

[67] J. Brzenziak *et al.* Conceptual design of a 2-Tesla superconducting solenoid for the Fermilab DØ detector upgrade. Technical report, FERMILAB-TM-1886, 1994.

[68] P. Baringer *et al.* Cosmic ray tests of the DØ preshower detector. *Nucl. Instrum. Meth.*, A469:295–310, 2001.

[69] M. Adams *et al.* Design Report for the Cental Preshower Detector for the DØ Upgrade. DØ-Note-3014, January 1996.

[70] A. Gordeev *et al.* Design Report for the Forward Preshower Detector for the DØ Upgrade. DØ-Note-3445, May 1998.

[71] L. Groers. DØ Calorimeter Upgrades for Tevatron Run II. DØ-Note-4240, October 2000. Proceedings for the IXth International Conference on Calorimetry in Particle Physics, Annecy, France, Oct 9-14, 2000.

[72] B. Abbott *et al.* (DØ Collaboration). Determination of the absolute jet energy scale in the DØ calorimeters. *Nucl. Instrum. Meth.*, A424:352–394, 1999.

[73] T. Diehl *et al.* Technical Design of the Cental Muon System. DØ-Note-3365, March 1998.

[74] G. Alexeev *et al.* Technical Design for the DØ Forward Muon Tracking Detector Based on Mini-Drift Tubes. DØ-Note-3366, December 1997.

[75] V. Abramov *et al.* Technical Design for the DØ Forward Trigger Scintillator Counters. DØ-Note-3237, November 1997.

[76] J. Abdallah *et al.* (DELPHI Collaboration). b-tagging in DELPHI at LEP. *Eur. Phys. J.*, C32:185–208, 2004.

[77] C. Clement *et al.* MuonID Certification for p14. DØ-Note-4350, February 2004.

[78] The DØ run quality database, `http://d0db-prd.fnal.gov/run/runquery.html`.

[79] T. Diehl V. M. Abazov and R. McCroskey. Good and Bad Muon Global Runs Early in Run II. DØ-Note 3938, April 2002.

[80] DØ trigger simulator (trigsim) homepage, `http://www-d0.fnal.gov/computing/trigsim/trigsim.html`.

[81] Pythia webpage, `http://www.thep.lu.se/~torbjorn/pythia.html`.

[82] The EvtGen package Homepage, `http://www.slac.stanford.edu/~lange/evtgen/`.

[83] H. L. Lai *et al.* (CTEQ Collaboration). Global QCD analysis of parton structure of the nucleon: Cteq5 parton distributions. *Eur. Phys. J.*, C12:375–392, 2000.

[84] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. GEANT 3 Users Guide. CERN-DD/EE/84-1.

[85] H. B. Prosper. Proceedings of the International Conference on Computing in High Energy Physics (CHEP'95), Rio de Janeiro, 18-22 September 1995.

[86] G. Punzi. Sensitivity of searches for new signals and its optimization. *ECONF*, C030908:MODT002, 2003.

[87] G. Zech. Upper limits in experiments with background or measurement errors. *Nucl. Instrum. Meth.*, A277:608, 1989.

[88] G. Feldman and R. Cousins. A Unified Approach to the Classical Satistical Analysis of Small Signals. *Phys. Rev. D*, 57:3873–3889, 1998.

[89] R. Cousins and V. L. Highland. Incorporating systematic uncertainties into an upper limit. *Nucl. Instrum. Meth.*, A320:331–335, 1992.

[90] C. Giunti and M. Laveder. The statistical and physical significance of confidence intervals. *hep-ex/0002020*, 2000.

[91] J. Conrad, O. Botner, A. Hallgren, and C. Perez de los Heros. Including systematic uncertainties in confidence interval construction for poisson statistics. *Phys. Rev.*, D67:012002, 2003.

[92] I. Bertram *et al.* A Recipe for the construction of confidence limits. FERMILAB-TM-2104, 2000.

[93] P. Ball and R. Zwicky. $B_{d,s} \to \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules revisited. *Phys. Rev.*, D71:014029, 2005.

[94] C. Q. Geng and C. C. Liu. Study of $B_s^0 \to (\eta, \eta', \phi)\, \ell\bar{\ell}$ decays. *J. Phys.*, G29:1103–1118, 2003.

[95] R. Bernhard and F. Lehner. Sensitivity Analysis of the rare decay $B_s^0 \to \phi\, \mu^+\mu^-$ with the DØ Detector. DØ-Note 4695, 31 January 2005.

[96] H. Albrecht *et al.* (ARGUS Collaboration). B Meson Decays into Charmonium States. *Phys. Lett.*, B199:451, 1987.

[97] F. Abe *et al.* (CDF Collaboration). Observation of $B^\pm \to \psi(2S)\, K^\pm$ and $B^0 \to \psi(2S)\, K^*(892)$ Decays and Measurements of B Meson Branching Fractions into J/ψ and $\psi(2S)$ final states. *Phys. Rev.*, D58:072001, 1998.

[98] S. J. Richichi *et al.* (CLEO Collaboration). Study of $B \to \psi(2S)\, K$ and $B \to \psi(2S)\, K^*(892)$ Decays. *Phys. Rev.*, D63:031103, 2001.

[99] D. Buskulic *et al.* (ALEPH Collaboration). First Measurement Of The B_s Meson Mass. *Phys.Lett.*, B308:425, 1993.

[100] B. Aubert *et al.* (BaBar Collaboration). Measurement of branching fractions for exclusive B decays to charmonium final states. *Phys. Rev.*, D65:032001, 2002.

[101] S. Marti i Garcia *et al.* A model of charge collection for irradiated p+n detectors. *Nucl. Instrum. Meth.*, A473:128–135, 2001.

[102] R. Lipton. Lifetime of the DØ Silicon Tracker. DØ-Note-4077, January 2003.

[103] J. Gardner *et al.* Results from Irradiation Tests on DØ Run II Silicon Detectors at the Radiation Damage Facility at Fermilab. DØ-Note-3962, March 2002.

[104] M. Moll *et al.* Relation between microscopic defects and macroscopic changes in silicon properties after hadron irradiation. *Nucl. Instrum. Meth.*, A 186:100–1102, 2002.

[105] S.M. Sze. *Semiconductor Devices, Physics and Technology*. John Wiley and Sons, 1985.

[106] H. Bichsel. Straggling in thin silicon detectors. *Reviews of Modern Physics*, 60(3):663–699, 1988.

Acknowledgments

This page is dedicated to all the people who made my dissertation possible, who supported and helped me arriving finally and happily at this last page.

First of all I am very thankful to Ueli Straumann for offering me the unique opportunity to work in his research group as a PhD student. He allowed me to pursue the research topic that interested me most, which meant to join the DØ collaboration and to work at FERMILAB.

I would like to express my gratitude to Frank Lehner for the initiation to the DØ experiment and for his continuous support over the past three years. It has been a real pleasure to work with you and I am looking forward to common projects in the future with you. I owe you also a deep debt for the steady proofreading of all the analyses notes and this thesis.

I am also indebted to Prof. Rick Van Kooten who has kindly agreed to be referee of this thesis.

The status of this analyses have been reported frequently to the B-Physics Group and have been reviewed by the Run II Editorial Board 23. Thanks to the members for their continuous help and support, especially to the chair of the Editorial Board Prof. Andrzej Zieminski.

This work would not have been possible without the work of a lot of people at the DØ experiment, the Tevatron collider and FERMILAB. I would like to thank everybody for their tireless efforts for an excellent fundamental research.

A special thanks goes to all my colleagues from Zürich and Chicago for the pleasant working atmosphere and all the fruitfull discussions. Especially I want to thank Olaf for proof-reading, Stefania and Johannes for the friendly atmosphere in the office and Matt for the nice distraction during our climbing trips.

Finally, I want to thank my parents, my sister and my grandmother for their tireless encouragement and support in all areas over all the years.

This work has been supported by the Swiss National Science Foundation and the University Research Fund.

Curriculum Vitae

Personal information

First Name:	Ralf
Middle Name:	Patrick
Last Name:	Bernhard

Education

2002–2005	**Dissertation** in Physics, University of Zurich, Switzerland *Search for Rare Decays of the B_s^0 Meson* *with the DØ Experiment* Supervisor: Prof. Dr. Ulrich Straumann
2002	**Physics Diploma** at University of Siegen, Germany
2001–2002	Diploma Thesis, University of Siegen, Germany *Eine neue Bestimmung der Formfaktoren im $K_L \to \pi\mu\nu$-Zerfall* *mit dem NA48-Detektor* Supervisor: Prof. Dr. Martin Holder
1999–2000	Exchange student at the ”Universita degli Studi di Pisa”, Italy
1996–1999	Study of Physics, University of Siegen, Germany
1995–1996	Fulltime vocational school to qualify for University, Bühl, Germany
1992–1995	Professional training with the company “R. BOSCH AG” to become an electronics engineer with specialization on apparatus technology