Search for $B_s^0 - \overline{B_s^0}$ Oscillations with the DELPHI Detector

Anders Westreng Borgland

Thesis submitted in partial fulfillment of requirements for the degree of Doctor Scientarium at the Faculty of Mathematics and Natural Sciences, University of Bergen.

June 1999

Search for $B_s^0 - \overline{B_s^0}$ Oscillations with the DELPHI Detector

Anders Westreng Borgland

Thesis submitted in partial fulfillment of requirements for the degree of Doctor Scientarium at the Faculty of Mathematics and Natural Sciences, University of Bergen.

June 1999

Abstract

Using four million hadronic Z^0 decays recorded by the DELPHI detector at LEP between 1991 and 1995, B_s^0 mesons have been reconstructed in their semileptonic decays $D_s^{\mp} l^{\pm}$. A high transverse momentum lepton was combined with an oppositely charged D_s meson in the decays $D_s^- \longrightarrow \phi\pi^-$, $K^{*0}K^-$, $\phi\pi^+\pi^-\pi^-$, $\phi\pi^-\pi^0$, $K^{*0}K^{*-}$, $\phi e^-\bar{\nu}_e$ and $\phi\mu^-\bar{\nu}_\mu$. In a total of 392 reconstructed candidates, 197 were found to be B_s^0 signal events.

The B_s^0 lifetime was measured to be

$$\tau(B_s^0) = 1.48\,^{+0.16}_{-0.14}\ \text{(stat) ps}\ ^{+0.06}_{-0.06}\ \text{(syst) ps}$$

An upper limit on the lifetime difference between the two physical B_s^0 states was found to be

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} \leq 56\% \quad @\ 95\% \ \text{C.L.}$$

This limit takes into account statistical uncertainties only.

Using the amplitude method, the following regions of the mass difference between the physical B_s^0 states are excluded:

$$\begin{aligned} 0 &\leq \Delta m_{B_s^0} \leq 4.4\,\text{ps}^{-1} \quad @\ 95\%\ \text{C.L.} \\ 6.0\,\text{ps}^{-1} &\leq \Delta m_{B_s^0} \leq 7.6\,\text{ps}^{-1} \quad @\ 95\%\ \text{C.L.} \end{aligned}$$

This includes both statistical and systematic uncertainties. The sensitivity was found to be

$$\Delta m_{B_s^0} = 6.7\,\text{ps}^{-1}$$

and the exclusion probability at $\Delta m_{B_s^0} = 7.6\,\text{ps}^{-1}$ is 45%.

Acknowledgments

This thesis is based on data taken with the DELPHI detector and I am therefore greatly indebted to all the DELPHI collaborators who over the years planned, built and improved the detector. My work has taken place in the context of the *B lifetimes / oscillation team* and I have benefited from discussions with, and criticism from, its members. In particular, I would like to thank Fabrizio Parodi for all his help and advice.

Ole M. Røhne was invaluable with his help with practical details in the beginning and also deserves praise for porting Delsim to Linux. I also received critical computer help on numerous occasions from Trond Myklebust, system manager *bis*.

During these years I have greatly benefited from discussions with both staff and other students in the Bergen HEP group, in particular Bjarne Stugu and Radek Skoda. I have also learned more about other topics in particle physics in discussions with Kristine Johansen and Katarina Pajchel. These last three also shared my working hours, making long working evenings more pleasant.

My thesis supervisor has been professor Gerald Eigen. His enthusiasm, advice and criticism have been a source of inspiration. I am especially thankful for the freedom I have enjoyed in my work. Thus, I have discovered all by myself that most *alleys* belong to the category *blind*.

I'll always have a special thought for Marine "qui a vécu un peu cette thèse aussi".

That workshops can have unpredictable consequences is my meeting with Daniela a proof of. Her belief in me and her constant encouragement helped me beyond words. I am glad I got the chance to meet you and I will miss you dearly.

That the end of a thesis can be filled with joy and not only work and stress is entirely due to Ingrid. Tre ting på en gang går ann.

My parents have always supported and helped me, not only during this thesis but throughout my studies. I would never have reached this point without them.

This thesis has been supported by a grant from the Norwegian Research Council.

The answer must be in the attempt.

Contents

1 Introduction **1**

2 Theoretical Motivation **5**

2.1 Why Quantum Field Theories? 5
2.2 Symmetries 6
2.2.1 The Lagrangian 6
2.2.2 Symmetries and Conservation Laws 7
2.2.3 Gauge Theories and Gauge Bosons 7
2.3 The Standard Model 10
2.3.1 General Description 10
2.3.2 Quantum ChromoDynamics 10
2.3.3 The Electroweak Sector 10
2.3.4 Masses in the Standard Model 13
2.3.5 The CKM Matrix 19
2.4 *CP* Violation 23
2.4.1 Discrete symmetries 23
2.4.2 The Kaon System 25
2.4.3 *CP* Violation and the CKM Matrix 27
2.5 $B_q^0 - \overline{B_q^0}$ Mixing Formalism 29
2.5.1 Mass and Decay Matrices 31
2.5.2 The Physical States 32
2.5.3 Time Evolution of B_q^0 Mesons 33
2.5.4 Lifetime Differences in the B-System 35
2.6 Theoretical Calculation of $\Delta m_{B_q^0}$ 36
2.7 Numerical Values 37
2.8 Constraints on the Unitarity Triangle 38

3 The DELPHI Experiment **40**

3.1 Detector Description 40
3.1.1 Tracking Detectors 42
3.1.2 Momentum Resolution 43
3.2 Particle Identification 43
3.2.1 Electron Identification 43
3.2.2 Muon Identification 47
3.2.3 Charged Hadron Identification 47

4 Experimental Motivation **55**
4.1 Experimental Constraints 55
4.2 B_s^0 Decay Modes and Possible Analyses 55
4.3 Kinematic Analysis Language 57

5 B_s^0 Candidate Selection **59**
5.1 Hadronic Event Selection and Track Quality Criteria 59
5.2 Data and Simulation Samples 61
5.3 Event Topology 61
5.4 D_s^- Selection 62
5.4.1 General Remarks 62
5.4.2 D_s^- Decay Modes 63
5.4.3 $D_s^- \longrightarrow \phi\pi^-$ 64
5.4.4 $D_s^- \longrightarrow K^{*0}K^-$ 65
5.4.5 $D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$ 67
5.4.6 $D_s^- \longrightarrow \phi\pi^-\pi^0$ 67
5.4.7 $D_s^- \longrightarrow K^{*0}K^{*-}$ 68
5.4.8 $D_s^- \longrightarrow \phi l^-\bar{\nu}_l$ 68
5.4.9 D_s^- Selection Criteria for All Decay Modes 69
5.5 B_s^0 Selection 70
5.6 Method for Choosing the Selection Criteria 70
5.7 Signal Acceptance Function 72
5.8 Sample Composition 72
5.8.1 Signal 72
5.8.2 Like-sign Candidates 74
5.8.3 *Cascade* Decays 82
5.8.4 Kinematic *Reflections* 83
5.9 Combinatorial Background Composition 84

6 Proper Time Reconstruction **86**
6.1 General Remarks 86
6.2 Momentum Resolution 87
6.2.1 Neutrino Momentum Estimate 87
6.2.2 B_s^0 Momentum Estimate 92
6.3 Decay Length Resolution 99
6.3.1 Decay Length Estimate 99
6.3.2 Improvement of the Decay Length Resolution 105
6.4 Proper Time Resolution 107
6.4.1 Time Dependent Proper Time Resolution 107
6.4.2 Time Integrated Proper Time Resolution 112
6.4.3 Proper Time Bias 118

7 B_s^0 Lifetime **119**
7.1 General Remarks 119
7.2 The Likelihood Function 119
7.3 Results 122

7.4 Systematic Uncertainties 124

8 B_s^0 Lifetime Difference **133**

9 $B_s^0 - \overline{B_s^0}$ Oscillations **137**
9.1 General Remarks 137
9.2 $b(\bar{b})$ Tagging at Production Time 137
9.2.1 Tagging Variables Description 137
9.2.2 Tagging Purity 138
9.3 The Likelihood Function 141
9.4 The Likelihood Function in the Case of No Sensitivity 142
9.4.1 Likelihood Referenced to the Minimum 144
9.4.2 Likelihood Referenced to Infinity 148
9.5 The Amplitude Method 148
9.5.1 The Amplitude 149
9.5.2 Exclusion Probability 149
9.5.3 Sensitivity 150
9.5.4 The Amplitude Method in Case of a Measurement 151
9.5.5 The Amplitude Method in Case of a Limit 151
9.6 Results 151
9.7 Systematic Uncertainties 151
9.8 The Asymmetry $N^{\text{no osc}} - N^{\text{osc}}$ 155

10 Conclusion **160**

A KAL Script **168**

B Likelihood Normalization **170**

C Conference Contributions **173**

List of Figures

1.1 Standard Model of elementary particle physics 2

2.1 Spontaneous symmetry breaking: The buckling needle 14
2.2 Higgs potential 16
2.3 The Unitarity Triangle 28
2.4 Box diagrams 30
2.5 Oscillating and non-oscillating components 39

3.1 Schematic layout of the DELPHI detector 41
3.2 Track parameter precision 44
3.3 Electron identification efficiency and purity 46
3.4 Muon identification efficiency 48
3.5 Cherenkov radiation in the RICH detector 49
3.6 DELPHI particle identification - RICH 50
3.7 Performance of the *tight* RICH tag 53
3.8 DELPHI particle identification - dE/dX 54

4.1 Time integrated mixing probability 56
4.2 $B_s^0 \longrightarrow D_s^- l^+ \nu_l$ Feynman diagram 57

5.1 Signal event topology 62
5.2 K^+K^- invariant mass distribution for the the $\phi\pi$ mode - 1994-1995 data set 66
5.3 K^+K^- invariant mass distribution for the the $\phi\pi$ mode - 1991-1993 data set 66
5.4 Lepton transverse momentum for the signal and the *cascade* background . . . 71
5.5 Generated $D_s^- l^+$ invariant mass for the the signal and the *cascade* background 71
5.6 Signal acceptance 73
5.7 Invariant mass distributions for $D_s^\mp$ and ϕ candidates for the entire data set . 75
5.8 Invariant mass distributions for $D_s^\mp$ candidates in the $\phi\pi^\mp$ and $K^{*0}K^\mp$ modes for the 1994-1995 data set 76
5.9 Invariant mass distributions for $D_s^\mp$ candidates in the $\phi\pi^\pm\pi^\mp\pi^\mp$ and $K^{*0}K^{*\mp}$ modes for the 1994-1995 data set 77
5.10 Invariant mass distributions for $D_s^\mp$ and ϕ candidates in the $\phi\pi^\mp\pi^0$ and $\phi l^\mp\bar{\nu}_l$ modes for the 1994-1995 data set 78
5.11 Invariant mass distributions for $D_s^\mp$ and ϕ candidates for the 1991-1993 data set 79
5.12 $B_d^0 \longrightarrow D^- D_s^+$ Feynman diagram 82
5.13 Reconstructed $K\pi\pi$ mass from the kinematic *reflection* decays 85

6.1 B_s^0 decay length resolution versus the B_s^0 decay length 87
6.2 Correction factor to the beam energy and number of jets in signal events . . . 89
6.3 Neutrino energy versus the $D_s^- l^+$ energy 90
6.4 Neutrino energy resolution 91
6.5 Generated energy for photons from the decay $D_s^{*-} \longrightarrow D_s^- \gamma$ 93
6.6 Relative B_s^0 momentum resolution 94
6.7 Relative $D_s^{(*)-}$ momentum resolution 95
6.8 Reconstructed D_s^- and D_s^{*-} mass as expected from Monte Carlo simulations 96
6.9 Signal momentum distributions for data and Monte Carlo simulations 97
6.10 Relative B_s^0 momentum resolution versus the B_s^0 momentum 98
6.11 Interaction region coordinates spread for the 1994-1995 data set 100
6.12 Primary vertex coordinates resolution for the 1994-1995 data set. 101
6.13 Primary vertex coordinates resolution for the 1991-1993 data set. 102
6.14 Generated angle between the B_s^0 and the $D_s^- l^+$ 103
6.15 B_s^0 decay length resolution 104
6.16 Primary vertex resolution (projected onto the $D_s^- l^+$ direction) 106
6.17 B_s^0 decay length resolution versus the angle between the lepton and the D_s^- . 108
6.18 Generated angle between the lepton and the D_s^- 109
6.19 B_s^0 reconstruction efficiency versus the angle between the lepton and the D_s^- 110
6.20 Generated B_s^0 momentum 111
6.21 B_s^0 proper time resolution versus the proper time 113
6.22 B_s^0 proper time resolution in different bins of $t_{\text{generated}}$ 114
6.23 B_s^0 proper time resolution 116
6.24 B_s^0 proper time resolution for cos $(\theta_{l^+ - D_s^-}) < 0.95$ 117

7.1 B_s^0 decay length for the entire data sample 126
7.2 B_s^0 proper time distribution for the entire data sample 127
7.3 B_s^0 proper time distribution for each decay mode for the entire data sample . 128
7.4 B_s^0 lifetime likelihood function 129
7.5 Fitted B_s^0 lifetime for each decay mode for the entire data set 130
7.6 Combinatorial background proper time distributions for each decay mode for the 1994-1995 data set 131
7.7 Combinatorial background proper time distributions for each decay mode for the 1991-1993 data set 132

8.1 Likelihood function for $\Delta\Gamma(B_s^0)/\Gamma(B_s^0)$ 135
8.2 Probability density function for $\Delta\Gamma(B_s^0)/\Gamma(B_s^0)$ 136

9.1 Distribution of the initial quark state identity tagging variable 139
9.2 Distribution of the jet charge 140
9.3 Observed $\Delta m_{B_s^0}$ likelihood function 143
9.4 Toy Monte Carlo simulations - Sensitive to the true value of $\Delta m_{B_s^0}$ 145
9.5 Toy Monte Carlo simulations - At the limit of sensitivity to the true value of $\Delta m_{B_s^0}$ 146
9.6 Toy Monte Carlo simulations - No sensitivity to the true value of $\Delta m_{B_s^0}$ . . . 147
9.7 Amplitude plot for a simulated sample of 1000 B_s^0 signal events 152

9.8 Likelihood plot for a simulated sample of 1000 B_s^0 signal events 152
9.9 Amplitude plot for a simulated sample of 250 B_s^0 signal events 153
9.10 Likelihood plot for a simulated sample of 250 B_s^0 signal events 153
9.11 Amplitude plot for a simulated sample of 100 B_s^0 signal events 154
9.12 Likelihood plot for simulated sample of 100 B_s^0 signal events 154
9.13 Amplitude plot for the entire data sample 156
9.14 Uncertainty on the oscillation amplitude for the entire data sample 157
9.15 Exclusion probability for the entire data sample 158
9.16 The asymmetry $(N^{\text{no osc}} - N^{\text{osc}})$ for the entire data sample 159

10.1 Different B_s^0 lifetime measurements 161
10.2 Sensitivity and excluded regions of $\Delta m_{B_s^0}$ for different analyses 162

List of Tables

3.1 Electron identification and misidentification probabilities . . . 45
3.2 Muon identification and misidentification probabilities . . . 47
3.3 Barrel RICH characteristics . . . 51

5.1 Number of hadronic Z^0 decays recorded by the DELPHI detector during LEP1 61
5.2 D_s^- selection criteria . . . 69
5.3 D_s^- kinematic selection criteria . . . 69
5.4 Number of $D_s^{\mp}l^{\pm}$ signal events for the entire data set . . . 80
5.5 Number of $D_s^{\mp}l^{\pm}$ signal events the in 1994-1995 data set . . . 80
5.6 Number of $D_s^{\mp}l^{\pm}$ signal events in the 1991-1993 data set . . . 80
5.7 Expected $D_s(\phi)$ mass resolutions, signal reconstruction efficiencies and expected number of signal events in the 1994-1995 data set . . . 81
5.8 Expected $D_s(\phi)$ mass resolution, signal reconstruction efficiencies and expected number of signal events in the 1991-1993 data set . . . 81
5.9 Fractions of electrons and muons from the B_s^0 candidate in the two data sets 81
5.10 $\overline{B_d^0}$ and B^- reconstruction efficiencies for the hadronic modes . . . 83
5.11 $\overline{B_d^0}$ and B^- reconstruction efficiencies for the semileptonic modes . . . 83
5.12 *Reflection* decay reconstruction efficiencies . . . 84

6.1 Linear fit constants for the estimate of the neutrino energy . . . 88
6.2 Neutrino energy resolution . . . 92
6.3 Relative B_s^0 momentum resolution for the 1994-1995 data set . . . 99
6.4 Relative B_s^0 momentum resolution for the 1991-1993 data set . . . 99
6.5 B_s^0 decay length resolution for the 1994-1995 data set . . . 105
6.6 B_s^0 decay length resolution for the 1991-1993 data set . . . 105
6.7 B_s^0 proper time resolution for the 1994-1995 data set . . . 112
6.8 B_s^0 proper time resolution for the 1991-1993 data set . . . 112
6.9 B_s^0 time integrated proper time resolution for the 1994-1995 data set . . . 115
6.10 B_s^0 time-integrated proper time resolution for the 1991-1993 data set . . . 115
6.11 B_s^0 proper time resolution . . . 115
6.12 Reconstruction biases . . . 118

7.1 Sample composition . . . 120
7.2 Combinatorial background proper time parameterization for the 1991-1993 data set . . . 122

7.3 Combinatorial background proper time parameterization for the 1994-1995 data set 122
7.4 Effective lifetimes for B_d^0 and B^- *cascade* decays - hadronic modes 123
7.5 Effective lifetimes for B_d^0 *cascade* decays - semileptonic modes 123
7.6 Separately fitted B_s^0 lifetime for each decay mode for the entire data sample . 123
7.7 Systematic uncertainties on the B_s^0 lifetime 124

9.1 Variations of the sensitivity 150

Chapter 1

Introduction

The goal of particle physics is the understanding of the ultimate building blocks of Nature and the forces that exist between them. Even though the *ultimate* building blocks have a certain habit of revealing still another layer underneath them, we are today able to claim that we understand Nature up to energies of 100 GeV, or, alternatively, down to 10^{-18}m. The model that is used has proven itself so successful that it is now just refered to as the *Standard Model*. Even though it has been tested on numerous occasions, its predictions have never been contradicted by experimental measurements.

The Standard Model is a quantum field theory based on *local gauge symmetries*. As can be seen in figure 1.1, particles are divided into two types. *Fermions* (leptons and quarks) are particles with spin 1/2. They come in three families and make up matter. *Bosons* are spin 1 particles and mediate the interactions. We know today of four interactions:

- *Gravitation* is much too weak at this level to have any importance and is not incorporated into the Standard Model. It is mediated by a massless spin-2 boson called the *graviton*. To make a truly unified theory of all interactions, gravitation must be included. This means, as a first step, unifying gravitation and quantum mechanics, a step which has been proven difficult to achieve.

- The *strong* interaction is felt only by quarks and is responsible for keeping the nucleus together. The corresponding charge is called *color*, thus giving the interaction its name: Quantum ChromoDynamics. The interaction is mediated by eight massless *gluons*.

- The *weak* interaction is felt by both quarks and leptons and is responsible for radioactive decays. It is mediated by the massive Z^0 and $W^{\pm}$ bosons.

- The *electromagnetic* interaction is mediated by the *photon*. The corresponding charge is the usual electric charge.

A startling success of the Standard Model has been its ability to unify interactions. Just like Maxwell showed that electricity and magnetism are two different manifestations of the same electromagnetic force, the electromagnetic force is in the Standard Model unified with the weak interaction. One consequence of this unification is the appearance of three massive vector bosons: Z^0 and $W^{\pm}$. Their existence and masses were predicted by the model and their subsequent discovery showed that the model rests on sound foundations.

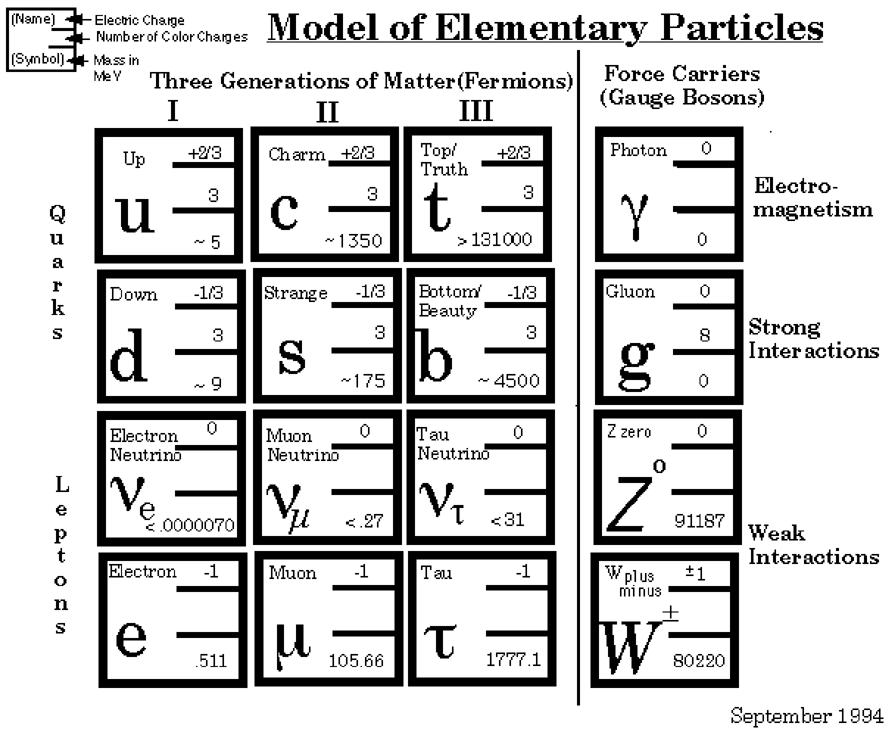


Figure 1.1: *Standard Model of elementary particle physics: Particles are divided into three families of spin 1/2 matter particles, and spin 1 bosons mediating the interactions. The top quark, with a mass of* (173.8 ± 5.2) GeV/c^2 *[1], was the last to be discovered.*

The fact that these bosons and all the fermions, with the possible exception of the neutrinos, are massive, however, is not easily accounted for. Naively introducing mass terms in the model breaks the gauge symmetry. It is, therefore, necessary to give particles mass via some other mechanism. The one most in favor today is called *spontaneous symmetry breaking* or the Higgs mechanism. It predicts the existence of a new boson called the Higgs boson. Unfortunately, it does not predict its mass and it has not yet been seen experimentally.

While the Higgs mechanism is successful in giving mass to the different particles, it has also the consequence of breaking the correspondence between the flavour eigenstates and weak eigenstates of the quarks. The corresponding rotation matrix is called the Cabibbo Kobayashi Maskawa (CKM) matrix. The fact that there are at least three quark families means that the matrix contains a complex phase. One consequence of this is the appearance of *CP* violation, which is related to the asymmetry between matter and antimatter in the universe.

Measuring the elements of the CKM matrix is, therefore, necessary to shed light on two of the least explained phenomena of the Standard Model. One of the most important CKM elements is the one giving the strength of the coupling between the top and down quarks, $|V_{td}|$. This element can, in principle, be extracted from the measurement of the mass difference $\Delta m_{B_d^0}$ between the physical B_d^0 states. Unfortunately, theoretical difficulties related to non-perturbative QCD effects make this difficult. It turns out that taking the ratio $\Delta m_{B_d^0}/\Delta m_{B_s^0}$ removes many of these problems and makes it easier to extract $|V_{td}|$. This means that even though $\Delta m_{B_d^0}$ has been precisely measured, it is not a very useful quantity before one is able to measure $\Delta m_{B_s^0}$. Related to the mass splitting is the lifetime difference between the two weak eigenstates $B^0_{s\,Light}$ and $B^0_{s\,Heavy}$.

Experimentally, B_s^0 mesons can be studied at both $p\bar{p}$ and e^+e^- colliders. By going to high energies the B_s^0 will be boosted and the typical distance between the B decay vertex and the primary interaction point will be several mm. With modern vertex detectors this decay vertex can be reconstructed. Hadron machines benefits from a high $b\bar{b}$ cross section, but suffers from a high track multiplicity in the final state. e^+e^- machines, however, presents a relative clean environment, thus facilitating the event reconstruction.

From 1989 to 1995 the Large Electron Positron (LEP) collider at CERN operated at the Z^0 resonance, resulting in over four million hadronic Z^0 being recorded by each of the four experiments DELPHI, OPAL, L3 and ALEPH.

The topic of this thesis is the possible measurement of the mass splitting and lifetime difference between the two weak eigenstates in the $B_s^0 - \overline{B_s^0}$ system, in addition to the measurement of the B_s^0 lifetime. It is based on data taken with the DELPHI detector between 1991 and 1995 and represents most of the data recorded at the Z^0 resonance.

The outline of the thesis is the following: Chapter 2 gives a description of the electroweak sector of the Standard Model, with special emphasis on explaining the relevant parts conceptually. Thus, it is shown how the CKM matrix appears as a necessary consequence of the way particles are given mass in the Standard Model. Then the connection to the neutral B meson system is given, and all the relevant experimental observables are given in detail. In this section units in which $\hbar = c = 1$ are used.

Chapter 3 gives a brief description of the relevant parts of the DELPHI detector and the particle identification algorithms. A motivation for the chosen analysis, along with a description of the data analysis tool that was adapted to DELPHI for this analysis, are then given in chapter 4. Chapter 5 gives a detailed description of the different selection criteria used to define the B_s^0 event sample. The proper time reconstruction is detailed in chapter 6.

Using this sample the B_s^0 lifetime is measured (chapter 7). Because of the small uncertainty, this is an interesting measurement in itself. However, since the B_s^0 lifetime already is known, it also provides us with a useful check of the event sample. Chapters 8 and 9 deal with extracting and setting a limit on the B_s^0 decay width difference $\Delta\Gamma_{B_s^0}$ and mass difference $\Delta m_{B_s^0}$.

Door meten tot weten.

Chapter 2

Theoretical Motivation

Symmetries play an important role in modern theoretical physics. It can be used to simplify calculations and sometimes to show the underlying order in a zoo of particles by revealing the true building blocks of Nature. But symmetries are not only a useful 'technicality', they can also be linked to physical phenomena. One consequence of symmetries is conserved quantities. Invariance under translation in space means that momentum is conserved. Similarly, energy is conserved if the system is invariant in time. There exists also another form of symmetry which is essential in particle physics: local gauge symmetries. These have the peculiar effect of introducing interactions in an otherwise non-interacting theory, i.e. the dynamics itself results from a symmetry.

The Standard Model of particle physics is a local gauge theory. Why it is so is not clear. It just seems to be a fact that Nature, for unknown reasons, decided to obey this symmetry. This means that it is only its experimental verification that tells us that the model is really describing Nature. These verifications have never put the model in doubt. In fact, with the recent colliders like LEP at CERN, the Standard Model has been tested and verified to an impressive precision.

This chapter is a motivation for the experimental analysis which is the main part of the thesis. It will, therefore, only give a description of the parts of the Standard Model which are relevant in this context. The emphasis is not put on the formal point of view. The purpose is more to explain the basic concepts and reveal the underlying physics in a simple as possible manner. Because of this, many calculations have been simplified and many technical details are not mentioned.

2.1 Why Quantum Field Theories?

In non-relativistic quantum mechanics one treats the quantization of single particles. Since particle physics is concerned with processes that occur at very small scales and at high energies, one may ask if not relativistic quantum mechanics can give an adequate description. In fact, the answer is negative and we are obliged to introduce a quantum field point of view. The main point is that at high energies we do not have a single-particle theory anymore. Because of Einstein's famous relation between energy and mass, $E = m\,c^2$, the theory must allow for particle-antiparticle pair creation if the energy is high enough. In addition to this, the Heisenberg uncertainty relation between energy and time, $\Delta E \cdot \Delta t = \hbar$, allows for the creation of *virtual* particles.

2.2 Symmetries

"Symmetries", to quote one theorist, *"are connected to conserved quantities, and without conserved quantities we just don't know how to calculate anything"*. One particular useful quantity in this context is the *Lagrangian*.

2.2.1 The Lagrangian

In any particular theory we are usually interested in the equations of motion. One way to get these is to introduce a function called the Lagrangian:

$$L = \text{Kinetic energy} - \text{Potential energy} = T - V \tag{2.1}$$

and a procedure to obtain the equations of motion from this function. In classical mechanics the fundamental quantity is the *action* S which is the time integral of the Lagrangian. In a field theory we write

$$S = \int L\, dt = \int \mathcal{L}(\phi, \partial_\mu \phi)\; d^4x \tag{2.2}$$

where the spatial Lagrangian density $\mathcal{L}$ is a function of the field(s) ϕ and its derivative and x is the space-time four vector. By using the *principle of least action,* which says the system evolves from time t_1 to t_2 along the path in configuration space for which S is an extremum (usually a minimum),

$$\delta S = 0 \tag{2.3}$$

one finds that the Lagrangian density must fulfill the Euler-Lagrange equation of motion for the field:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{2.4}$$

There will be one such equation for each field.

As long as we can write down the Lagrangian[1] for a system, this procedure will automatically give us the corresponding equations of motion. For example, to obtain Newton's Second law of motion in classical mechanics,

$$\vec{F} = m\frac{d\vec{v}}{dt} \tag{2.5}$$

where m is the mass and $\vec{v}$ the velocity of the particle, one would use

$$L = \frac{1}{2}mv^2 - V \tag{2.6}$$

[1]To be precise, one should in fact say *a Lagrangian* since we can always add a 4-divergence to it without affecting the Euler-Lagrange equations of motion.

and

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial r} = 0 \tag{2.7}$$

assuming that the velocity is not a function of the distance r, and that the force can be derived from a scalar potential V with $\vec{F} = -\nabla V$.

2.2.2 Symmetries and Conservation Laws

The main advantage with the Lagrangian approach is that all the symmetry-properties are contained in the Lagrangian density $\mathcal{L}$, and that they provide us with an easy bridge to conservation laws.

Let us consider a continuous variation of the field ϕ. Since any finite transformation can be built up of through repeated use of infinitesimal transformations, only the latter is considered:

$$\phi \longrightarrow \phi^{'}(x) = \phi(x) + \alpha\Delta\phi(x) \tag{2.8}$$

where α is an arbitrary small parameter.

This transformations is called a *symmetry* if it leaves the equations of motion invariant. When the above expression is put into the equations of motion it leads to

$$\partial_\mu j^\mu(x) = 0 \tag{2.9}$$

where

$$j^\mu(x) = \frac{\partial_\mu \mathcal{L}}{\partial(\partial_\mu \phi)}\Delta\phi \tag{2.10}$$

i.e. the symmetry leads directly to a *conserved current* $j^\mu(x)$. This can be rewritten by saying that the *charge*

$$Q = \int_{\text{All space}} j^0 d^3x \tag{2.11}$$

is a constant in time. That the invariance of the Lagrangian density $\mathcal{L}$ under a continuous one-parameter transformation implies a conserved quantity is called *Noether's theorem [2].*

From this theorem it follows that the invariance of the Lagrangian density under translations with respect to time and space leads to conservation of energy and momentum, respectively, while invariance under rotations leads to conservation of angular momentum.

2.2.3 Gauge Theories and Gauge Bosons

In addition to the invariance of the Lagrangian density under continuous space-time transformations, there are also invariances under a more abstract type of transformations: *phase transformations.* These will lead to gauge theories and the necessity of introducing gauge bosons.

Classical Electrodynamics

In classical electrodynamics the electromagnetic field can be written as a function of a scalar and a vector potential. Since the potentials are not observables, we are free to add terms to then *as long as the fields themselves are not affected.* This is called *gauge invariance.* We treat the potentials as mathematical objects that we must perform operations on to get the physical fields. This gives us an additional liberty which is used in classical electrodynamics to simplify the calculations by a clever choice of gauge. The gauge invariance also has dynamical properties.

Gauge Invariance

It is not possible to measure an absolute potential, only potential differences are accessible. This means that a constant term can be added to the potential without any observational consequences. Two possibilities exist.

The same *global* term can be added to the potential at all points in space. This is somewhat unsatisfactory because Einstein's theory of Special Relativity tells us that nothing can travel faster than the speed of light. This seems to be in contradiction with changing all points in space at the same time.

The second possibility is to change the potential at all points in space by an amount that depends on the position of the point. This is a *local* transformation and is consistent with relativity. One can also say that this is the most general transformation we can write down and that there is no reason to restricts oneself to global transformations. On the other side, it seems to have a measurable effect. If the potential is changed at point A and point B with different amounts, the potential difference will no longer be the same. Point A must, therefore, communicate to point B that it has been changed and by the amount it has been changed. This is done by sending out a particle which in the case of the electromagnetic field is the photon. Requiring invariance under *local gauge transformations* leads in a natural way from a non-interacting theory to an interacting one, with particles communicating the interaction in question.

Gauge Invariance in a Formal Way

Noether's theorem says symmetries lead to conserved quantities and the natural question is what conserved quantity is implied from global and local gauge invariance. From the previous example we expect that it should be linked to the electric charge since the fields stay the same, In addition, in this special case the conserved charge should also be the same for global and local transformations. This turns out to be verified. Invariance under global phase transformations of the type

$$U = e^{i\alpha Q} \tag{2.12}$$

leads to the conservation of electric[2] charge Q. Replacing this with a local transformation only reproduces the same charge conservation.

[2] This conservation law also applies to other types of charges than the electric. It is assumed that the field is complex.

To see how gauge bosons come into the picture, let us be more specific and look at the non-interacting Quantum ElectroDynamic (QED) Lagrangian. For the complex Dirac field $\psi(x)$ the Lagrangian is

$$\mathcal{L} = \bar{\psi}(x)\,(i\gamma^{\mu}\partial_{\mu} - m)\,\psi(x) \tag{2.13}$$

where the adjoint field is $\bar{\psi} = \psi^{\dagger}\gamma^{0}$

Any global phase transformation of the type

$$\psi(x) \longrightarrow \psi'(x) = e^{i\alpha}\psi(x) \tag{2.14}$$

will have no effect on the Lagrangian. These transformations define the unitary group of dimension one, U(1). However, in a local transformation

$$\psi(x) \longrightarrow \psi'(x) = e^{i\alpha(x)}\psi(x) \tag{2.15}$$

the term in the Lagrangian involving a derivative will cause problems. In fact, the derivative in the direction of the vector η^{μ} is defined by

$$\eta^{\mu}\partial_{\mu}\psi = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\psi(x+\epsilon n) - \psi(x)\right] \tag{2.16}$$

Since the local phase transformation depends on the point x, the field transforms differently for $\psi(x+\epsilon n)$ and $\psi(x)$ and the derivative is not well defined. This is the same problem as in the previous example. There it was solved by introducing a communicating particle between the field at the two different points A and B. Here, for more formal reasons, something similar must be done.

We want, in our definition of the derivative, to subtract the values of the field $\psi(x)$ at neighboring points which means a compensating factor for the phase difference between the points must be introduced [3]. The two points must be connected in some way. This can be done by introducing a scalar $U(x,y)$ that depends on the two points and which transforms as

$$U(x,y) \longrightarrow U^{'}(x,y) = e^{i\alpha(y)}\,U(x,y)\,e^{-i\alpha(x)} \tag{2.17}$$

The field $\psi(x)$ and the scalar quantity $U(x,y)$ now transform in the same way and a derivative can be defined as

$$\eta^{\mu}\partial_{\mu}\psi = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\psi(x+\epsilon n) - U(x+\epsilon n, x)\psi(x)\right] \tag{2.18}$$

This is called the *covariant derivative.*

Assuming that $U(x,y)$ is a continuous function of x and y, it can be expanded as

$$U(x+\epsilon n, x) = 1 - ie\,\epsilon\eta^{\mu}A_{\mu}(x) + \mathcal{O}(\epsilon^{2}) \tag{2.19}$$

where the constant e has been arbitrarily extracted. A new vector field $A_{\mu}(x)$ has been introduced. In the Lagrangian this field will be sandwiched between the two Dirac fields. It is nothing else than the photon, the particle communicating the interaction.

The appearance of a vector field when we require invariance under local gauge transformations is a general feature, i.e. it has *dynamical* implications and leads to *gauge bosons.* In this case the *covariant derivative* must be used instead of the usual derivative:

$$\partial_\mu \psi(x) \longrightarrow D_\mu \psi(x) = \partial_\mu \psi(x) + ieA_\mu(x)\psi(x) \tag{2.20}$$

The covariant derivative transforms in the same way as the field $\psi(x)$ itself. The meaning of the 'arbitrary' parameter e is now clear. It is the electric charge. One also finds that the vector field A_μ transforms as

$$A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu \alpha(x) \tag{2.21}$$

i.e. it is seen explicitly that the transformation depends directly on the 'rate of change' of the parameter α, something which was anticipated in our example. In general, there is one field associated to each generator of the group. The group U(1) has one generator.

2.3 The Standard Model

2.3.1 General Description

The Standard Model is a local gauge theory based on the symmetry

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

and describes the interactions between *fermions* and *gauge bosons.* The theory is in fact composed of two distinct parts: Quantum ChromoDynamicsy and the electroweak part.

2.3.2 Quantum ChromoDynamics QCD

QCD is based on the symmetry group $SU(3)_C$ color. Even though the strong interaction is not responsible for neutral B meson oscillations, QCD corrections to the relevant weak Feynman diagrams must be included. This means that (non-perturbative) QCD will become important when one tries to extract useful information of the measurements. QCD will, therefore, only be further described in connection to neutral B meson oscillations.

2.3.3 The Electroweak Sector

In the 1960s a unification of the electromagnetic and the weak interactions were suggested by Glashow, Weinberg and Salam [4]. The unified theory is a local gauge theory based on the symmetry group $SU(2)_L \times U(1)_Y$. The main characteristics are:

- Fermions have different transformation-properties depending on their *handedness.* The *left* and *right* handed components of a field are obtained from

$$\psi^{L,R}(x) = \frac{1}{2}(1 \mp \gamma_5)\,\psi(x) \;=\; P^{L,R}\,\psi(x) \tag{2.22}$$

  where $P^{L,R}$ is called the *helicity* projection operator. The helicity is the spin component on the momentum of the particle: $\frac{\vec{p}\cdot\vec{s}}{|\vec{p}|}$. Helicity is only a conserved quantity for massless particles since a Lorentz boost of a massive particle can change the sign of $\vec{p}$ and, therefore, of the helicity.

Left handed fields transform as a doublet under the symmetry group $SU(2)_L$ while the right handed fields are singlets under this transformation. Assuming the neutrinos to be massless (although there is recent experimental evidence for a small, but non-zero neutrino mass [5]), i.e. no need for right handed neutrinos, we have:

$$\text{Leptons:} \begin{pmatrix} \nu_l \\ e_l \end{pmatrix}^L, \; e_l^R \quad \text{where } \nu_l = \nu_e, \nu_\mu, \nu_\tau \quad \text{and } e_l = e^-, \mu^-, \tau^-$$
$$\text{Quarks:} \begin{pmatrix} u_i \\ d_i \end{pmatrix}^L, \; u_i^R, \, d_i^R \quad \text{where } u_i = u, c, t \quad \text{and } d_i = d, s, b$$

- The $SU(2)_L$ part of the Lagrangian is invariant under the local gauge transformation[3]

$$\Psi_l^L(x) \longrightarrow \Psi_l^{\prime L}(x) = e^{ig\frac{\tau_j}{2}\omega_j(x)}\, \Psi_l^L(x) \tag{2.23}$$
$$\psi_l^R(x) \longrightarrow \psi_l^{\prime R}(x) = \psi_l^R(x) \tag{2.24}$$

where

$$\Psi_l^L(x) = \begin{pmatrix} \psi_{\nu_l}^L(x) \\ \psi_l^L(x) \end{pmatrix}, \quad l = e^-, \mu^-, \tau^- \tag{2.25}$$

and $\omega_{j=1,2,3}(x)$ are three arbitrary real differentiable functions of x, g is a real constant which we can identify as the coupling constant of $SU(2)_L$ and $\tau_{j=1,2,3}$ are the Pauli spin matrices. The right handed fields are singlets under this transformation, reflecting the fact that only left handed fields participate in the weak interaction. These transformations lead to three conserved charges called *weak isospin* $I^W_{i=1,2,3}$

One important difference with respect to QED is that the Pauli matrices do not commute. This means that $SU(2)_L$ is a *non-Abelian* gauge theory. All elements in the U(1) group commute so QED is an *Abelian* gauge theory. The main difference between the two is that in Abelian gauge theories the gauge bosons do not couple directly to themselves (i.e. a photon couples to electric charge, but not to itself since the photon is neutral), while in non-Abelian gauge theories the gauge bosons possess the relevant charge and can couple directly to other gauge bosons. $SU(3)_C$ is also a non-Abelian gauge theory and triplet and quartet gluon couplings are possible.

- The $U(1)_Y$ part of the Lagrangian is invariant under the local gauge transformation

$$\psi_l(x) \longrightarrow \psi_l^{\prime}(x) = e^{ig'Yf(x)}\, \psi_l(x) \tag{2.26}$$

where $f(x)$ is an arbitrary real differentiable function, g' is the coupling constant of $U(1)_Y$ and Y is the weak hypercharge. The weak hypercharge satisfies the Gell-Mann-Nishijima relation

$$Y = Q - I_3^W \tag{2.27}$$

[3]In the following only the lepton part is written out. The quark part can be written in an analogue way

where Q is the electric charge in units of the electron charge and I_3^W is the third component of the weak isospin.

- Introducing the covariant derivatives, four gauge bosons will appear. For $SU(2)_L$ there are the three $W^\mu_{i=1,2,3}(x)$ while for $U(1)_Y$ there is $B^\mu(x)$. Formally, this reflects that the group $SU(2)_L$ has three generators while $U(1)_Y$ has one.

For the moment this model seems to be composed of two distinct parts. The $W^\mu_{i=1,2,3}(x)$ are invariant under $U(1)_Y$ gauge transformations while $B^\mu(x)$ is invariant under $SU(2)_L$ gauge transformations. The total Lagrangian is, therefore, invariant under $SU(2)_L \times U(1)_Y$ gauge transformations. There is, however, an indication that this model may connect the weak and electromagnetic interactions.

It was mentioned that there are three conserved weak isospin charges. Two of the corresponding currents can couple electrically neutral with electrically charged fermions and are, therefore, called *charged currents*. The third one couples either electrically neutral or electrically charged fermions, i.e. a *neutral current*. The last part bears great resemblance to the usual QED current. We have:

$$J_3^\mu(x) = \frac{1}{2}[\bar{\psi}^L_{\nu_l}(x)\gamma^\mu\psi^L_{\nu_l}(x) - \bar{\psi}^L_l(x)\gamma^\mu\psi^L_l(x)] \tag{2.28}$$

$$J^\mu_{QED}(x) = -e\,\bar{\psi}_l(x)\gamma^\mu\psi_l(x) \tag{2.29}$$

We, therefore, write out the part of the Lagrangian containing the interactions between the fermion fields and the gauge bosons:

$$\mathcal{L} = i[\bar{\Psi}^L_l(x)\gamma^\mu D_\mu \Psi^L_l(x) \; + \; \bar{\psi}^R_l(x)\gamma^\mu D_\mu \psi^R_l(x)] \tag{2.30}$$

In analogy with equation (2.18) we have to introduce four real fields, $W^\mu_j, j = 1,2,3$ for $SU(2)_L$ and B^μ for $U(1)_Y$, and the covariant derivatives become:

$$D^\mu \Psi^L_l(x) = [\partial^\mu \; + \; ig\frac{\tau_j}{2}W^\mu_j \; + \; ig'YB^\mu(x)]\,\Psi^L_l(x) \qquad j = 1,2,3 \tag{2.31}$$

$$D^\mu \psi^R_l(x) = [\partial^\mu \; + \; ig'YB^\mu(x)]\,\psi^R_l(x) \tag{2.32}$$

The idea is now to go from these four real gauge fields to the four physical fields. We define four new fields, two charged and two neutral, by:

$$\begin{pmatrix} W^+_\mu \\ W^-_\mu \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & +i \end{pmatrix}\begin{pmatrix} W^1_\mu \\ W^2_\mu \end{pmatrix} \tag{2.33}$$

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} \tag{2.34}$$

The angle θ_W specifies the mixing of the fields and is called the *Weinberg angle*. In terms of the two coupling constants g and g' it becomes

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 \; + \; g'^2}} \tag{2.35}$$

By rewriting the Lagrangian as a function of the four new fields, and by requiring that the field A_μ is the usual electromagnetic field, we get

$$e = g\sin\theta_W \tag{2.36}$$
$$e = g'\cos\theta_W \tag{2.37}$$

The coupling g' can be eliminated and the Lagrangian can be rewritten using the coupling constant g and the Weinberg angle.

The Weinberg angle is an interesting quantity. It tells us how strong the coupling between the weak and electromagnetic interactions is. For $\theta_W = 0$ two interactions are completely decoupled. This is ruled out by experimental measurements [1] which means that the weak and electromagnetic interactions have been unified!

What can also be verified by measuring this angle is that our choice of gauge group is correct. The value of $\sin^2\theta_W$ predicted by $SU(2)_L \times U(1)_Y$ agrees to a high precision with the experimental result. This is not the case if we choose other gauge groups. It should be mentioned that this model also predicted both the existence of three new particles, $W^\pm$ and the second neutral field, Z^0, and their mass. They were found where they were expected to be by the UA1 and UA2 collaborations at CERN [6]. Since then, the electroweak model has been tested to a very high precision numerous times, especially at LEP. During the LEP1 phase the collider functioned as a "Z^0 factory" producing over 4 million hadronic Z^0 per experiment. During the current LEP2 phase it is also able to produce W^+W^- pairs.

The $SU(2)_L \times U(1)_Y$ model of electroweak interactions has never been contradicted by experimental results.

2.3.4 Masses in the Standard Model

It turns out that the question of giving the gauge bosons and the fermions masses, while keeping the photon massless, is far from a trivial question.

The intuitive choice of gauge boson mass terms in the Lagrangian is something like

$$m_W^2\, W_\mu^\dagger(x)W^\mu(x) \;+\; \frac{1}{2}m_Z^2 Z_\mu(x)Z^\mu(x) \tag{2.38}$$

Unfortunately, these terms are neither invariant under $SU(2)_L$ nor $U(1)_Y$ gauge transformations. They also lead to problems with the *renormalizability* of the theory.

A theory is said to be *renormalizable* if its predictions in terms of a finite number of parameters remain finite when all cut-offs appearing in the theory are removed (formally taken to infinity). Any theory pretending to make predictions about Nature must be renormalizable.

One can also add a similar term for the fermion masses, but again it is not $SU(2)_L \times U(1)_Y$ invariant. So our choice seems to be to either insist that the particles are massless, thus contradicting experimental observations, or break the $SU(2)_L \times U(1)_Y$ invariance and remain with a non-renormalizable, i.e. non-physical, theory. Fortunately, there exists a mechanism to introduce masses without spoiling the gauge invariance. Using this approach it is possible to predict the correct $W^\pm$ and Z^0 masses, and keep the photon massless.

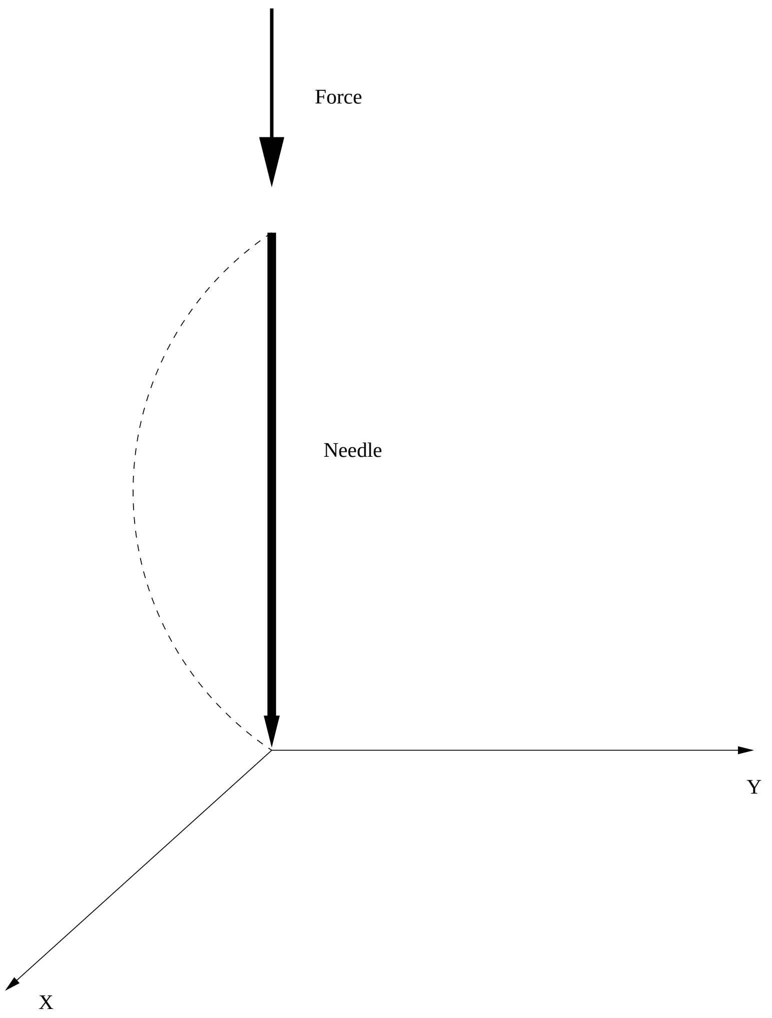


Figure 2.1: *Example of spontaneous symmetry breaking: The needle is initially in a perfect symmetric state around the z-axis. By applying pressure on the needle it bends into a new ground state. The initial symmetry is broken, but it is impossible to predict which new state the needle will be in (which way it will bend). The breaking of the rotational symmetry around the z-axis is, therefore, said to be* spontaneous.

Spontaneous Symmetry Breaking

The concept of spontaneous symmetry breaking can be illustrated with a simple example: the buckling of a compressed needle (see figure 2.1). Initially, the needle stays in the position x = y = 0, i.e. there is a perfect symmetry around the z-axis. When the force becomes too strong the needle bends since the energy in this state is lower than in the initial one. The rotational symmetry around the z-axis is broken. The needle can bend in any direction and still be in a state with the same energy, i.e. it is impossible to predict which (asymmetric) ground state the needle will choose. The breaking of the rotational symmetry around the z-axis is, therefore, said to be *spontaneous.*

Mass-generation with Spontaneous Symmetry Breaking and Perturbation Theory

To see how one can generate masses with the mechanism of spontaneous symmetry breaking, let us consider the following Lagrangian for a scalar field ϕ:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 \ - \ \frac{1}{2}\mu^2\phi^2 \ - \ \frac{1}{4}\lambda\phi^4 \tag{2.39}$$

where $\lambda > 0$ and $\mu^2 < 0$ (see figure 2.2 for a two-dimensional version of this). The field has to be a scalar field if we require the vacuum states to be invariant under Lorentz transformations and under translations.

Since μ^2 is negative, the mass term $-\frac{1}{2}\mu^2\phi^2$ in the Lagrangian has the wrong sign. The potential has two minima $\pm v$, where $v = \sqrt{-\mu^2/\lambda}$, while $\phi = 0$ is no longer the energy minimum. Since we want to use perturbation theory, the field is expanded around one of the minima (we are free to chose any of them):

$$\phi(x) \ = \ v \ + \eta(x) \tag{2.40}$$

where $\eta(x)$ is the quantum fluctuation around the minimum v. Inserting this into the original Lagrangian, terms like

$$-\lambda v^2 \eta^2 \tag{2.41}$$

appear, i.e. the scalar field $\eta(x)$ has a mass term with the right sign. This is, at first sight, strange because expanding the original field around v should not change the physics! The explanation is that if we could solve the Lagrangians *exactly* they would have the same solutions. However, in particle physics we are not able to do that and must resort to *perturbation theory*. Since the original point $\phi = 0$ is an unstable point, any expansion around that point will give a perturbation series that does not converge. The point $\phi = v$ is a stable point, however, and we are free to expand around this point. This means that the correct description in perturbation theory of the physics is given by a Lagrangian expressed in terms of the field $\eta(x)$. In this operation, the initial symmetry has been spontaneously broken by choosing a stable ground state $\phi = +v$ and at the same time generated a mass for the field $\eta(x)$.

Equation (2.40) can be used to derive a condition for spontaneous symmetry breaking in a quantized theory. As there are no particles in the vacuum, we have

$$< 0|\phi(x)|0 >= v \tag{2.42}$$

i.e. the field must have a *non-vanishing vacuum expectation value* (this is also visible in figure 2.2).

The General Principles of the Higgs Mechanism

The previous mechanism can easily be extended to require invariance under $U(1)$ local gauge transformations. We know this will introduce a real, massless vector field A_μ via the covariant derivative

$$D_\mu\phi(x) \ = \ [\partial_\mu \ + \ iqA_\mu(x)]\,\phi(x) \tag{2.43}$$

where the field transforms as

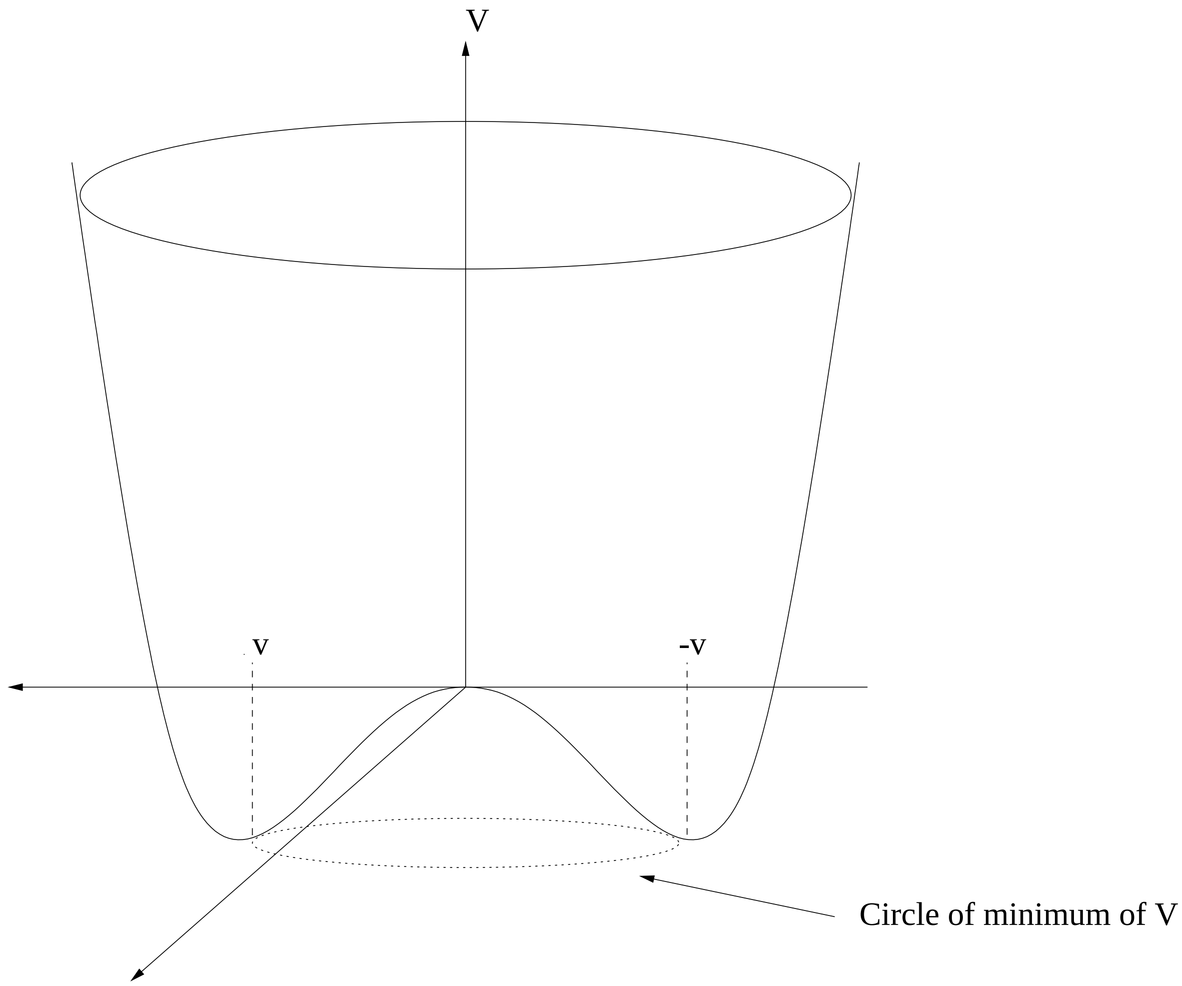


Figure 2.2: *Higgs potential: The potential has an unstable local maximum for field values of 0 while the minimum occurs at a circle of non-zero values of the two fields.*

$$\phi(x) \longrightarrow \phi^{'}(x) = \phi(x)e^{-iqf(x)} \tag{2.44}$$

Here $\phi(x)$ is a complex scalar field. The Lagrangian density becomes

$$\mathcal{L}(x) = [D^{\mu}\phi(x)]^{*}[D^{\mu}\phi(x)] \ - \ \mu^{2}|\phi(x)|^{2} \ - \ \lambda|\phi(x)|^{4} \tag{2.45}$$

where the potential is called the *Higgs potential.*

By expanding around the stable minimum v, the field $\phi(x)$ can be written as

$$\phi(x) = \frac{1}{\sqrt{2}}[v \ + \ \sigma(x) \ + \ i\eta(x)] \tag{2.46}$$

where $\sigma(x)$ and $\eta(x)$ are two real fields. However, one can always choose the gauge transformation of the field $\phi(x)$ such that one eliminates the field $\eta(x)$, and $\phi(x)$ becomes

$$\phi(x) = \frac{1}{\sqrt{2}}[v \; + \; \sigma(x)] \tag{2.47}$$

This gauge is called the *unitary gauge*. It can be shown that if the field $\eta(x)$ is kept, the Lagrangian will have too many degrees of freedom compared to the original one (five versus four). These additional degrees of freedom will have to be eliminated.

By inserting this into the original Lagrangian we get terms like

$$-\frac{1}{2}(2\lambda v^2)\,\sigma^2(x), \quad +\frac{1}{2}(qv)^2 A_\mu(x)A^\mu(x) \tag{2.48}$$

i.e. having started out with a *complex* scalar field and a *massless* vector field, we have ended up with a *real* scalar field and a *massive* vector field. One of the two degrees of freedom of the original complex field $\phi(x)$ has been transfered to the massless vector field and it has acquired a mass in the process. One way to see this is to notice that a massless vector field has two degrees of freedom while a massive vector field has three.

We also notice that in the process a massive real scalar field $\sigma(x)$ appeared. The particle associated with this field is called the *Higgs boson.* It should also be emphasized that this mass generating process does not break the gauge invariance of the Lagrangian. This mechanism is known as the *Higgs mechanism.*

The Higgs Mechanism and the Electroweak Model

We need to give masses to the weak gauge bosons while keeping the photon massless. We, therefore, introduce a weak isospin Higgs doublet

$$\Phi(x) = \begin{pmatrix} \phi_a(x) \\ \phi_b(x) \end{pmatrix} \tag{2.49}$$

where $\phi_a(x)$ and $\phi_b(x)$ are two complex scalar fields. $\Phi(x)$ transforms in the usual way under $SU(2)_L$ and $U(1)_Y$. Just like in the previous section, we choose a specific ground state $\Phi_0(x)$ and thus break spontaneously the symmetry. In the unitary gauge the field becomes

$$\Phi_0(x) = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{2.50}$$

where

$$v = \sqrt{-\mu^2/\lambda} \quad > \; 0 \tag{2.51}$$

Even though this ground state is not invariant under $SU(2)_L \times U(1)_Y$, it is invariant under $U(1)_{EM}$ electromagnetic gauge transformations. This ensures a massless photon and conserves electric charge.

$$SU(2)_L \times U(1)_Y \overset{S.S.B.}{\longrightarrow} U(1)_{EM} \tag{2.52}$$

We now perturb around this particular ground state by introducing four real fields

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \eta_1(x) + i\,\eta_2(x) \\ v + \sigma(x) + i\,\eta_3(x) \end{pmatrix} \tag{2.53}$$

Three of these fields disappear by transfering their degrees of freedom to the vector bosons, thus giving mass to the W^+, W^- and Z^0. The photon remains massless while one of the four original fields survives and gives rise to a massive neutral spin-0 particle, the Higgs boson:

$$\Phi(x) = \begin{pmatrix} 0 \\ (v + \sigma(x))/\sqrt{2} \end{pmatrix} \tag{2.54}$$

The particle masses become:

$$\begin{aligned} m_H &= \sqrt{-2\mu^2} && (2.55) \\ m_Z &= \frac{m_W}{\cos\theta_W} && (2.56) \\ m_W &= \frac{1}{2} v g = \frac{1}{2} v \frac{e}{\sin\theta_W} && (2.57) \\ m_\gamma &= 0 && (2.58) \end{aligned}$$

The predicted values for the gauge bosons agree with the experimental observations. This was a huge success for the model. For the Higgs mass things are a little different though. We can express v as a function of known constants, but there is nothing in the model that predicts the value of λ. We, therefore, can not predict μ^2. Since this is the parameter giving the Higgs mass we can't go any further.

Fermion Masses

The fermion masses are introduced via Yukawa couplings between the fermion and Higgs fields. For the massive leptons (we here assume the neutrinos to be massless) the interaction Lagrangian is written

$$\mathcal{L}_{\text{Yukawa}} = g_l \, \bar{\Psi}_l^L(x) \, \psi_l^R(x) \, \Phi(x) \; + \; h.c. \tag{2.59}$$

where g_l is a dimensionless coupling constant, and $l = e, \mu, \tau$. After the symmetry breaking the Lagrangian becomes

$$\mathcal{L}_{\text{Yukawa}} = -m_l \, \bar{\psi}_l^L \, \psi_l^R \, (1 + \frac{\sigma(x)}{v}) \; + \; h.c. \tag{2.60}$$

where the fermion mass is

$$m_l = \frac{-v g_l}{\sqrt{2}} \tag{2.61}$$

We see that the coupling of the Higgs to the fermions is proportional to the fermion mass.

Quark Masses

In the previous paragraph we saw how to write the Yukawa couplings for the massive fermions when we assumed that the neutrinos are massless. To give mass to quarks we must do something similar. Since all the quarks are massive, however, there will be one difference.

The structure of the coupling of the Higgs field to the fermion fields does not follow from any gauge symmetry and there are no constraints on it. In particular, there is nothing that says that we should have quark flavor conservation. This means that in the quark sector we must write down the most general form of the coupling which allows for quark mixing. The interaction Lagrangian becomes

$$\mathcal{L}_{Yukawa} = -(\bar{u}_L^{'} m_u u_R^{'} \; + \; \bar{d}_L^{'} m_d d_R^{'})\,(1 + \frac{\sigma(x)}{v}) \; + \; h.c. \tag{2.62}$$

where

$$u_{L,R}^{'} = \begin{pmatrix} u^{'} \\ c^{'} \\ t^{'} \end{pmatrix}_{L,R} \qquad d_{L,R}^{'} = \begin{pmatrix} d^{'} \\ s^{'} \\ b^{'} \end{pmatrix}_{L,R} \tag{2.63}$$

and m_u, m_d are 3×3 complex matrices whose elements are defined in an analogue way to the lepton masses, i.e.

$$m_{u_{ij}} = \frac{-v}{\sqrt{2}} Y_{u_{ij}} \quad m_{d_{ij}} = \frac{-v}{\sqrt{2}} Y_{d_{ij}} \tag{2.64}$$

where $Y_{u_{ij}}$ and $Y_{d_{ij}}$ are the Yukawa coupling constants. These matrices are not diagonal. We define new physical quark fields that diagonalize the mass matrices:

$$u_{L,R} = V_{L,R}^{u}\, u_{L,R}^{'} \tag{2.65}$$

$$d_{L,R} = V_{L,R}^{d}\, d_{L,R}^{'} \tag{2.66}$$

where $V_{L,R}^{u,d}$ are unitary matrices. The interaction Lagrangian becomes

$$\mathcal{L}_{Yukawa} = -[\bar{u}_L m_{Diag}^{u} u_R \; + \; \bar{d}_L m_{Diag}^{d} d_R]\,(1 + \frac{\sigma(x)}{v}) \tag{2.67}$$

where

$$m_{Diag}^{u,d} = V_L^{u,d}\, m^{u,d}\, V_R^{\dagger u,d} \tag{2.68}$$

are the diagonal mass matrices.

Since there are no right handed neutrinos, the mass matrix in the leptonic sector is diagonal and lepton flavor is conserved.

2.3.5 The CKM Matrix

The rotation of the quark fields affects only parts of the electroweak sector. We see that currents only involving quarks of the same type (i.e. *up* or *down*) will not be affected at all,

$$\bar{u_L} u_L = \bar{u}_L^{'} V_L^{u\dagger} V_L^{u} u_L^{'} = \bar{u}_L^{'} u_L^{'} \tag{2.69}$$

since $V_L^{u\dagger} V_L^{u} = 1$. This means that there is no quark mixing involving the neutral current. In the Z^0 sector this is confirmed by the absence of Flavor Changing Neutral Currents (FCNC) at tree level. The coupling of quark fields to the $W^{\pm}$, however, involves both up and down type quarks and quark mixing will occur. The charged weak current couplings can be written as

$$\mathcal{L}_W = \frac{g}{\sqrt{2}} \left(\bar{u}_L^{'} \gamma^\mu W_\mu^+ d_L^{'} \; + \bar{d}_L^{'} \gamma^\mu W_\mu^- u_L^{'} \right) = \frac{g}{\sqrt{2}} \left(\bar{u}_L \gamma^\mu W_\mu^+ V d_L \; + \bar{d}_L \gamma^\mu W_\mu^- V^\dagger u_L \right) \quad (2.70)$$

where we have introduced the matrix

$$V = V_L^u V_L^{\dagger d} \quad (2.71)$$

This matrix is a 3×3 unitary matrix describing the quark mixing and is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \quad (2.72)$$

The different elements of the matrix give the coupling strength between the different quarks.

The CKM matrix has its origin in the fermion mass generation part of the Standard Model. Unfortunately, the parameters that describe these masses and matrix elements are fundamental parameters of the theory. The only way to get their numerical values are to measure them experimentally. It is, therefore, of fundamental importance to measure all the matrix elements as precisely as possible. This will enable us to examine the internal consistency of the model and also get some hints about the quark mass matrix and, thus, get a better understanding of how the fermions acquire mass.

Number of Free Parameters in the CKM Matrix

The CKM matrix is the product of two complex unitary matrices and is thus also a complex unitary matrix. This means that in the general case of N families it has $2 \times (N \times N) = 2N^2$ real parameters. Unitarity removes N^2 of these by imposing the same number of constraints. The quark phases are not defined in an unique way since they are not observables. This gives another $2N$ constraints. However, the overall phase is redundant so there are in fact only $2N$-1 constraints. This leaves $2N^2 - N^2 - (2N-1) = (N-1)^2$ free real parameters.

These can be decomposed into $\frac{1}{2}N(N-1)$ mixing angles and $\frac{1}{2}(N-1)(N-2)$ phases. For two quark families the matrix is real with one free parameter: the *Cabibbo angle* [9]. For the case of three families there are 3 angles and 1 phase. In the Standard Model *CP* violation is directly linked to the existence of a complex phase in the quark mixing matrix, i.e. it can only occur in a world with three or more quark families. This was the original motivation for Kobayashi and Maskawa to postulate the existence of a third quark family at a time when only two were known [10].

Parameterizations of the CKM Matrix

There are several possible parameterizations of the CKM matrix. Two of them will be detailed here. Since there are three mixing angles and one phase, the matrix can be written as the product of three rotation matrices where one of these has a phase. This is the parameterization adopted by the PDG [1]:

$$V_{CKM} = \mathcal{R}_{23}(\theta_{23}, 0) \times \mathcal{R}_{13}(\theta_{13}, -\delta) \times \mathcal{R}_{12}(\theta_{12}, 0) \quad (2.73)$$

$$V_{CKM} = \begin{bmatrix} C_{12}C_{13} & S_{12}C_{13} & S_{13}e^{-i\delta} \\ -S_{12}C_{23} - C_{12}S_{23}S_{13}e^{i\delta} & C_{12}C_{23} - S_{12}S_{23}S_{13}e^{i\delta} & S_{23}C_{13} \\ S_{12}S_{23} - C_{12}C_{23}S_{13}e^{i\delta} & -C_{12}S_{23} - S_{12}C_{23}S_{13}e^{i\delta} & C_{23}C_{13} \end{bmatrix} \quad (2.74)$$

where ij are quark indices, $\mathcal{R}_{ij}(\theta_{ij}, \delta)$ is a rotation in the (ij) plane of an angle θ_{ij} and phase δ, $C_{ij} = \cos\theta_{ij}$ and $S_{ij} = \sin\theta_{ij}$.

Even though this parameterization is correct, it hides an important fact. The couplings are of varying strength, with the coupling inside the same family being the strongest ones and then weakening the further away one gets from the diagonal. To better illustrate this important feature Wolfenstein [7] proposed an approximate parameterization

$$V_{CKM} = \begin{bmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4) \quad (2.75)$$

where

$$\lambda = \sin\theta_{12} \quad (2.76)$$

$$A = \frac{\sin\theta_{23}}{\lambda^2} \quad (2.77)$$

$$\rho = \frac{\sin\theta_{13}}{A\lambda^3}\cos\delta \quad (2.78)$$

$$\eta = \frac{\sin\theta_{13}}{A\lambda^3}\sin\delta \quad (2.79)$$

The angle θ_{12} is the already mentioned Cabibbo angle.

Measurement of the CKM Matrix Elements

The values, and the precision with which the CKM matrix elements are known, vary a lot. In principle, the matrix elements can all be determined by weak decays of quarks or from deep inelastic neutrino scattering. The present knowledge of the matrix elements comes from [1]:

- Nuclear β-decays give
$$|V_{ud}| = 0.7740 \pm 0.0010$$
- K_{e3} decays $K^+ \longrightarrow \pi^0 e\nu_e$ and $K_L^0 \longrightarrow \pi^- e^+\nu_e$ give
$$|V_{us}| = 0.2196 \pm 0.0023$$
Semileptonic hyperon decays can also be used.
- Measuring the lepton spectrum in the decay $b \longrightarrow ul\bar{\nu}_l$ above the endpoint of $b \longrightarrow cl\bar{\nu}_l$ gives
$$|V_{ub}/V_{cb}| = 0.08 \pm 0.02$$
Exclusive channels of semileptonic B decays like $B \longrightarrow \rho l\nu_l$, $\pi l\nu_l$ can also be used.

- Single charm production in deep inelastic neutrino (antineutrino) - nucleon scattering $\nu_\mu d \longrightarrow \mu^- c$ gives

$$|V_{cd}| = 0.224 \pm 0.016$$

- D_{e3} decays $D \longrightarrow \bar{K} e^+ \nu_e$ give

$$|V_{cs}| = 1.04 \pm 0.16$$

  W decays [8] (now available at LEP2) can also be used.

- Semileptonic decays of B mesons to charmed mesons $B \longrightarrow \overline{D^*} l^+ \nu_l$ and inclusive semileptonic B decays give

$$|V_{cb}| = 0.0395 \pm 0.0017$$

- Semileptonic decays of the top quark to the b quark give

$$\frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} = 0.99 \pm 0.29$$

- From B mixing one can set a limit [1] on the ratio

$$|V_{td}/V_{ts}| \; < \; 0.27$$

Numerically we have [1], assuming unitarity and three generations,

$$V_{CKM} = \begin{bmatrix} 0.9745 - 0.9757 & 0.219 - 0.224 & 0.002 - 0.005 \\ 0.218 - 0.224 & 0.9736 - 0.9750 & 0.036 - 0.046 \\ 0.004 - 0.014 & 0.034 - 0.046 & 0.9989 - 0.9993 \end{bmatrix}$$

Experimental Motivation for the CKM Matrix

From the previous sections one gets the impression that the CKM matrix is purely theoretically motivated. There is, however, also an experimental motivation.

At the time when only the three quarks u, d and s were known, it was observed that the strangeness-changing process $s \longrightarrow u \, + \, W$ was weaker than the corresponding strangeness-conserving process $d \longrightarrow u + W$. It was also observed that the strangeness-conserving process was weaker than the corresponding leptonic transition. N. Cabibbo [9] then suggested in 1963 that the strangeness-changing process contained a factor $\sin\theta_c$ while the strangeness-conserving process contained a factor cos θ_c. The angle θ_c is called the Cabibbo angle. The rotated d quark could then be written as

$$d \cos\theta_c \; + \; s \sin\theta_c \tag{2.80}$$

while the u quark remained unrotated. The smallness of the Cabibbo angle could then explain the different strength between the two processes.

This model worked fine, but still had one problem. It predicted a rate of $K^0 \longrightarrow \mu^+\mu^-$ decay that was too large compared to the experimental limit. This was a specific example of the fact that the Cabibbo model allows for FCNC at tree level, something which is not observed.

The neutral current can be written as

$$u\bar{u} + d\bar{d}\cos^2\theta_c + s\bar{s}\sin^2\theta_c + (d\bar{s} + s\bar{d})\cos\theta_c\sin\theta_c \tag{2.81}$$

where the last term is responsible for the FCNC. To solve this problem a fourth quark was introduced in 1970 by Glashow, Iliopoulos and Maiani [11]. This quark, called the c quark, should couple to the d quark with a factor -sin θ_c and to the s quark with a factor cos θ_c to cancel out the corresponding FCNC term. The weak quark eigen states $(d^{'}, s^{'})$ are, therefore, rotated with respect to the flavor eigenstates (d, s), with the rotation matrix given by

$$\begin{pmatrix} d^{'} \\ s^{'} \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{2.82}$$

while u, c remain unrotated. This is a 2×2 version of the CKM matrix.

The introduction of the c quark, which was done for purely theoretical reasons, was confirmed with the experimental discovery of $J/\psi(c\bar{c})$ in 1974 [12].

2.4 *CP* Violation

The previous sections gave examples of continuous symmetries and how they are linked to conserved quantities. It was also shown how more abstract local gauge symmetries have dynamical implications. There exists also a third type of symmetries, associated with the dynamics of the interactions, which are the discrete symmetries.

2.4.1 Discrete symmetries

Charge conjugation

Charge conjugation C is the operation of changing a particle into the corresponding antiparticle. This is obtained by changing the sign of all internal quantum numbers, but leaving spin and momentum unchanged.

$$C|f(\vec{p},\vec{s})> = \eta_c|\bar{f}(\vec{p},\vec{s})> \tag{2.83}$$

where $\vec{p}$ and $\vec{s}$ are the momentum and spin, respectively, of the particle and η_c is a phase factor.

Parity

Parity P changes the position vector $\vec{r}$ into $-\vec{r}$. This means that the momentum changes sign while the orbital momentum, and by analogy the spin, does not.

$$P|f(\vec{p},\vec{s})> = \eta_P|f(-\vec{p},\vec{s})> \tag{2.84}$$

where η_P is a phase factor called the *intrinsic parity*. The parity operation also changes the *handedness* of the particle by flipping the sign of the helicity, i.e. it transforms a left handed particle into a right handed one.

Time Reversal

Time reversal T changes the time t into $-t$. This means the momentum and the spin change sign while the helicity stays the same.

$$T|f(\vec{p},\vec{s})>=\eta_P^s|f(-\vec{p},-\vec{s})>^* \tag{2.85}$$

where η_T^s is a phase factor that depends on the spin [13], [14]. The reason the wave function has to be complex conjugated can bee seen from the Schrödinger equation

$$i\,\frac{\partial\psi(\vec{r},t)}{\partial t} \;=\; -\frac{\hbar^2}{2m}\,\nabla^2\psi(\vec{r},t) \tag{2.86}$$

The wave function

$$\psi(\vec{r},t)=e^{i(\vec{p}\vec{r}\,-\,Et)} \tag{2.87}$$

where $E=\frac{\vec{p}^2}{2m}$ satisfies the Schrödinger equation while $\psi(\vec{r},-t)$ does not. However, the complex conjugate of the wave function $\psi(\vec{r},-t)^*$ satisfies it. The physical interpretation of this wave function is that it describes the same particle moving in the opposite direction.

CPT Invariance

The reason these three symmetries are so important comes from the fact that under very general hypotheses of *causality* and *locality* a field theory must be invariant under the combined operation CPT [15]. From this it can be derived that particles and anti-particles have the same mass and lifetime. One experimental verification of this comes from the mass difference between the K^0 and $\overline{K^0}$ which is found to be [1]

$$\frac{m_{\bar{K}^0}-m_{K^0}}{m_{K^0}}\leq 9\times 10^{-18}$$

CP Violation

It is an experimental fact that charge conjugation and parity are conserved in the strong and electromagnetic interactions. It is, however, possible to introduce a *CP* violating term in the QCD Lagrangian without violating any known conservation law or symmetry. There are no satisfactory explanation for the fact that this term is either zero or extremely small.

The weak interaction violates parity maximally since there are no right handed neutrinos or left handed anti-neutrinos. It also violates charge conjugation. The combined operation of C and P, however, seemed to be conserved also in weak interactions.

This caused a problem because we do expect *CP* violation in Nature. Believing that the universe was created from energy, one would expect equal amounts of matter and anti-matter. However, we live in a matter dominated universe. Even if we believe that matter and anti-matter somehow got separated, there should be boundaries where they meet and annihilate to create energy. No such energetic boundary region has ever been observed in the universe. It can be shown that both *C* and *CP* violation are necessary to create the baryon asymmetry of the universe [16], [17]. It was, therefore, exciting when *CP* violation was observed in 1964 in the neutral kaon system [18].

2.4.2 The Kaon System

To solve a certain number of mysteries with two new resonances, Λ^0 and K^0, discovered in the 1950s, the concept of *strangeness* was introduced by Gell-Mann. This was a quantum number that was conserved in the strong and electromagnetic, but not in the weak, interactions. One consequence was that the neutral kaon was different from its antiparticle. This meant that there could be oscillations between the two, i.e. a K^0 could oscillate into a $\overline{K^0}$. One way to see this is to notice that they have a common final state $\pi^+\pi^-$, and we can, therefore, have [19]

$$K^0 \longrightarrow \pi^+\pi^- \longrightarrow \overline{K^0}$$

This means that the physical states (i.e. the weak/mass eigenstates) are superpositions of the two strong/flavour eigenstates:

$$|K_1> = \frac{1}{\sqrt{2}}(|K^0> + |\overline{K^0}>) \tag{2.88}$$

$$|K_2> = \frac{1}{\sqrt{2}}(|K^0> - |\overline{K^0}>) \tag{2.89}$$

These two states are *CP* eigenstates. Using the convention $CP|K^0>=|\overline{K^0}>$ we have

$$CP|K_1>= +|\overline{K_1}> \tag{2.90}$$

$$CP|K_2>= -|\overline{K_2}> \tag{2.91}$$

This has dramatic consequences on the decay to the already mentioned state $\pi^+\pi^-$. The state has *CP* eigenvalue +1 which means that (assuming *CP* is conserved) only the $|K_1>$ state is allowed to decay into it. The $|K_2>$ state can then only decay into more complicated final states like the three pion state. Since the phase space in this case is very limited, the $|K_2>$ state lives much longer than the $|K_1>$ state. This lifetime difference is observed experimentally and the two states are quite naturally called K_S (*short*) and K_L (*long*), with

$$\tau_{K_S} \approx 0.9\,10^{-10}s \qquad \tau_{K_L} \approx 5\,10^{-8}s$$

It should be emphasized that while *CP* conservation[4] is responsible for the lifetime difference *per se*, the fact that it is so huge is a 'kinematic accident' due to the limited number of common final states and the limited phase space available in the three pion state. In the B-meson system things will be rather different.

CP Violation in the Neutral Kaon System

One way to verify that *CP* really is conserved is to use the large lifetime difference. Starting out with a beam containing both K_S and K_L, it is sufficient to wait a time $t \gg \tau_{K_S}$ to have a pure K_L beam. One can then search for the *CP* forbidden decay

$$K_L \longrightarrow \pi^+\pi^-$$

[4] As we shall see in the next section, *CP* is only approximately conserved so 'the nearly complete *CP* conservation' is more precise.

This was done in 1964 [18] and the CP forbidden decay of the K_L was observed. The CP violation is small, however, at the 10^{-3} level. But it means that the equivalence between the CP eigenstates and the weak eigenstates is no longer a valid assumption. The short and long states must now be redefined to be

$$|K_S> = \frac{|K_1> + \epsilon|K_2>}{\sqrt{1 + |\epsilon|^2}} \tag{2.92}$$

$$|K_L> = \frac{|K_2> + \epsilon|K_1>}{\sqrt{1 + |\epsilon|^2}} \tag{2.93}$$

where ϵ is a complex number of magnitude $|\epsilon| = (2.275 \pm 0.019) \times 10^{-3}$. This parameter quantifies CP violation due to the *mixing*[5].

One can also look for *direct* CP violation, i.e. where the rate for a process is different from the one for the corresponding CP conjugated one. The parameter ϵ', expressing the strength of the direct CP violation, can be found from

$$1 \;+\; 6\,\mathrm{Re}\,(\epsilon'/\epsilon) \;=\; \frac{(K_L \longrightarrow \pi^+\pi^-)/(K_S \longrightarrow \pi^+\pi^-)}{(K_L \longrightarrow \pi^0\pi^0)/(K_S \longrightarrow \pi^0\pi^0)} \tag{2.94}$$

The following experimental results were obtained:

$$\mathrm{Re}\;(\epsilon'/\epsilon) = (23 \pm 6.4) \times 10^{-4} \quad \text{(CERN NA31 [21])}$$

$$\mathrm{Re}\;(\epsilon'/\epsilon) = (7.4 \pm 5.9) \times 10^{-4} \quad \text{(Fermilab E731 [22])}$$

The NA31 result is 3.6σ away from 0, thus confirming the existence of direct CP violation. However, because of the large discrepancy between the results from the two experiments, a second generation of experiments were planned and realized, improving both the sensitivity and increasing the statistics. This year, based on 20% of their data, the Fermilab E832 collaboration announced a new and more precise result which *confirms* the NA31 result:

$$\mathrm{Re}\;(\epsilon'/\epsilon) \;=\; (28.0 \pm 4.1) \times 10^{-4} \quad \text{(Fermilab E832 [23])}$$

This implies the death of the *Superweak* model [20], a model introducing a new $\Delta S = 2$ interaction. In this model CP violation only originates from the mixing and there is no direct CP violation.

Time-reversal violation

Assuming CPT invariance, CP violation would indicate a corresponding time-reversal violation and many experimental searches have been performed. In 1998 the CPLEAR collaboration at CERN presented the first direct measurement of time-reversal violation, by looking at the transitions $K^0 \longrightarrow \overline{K^0}$ and $\overline{K^0} \longrightarrow K^0$ [24]. They found a relative asymmetry of $(6.6 \pm 1.6) \times 10^{-3}$. From CPT invariance one would expect 6.2×10^{-3}, i.e. the two numbers are in good agreement.

[5] To be precise, the parameter ϵ introduced this way depends on the choice of phase convention and ϵ itself is therefore not an observable. The proper way to define it is $\epsilon \equiv \frac{A(K_L \longrightarrow (\pi\pi)_{I=0})}{A(K_S \longrightarrow (\pi\pi)_{I=0})} = (2.275 \pm 0.019)\,10^{-3}\; e^{i\,43.4^0}$ [1] where A is the amplitude of the process and I is the isospin of the two pion state.

2.4.3 *CP* Violation and the CKM Matrix

CP violation is not *explained* in the Standard Model. It is, however, *incorporated* via the CKM matrix. Let us, therefore, take a closer look at the properties of this matrix.

The Unitarity Triangle

The CKM matrix, being the product of two unitary matrices, is itself a unitary matrix:

$$V_{CKM} V_{CKM}^{\dagger} = 1 \tag{2.95}$$

Three independent equations of matrix elements can be extracted from this, with each equation describing a triangle in the complex $(\rho - \eta)$ plane. Numerically, two of the triangles are almost flat which makes it difficult to extract any information from them. The third one, however, has sides of the same order (see figure 2.3)

$$V_{ud} V_{ub}^{*} \; + \; V_{cd} V_{cb}^{*} \; + \; V_{td} V_{tb}^{*} \; = \; 0 \tag{2.96}$$

This triangle is called *The Unitarity Triangle* and is directly linked to both new physics and *CP* violation. If the triangle is not closed, i.e. the sum of the three angles $\alpha + \beta + \gamma \neq \pi$, then the unitarity of the CKM matrix has to be reconsidered. This would be an indication of new families. We also know that *CP* violation can only occur if $\eta \neq 0$, i.e. the triangle must not be flat. We shall later see that there are for the moment no strict constraints on the position of the apex of the triangle.

To measure *CP* violation there are two possible strategies. They can both be used to overconstrain the triangle and thus find any possible internal inconsistencies.

The angles can be measured directly from CP asymmetries using *CP* eigenstates that are common to both B_q^0 and $\overline{B_q^0}$ (q = d,s) as final states:

$$\frac{\text{Prob}\,(B_q^0 \longrightarrow f) \; - \; \text{Prob}\,(\overline{B_q^0} \longrightarrow f)}{\text{Prob}\,(B_q^0 \longrightarrow f) \; + \; \text{Prob}\,(\overline{B_q^0} \longrightarrow f)} \tag{2.97}$$

For $B_d^0 \longrightarrow \psi K_S^0$ this asymmetry is proportional to $\sin 2\beta$, for $B_d^0 \longrightarrow \pi^+\pi^-$ this gives $\sin 2\alpha$ and for $B_s^0 \longrightarrow \rho K_S^0$ one obtains $\sin 2\gamma$.

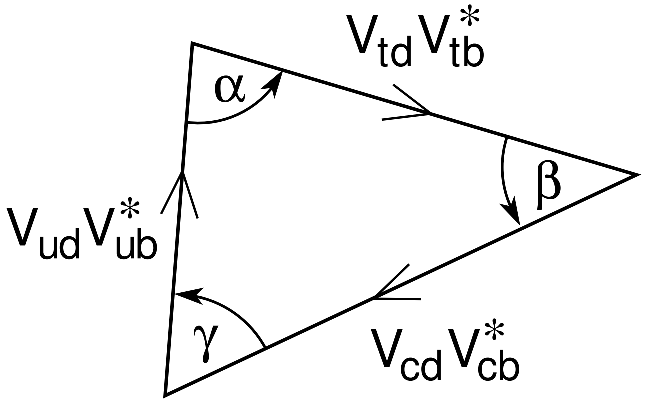
Using this method, the CDF Collaboration at Fermilab has measured the value of $\sin 2\beta$ to be [26]

$$\sin 2\beta \; = \; 0.79^{+0.41}_{-0.44} \text{ (stat. + syst.)}$$

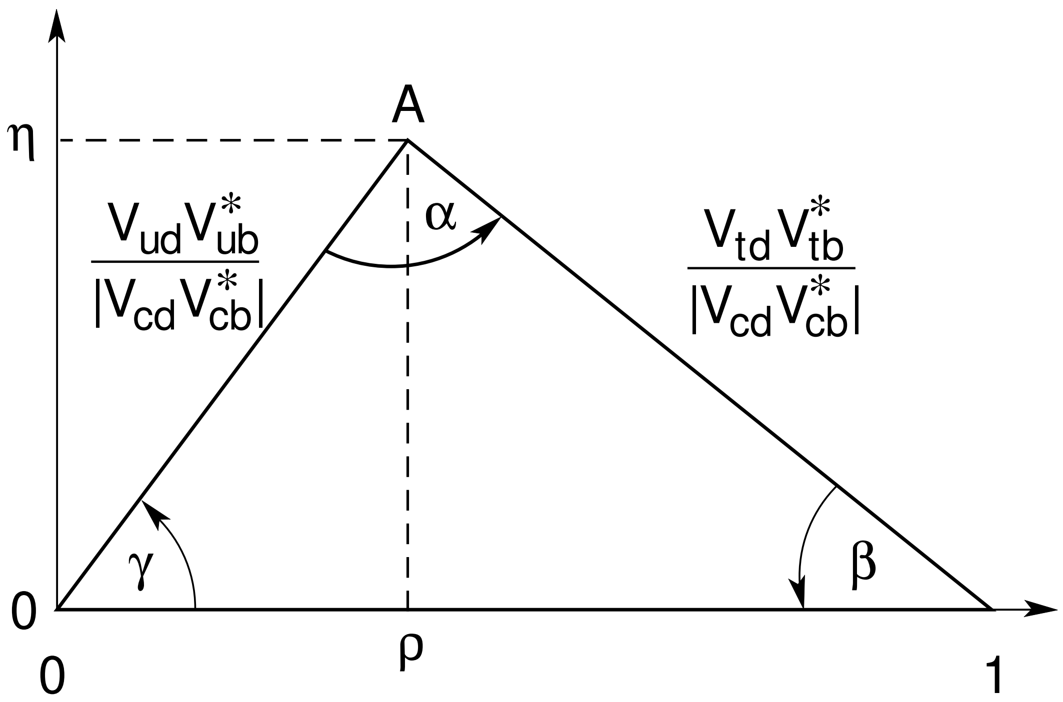
Using this value, they calculate a 93% confidence interval of

$$0.0 \; < \; \sin 2\beta \; < \; 1.00$$

The second approach is to measure the sides of the triangle. One of them is proportional to the matrix element V_{td}. This element can be obtained from $B_d^0 - \overline{B_d^0}$ mixing. Unfortunately, as we shall see, this direct extraction of V_{td} comes with rather large theoretical uncertainties. One way to reduce these uncertainties involves the matrix element V_{ts} which in an analogous way can be extracted from $B_s^0 - \overline{B_s^0}$ mixing.



(a)



(b)

Figure 2.3: *The Unitarity Triangle: The unitarity condition of the CKM matrix can be visualized as a triangle in the complex $(\rho - \eta)$ plane. Two of these triangles are expected to be nearly flat. The third is expected to have sides of the same order, however, and is shown in figure a). Figure b) shows the same triangle, but with one side normalized to one [25].*

Why *CP* Violating Effects are Expected to be Large in the B Meson System

Since the CKM matrix can be parameterized in several different ways, it is interesting to see which invariant quantities that exist. The magnitude of the coupling between the quarks, $|V_{ij}|$, is one. Another one is the imaginary part of $\lambda_{\alpha\beta ij}$ [27]:

$$\mathrm{Im}\,(\lambda_{\alpha\beta ij}) \;=\; \mathrm{Im}\,(\,V_{\alpha i}\,V^*_{\alpha j}\,V_{\beta j}\,V^*_{\beta i}) \tag{2.98}$$

where $\alpha \neq \beta,\ i \neq j$.

In the Wolfenstein parameterization this becomes

$$\mathrm{Im}\,(\,V_{ud}\,V_{tb}\,V^*_{ub}\,V^*_{td}) \;=\; A^2\lambda^6\eta \tag{2.99}$$

This is twice the area of the unitarity triangle, i.e. *CP* violating effects will be proportional to this area. Since the Wolfenstein parameterization is only approximate, however, one must be careful which CKM elements one uses. For example, Im $(V_{cb}V_{us}V^*_{ub}V^*_{cs}) = 0$.

To estimate the importance of the *CP* violating effects it will have to be scaled by the width of the process. For K mesons the dominant quark subprocess is $s \longrightarrow u$. We have $V_{us} = \lambda$ and the resonance width goes like λ^2, i.e. the *CP* violating effects are proportional to

$$A^2\lambda^4\eta$$

For D mesons the dominant decay is $c \longrightarrow s$, $V_{cs} \sim 1$ and the *CP* violating effects are proportional to

$$A^2\lambda^6\eta$$

For B mesons the dominant decay is $b \longrightarrow c$, $V_{cb} = A\lambda^2$ and the *CP* violating effects are proportional to

$$\lambda^2\eta$$

CP violating effects are thus expected to be larger for B mesons than for K, and especially D, mesons. This is, however, an 'accident' since the dominant quark decay for D mesons stays in the same family and thus have a CKM matrix element close to 1.

2.5 $B^0_q - \overline{B^0_q}$ Mixing Formalism

The formalism developed in this section is valid for any neutral particle - antiparticle system where the two are distinct.

Box Diagram

The Feynman diagrams responsible for the flavor oscillations are the second order weak box diagrams shown in figure 2.4.

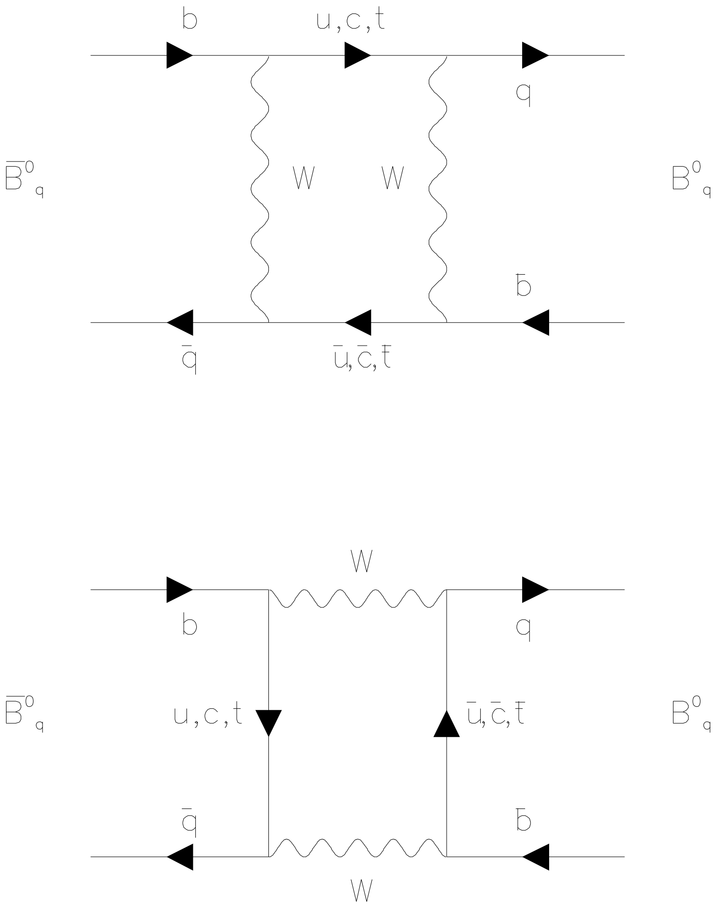


Figure 2.4: *Box diagrams contributing to B^0_q mixing. The mass splitting $\Delta m_{B^0_q}$ is given by processes involving virtual intermediate states. These are completely dominated by the top quark because of its large mass. The decay width difference $\Delta\Gamma$ is given by processes involving real intermediate states, i.e. they do not receive any contributions from the top quark.*

2.5.1 Mass and Decay Matrices

A time dependent neutral B meson state $|B^0(t)>$ can be written as a superposition of the two flavor eigenstates $|B^0_q>$ and $|\overline{B^0}_q>$:

$$|B^0(t)> = \ a(t)\,|B^0_q> \ + b(t)\,|\overline{B^0_q}> \tag{2.100}$$

In this subspace the Schrödinger equation becomes

$$\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} \ = \ \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix} \tag{2.101}$$

where the 2×2 effective Hamiltonian $\mathbf{H}$ can be split into two parts [28]:

$$\mathbf{H} \ = \ \mathbf{M} \ - \ \mathrm{i}\frac{\mathbf{\Gamma}}{2} \ = \ \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \ - \ \frac{\mathrm{i}}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \tag{2.102}$$

The mass matrix $\mathbf{M}$ and the decay matrix Γ are two Hermitian matrices since they correspond to measurable quantities. The Hamiltonian $\mathbf{H}$ is not Hermitian however. Since the particles decay, probability is not conserved. We have $M^\dagger = M$, $\Gamma^\dagger = \Gamma$ and $H^\dagger \neq H$.

The mass matrix elements M_{11} and M_{22} are due to the mass of the constituents and their binding energy and correspond to the mass of the two flavor eigenstates. The diagonal elements of the decay matrix Γ are the decay widths of the flavor eigenstates.

The mass matrix $\mathbf{M}$, being the dispersive part of the Hamiltonian, is determined by virtual intermediate states. Because of the large top mass this means that the box diagram with the t quark is dominant and other quark contributions can be neglected. The decay matrix Γ, being the absorptive part of the Hamiltonian, comes from real intermediate states. This means that there are no contributions to the off-diagonal elements of Γ from intermediate states containing the top quark. Because of the large top mass and the large value of the b to t quark coupling, we have $|\Gamma_{12}| \ll |M_{12}|$. The lifetime difference is much smaller than the mass difference. This is the opposite of the situation in the neutral kaon system.

CPT Invariance

The invariance under the combined operation CPT has some consequences for the matrix elements. We have

$$< B^0_q|H|f> = < B^0_q|(CPT)^\dagger H(CPT)|f> = \ \eta_{CP} \ < \overline{B^0_q}|T^\dagger HT|\bar{f}> = \ \eta_{CP} \ < \bar{f}|H|\overline{B^0_q}> \tag{2.103}$$

where η_{CP} is a phase factor.

By choosing $|f>=|B^0_q>$ the diagonal elements satisfy

$$\begin{aligned} M_{11} &= M_{22} &&= m_{B^0_q} \\ \Gamma_{11} &= \Gamma_{22} &&= \Gamma_{B^0_q} \end{aligned} \tag{2.104}$$

By choosing $|f>=|\overline{B^0_q}>$ the off-diagonal elements satisfy

$$
\begin{aligned}
M_{21} &= M_{12}^* \\
\Gamma_{21} &= \Gamma_{12}^*
\end{aligned}
\tag{2.105}
$$

If only invariance under CP is required, this becomes

$$
\begin{aligned}
M_{21} &= M_{12} \\
\Gamma_{21} &= \Gamma_{12}
\end{aligned}
\tag{2.106}
$$

with all four matrix elements real.

2.5.2 The Physical States

To obtain the weak eigenstates, the Hamiltonian H has to be diagonalized. Calling the two eigenstates B_{qL}^0 and B_{qH}^0 (*Light* and *Heavy*), they can be expressed in the basis of the flavor eigenstates by

$$
\begin{aligned}
|B_{qL}> &= \frac{(1+\epsilon)\,|B_q^0> \;+\;(1-\epsilon)\,|\overline{B_q^0}>}{\sqrt{2(1+|\epsilon|^2)}} \\
|B_{qH}> &= \frac{(1+\epsilon)\,|B_q^0> \;-\;(1-\epsilon)\,|\overline{B_q^0}>}{\sqrt{2(1+|\epsilon|^2)}}
\end{aligned}
\tag{2.107}
$$

where the complex parameter ϵ expresses the degree of CP violation.

The eigenstates $|B_{qL}>$ and $|B_{qH}>$ have well defined masses and decay widths:

$$
\begin{aligned}
M_{H,L} &= M_{11} \;\pm\; \mathrm{Re}\sqrt{(\mathrm{M}_{12}-\frac{\mathrm{i}}{2}\Gamma_{12})(\mathrm{M}_{12}^*-\frac{\mathrm{i}}{2}\Gamma_{12}^*)} \\
\Gamma_{H,L} &= \Gamma_{11} \;\mp\; 2\mathrm{Im}\sqrt{(\mathrm{M}_{12}-\frac{\mathrm{i}}{2}\Gamma_{12})(\mathrm{M}_{12}^*-\frac{\mathrm{i}}{2}\Gamma_{12}^*)}
\end{aligned}
\tag{2.108}
$$

The mass and decay width differences then become

$$
\begin{aligned}
\Delta m &= M_H \;-\; M_L \;=\; 2\mathrm{Re}\sqrt{(\mathrm{M}_{12}-\frac{\mathrm{i}}{2}\Gamma_{12})(\mathrm{M}_{12}^*-\frac{\mathrm{i}}{2}\Gamma_{12}^*)} \\
\Delta\Gamma &= \Gamma_H \;-\; \Gamma_L \;=\; -4\mathrm{Im}\sqrt{(\mathrm{M}_{12}-\frac{\mathrm{i}}{2}\Gamma_{12})(\mathrm{M}_{12}^*-\frac{\mathrm{i}}{2}\Gamma_{12}^*)}
\end{aligned}
\tag{2.109}
$$

CP Conservation

In the case of CP conservation (i.e. M_{12} and Γ_{12} real), $\epsilon = 0$ and the weak eigenstates also become CP eigenstates:

$$|B_{qL}> = \frac{|B_q^0> \; + \; |\overline{B_q^0}>}{\sqrt{2}}$$
$$|B_{qH}> = \frac{|B_q^0> \; - \; |\overline{B_q^0}>}{\sqrt{2}} \tag{2.110}$$

where

$$CP|B_{qL}> = \; +|B_{qL}>$$
$$CP|B_{qH}> = \; -|B_{qH}> \tag{2.111}$$

The CP even state is the lightest state.

2.5.3 Time Evolution of B_q^0 Mesons

The B mesons are produced by the strong interaction which means that initially we have either a $|B_q^0>$ or a $|\overline{B_q^0}>$. Starting out with a $|B_q^0>$, and inverting equation (2.107) to express the flavor eigenstates in the basis of the weak eigenstates, we can write the initial condition as

$$|B(t=0)> = \; |B_q^0> = \; \frac{\sqrt{2(1+|\epsilon|^2)}}{2(1+\epsilon)} \; (|B_{qH}^0> \; + |B_{qL}^0>) \tag{2.112}$$

At time t we have

$$|B(t)> = \; \frac{\sqrt{2(1+|\epsilon|^2)}}{2(1+\epsilon)} \; (e^{-i(m_H - \frac{i}{2}\Gamma_H)t}|B_{qH}^0> \; + (e^{-i(m_L - \frac{i}{2}\Gamma_L)t}|B_{qL}^0>) \tag{2.113}$$

In the basis of the flavor eigenstates this becomes

$$|B(t)> = \; A_{B_q^0 \longrightarrow B_q^0}(t)|B_q^0> \; + A_{B_q^0 \longrightarrow \overline{B_q^0}}(t)|\overline{B_q^0}> \tag{2.114}$$

where we have introduced the transition amplitudes

$$A_{B_q^0 \longrightarrow B_q^0}(t) = \quad \tfrac{1}{2}\,(e^{-i(m_H - \frac{i}{2}\Gamma_H)t} \; + \; e^{-i(m_L - \frac{i}{2}\Gamma_L)t})$$
$$A_{B_q^0 \longrightarrow \overline{B_q^0}}(t) = \; \tfrac{1}{2}\,(e^{-i(m_H - \frac{i}{2}\Gamma_H)t} \; - \; e^{-i(m_L - \frac{i}{2}\Gamma_L)t})\,\tfrac{1-\epsilon}{1+\epsilon} \tag{2.115}$$

The time dependent probability that a B_q^0 oscillates into a $\overline{B_q^0}$ is then given by

$$\mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t) \; = \; |<\overline{B_q^0}|B(t)>|^2 \; = \; |A_{B_q^0 \longrightarrow \overline{B_q^0}}(t)|^2 \tag{2.116}$$

This implies that

$$\mathcal{P}_{B_q^0 \longrightarrow B_q^0}(t) \; = \; \frac{1}{2}e^{-\Gamma_{B_q^0}t}\;(\cosh\frac{\Delta\Gamma_{B_q^0}}{2}t \; + \; \cos\Delta m_{B_q^0}t) \; = \; \mathcal{P}_{\overline{B_q^0} \longrightarrow \overline{B_q^0}}(t) \tag{2.117}$$

and

$$
\begin{aligned}
\mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t) &= \frac{1}{2}\left|\frac{1-\epsilon}{1+\epsilon}\right|^2 e^{-\Gamma_{B_q} t}\,(\cosh\frac{\Delta\Gamma_{B_q^0}}{2}t \;-\; \cos\Delta m_{B_q^0}\,t) \\
\mathcal{P}_{\overline{B_q^0} \longrightarrow B_q^0}(t) &= \frac{1}{2}\left|\frac{1+\epsilon}{1-\epsilon}\right|^2 e^{-\Gamma_{B_q} t}(\,\cosh\frac{\Delta\Gamma_{B_q^0}}{2}t \;-\; \cos\Delta m_{B_q^0}\,t)
\end{aligned}
\tag{2.118}
$$

where

$$
\Gamma \;=\; \frac{\Gamma_L \;+\; \Gamma_H}{2} \tag{2.119}
$$

is the mean decay width.

For the neutral kaons the parameter ϵ was an indication of CP violation having its origin in the particle-antiparticle mixing. Here it can be seen explicitly. For $\epsilon \neq 0$ we have

$$
\mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t) \;\neq\; \mathcal{P}_{\overline{B_q^0} \longrightarrow B_q^0}(t) \tag{2.120}
$$

Time Integrated Mixing

The time integrated probability for a B_q^0 meson to oscillate before it decays is defined as

$$
\chi_q \;=\; \frac{\int_0^\infty \mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t)\;dt}{\int_0^\infty \mathcal{P}_{B_q^0 \longrightarrow B_q^0}(t)\;dt \;+\; \int_0^\infty \mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t)\;dt} \tag{2.121}
$$

Neglecting any CP violation effects this becomes

$$
\chi_q \;=\; \frac{x_q^2 \;+\; y_q^2}{2\,(1+x_q^2)} \tag{2.122}
$$

where we have introduced the two dimensionless quantities

$$
x_q \;=\; \frac{\Delta m_{B_q^0}}{\Gamma_{B_q^0}} \qquad y_q \;=\; \frac{\Delta\Gamma_{B_q^0}}{2\Gamma_{B_q^0}} \tag{2.123}
$$

x_q is the parameter that quantifies the mixing.

For B_d^0 oscillations one can neglect y_d with respect to x_d. And even if the decay width difference in the B_s^0 system is not negligible, y_s is still small compared to x_s [1], [29]. Neglecting CP violation, this gives the approximate oscillation formulas

$$
\begin{aligned}
\mathcal{P}_{B_q^0 \longrightarrow B_q^0}(t) \;=\; \mathcal{P}_{\overline{B_q^0} \longrightarrow \overline{B_q^0}}(t) &= \frac{1}{2}\,\Gamma_{B_q^0}\,e^{-\Gamma_{B_q^0} t}\,(1 \;+\; \cos\Delta m_{B_q^0}\,t) \\
\mathcal{P}_{B_q^0 \longrightarrow \overline{B_q^0}}(t) \;=\; \mathcal{P}_{\overline{B_q^0} \longrightarrow B_q^0}(t) &= \frac{1}{2}\,\Gamma_{B_q^0}\,e^{-\Gamma_{B_q^0} t}\,(1 \;-\; \cos\Delta m_{B_q^0}\,t)
\end{aligned}
\tag{2.124}
$$

and

$$
\chi_q \;=\; \frac{x_q^2}{2\,(1+x_q^2)} \tag{2.125}
$$

If we do not distinguish between a B_q^0 and a $\overline{B_q^0}$ we recover the usual exponential formula for a decay (for this purpose they have been normalized with $\Gamma_{B_q^0}$).

From the measurement of the time integrated oscillation probability we can extract x_q and, therefore, $\Delta m_{B_q^0}$. This will only work if the mass difference is not too large. For large values of x_q, χ_q goes very quickly to 0.5 and at a certain point it looses all sensitivity (within any reasonable experimental resolution) to the exact value of x_q. As we shall later see, the B_d^0 meson belongs to the first category while in the B_s^0 system the mass difference is expected to be much too large to enable us to use the time integrated oscillation probability.

Asymmetries

From the mixing probabilities one can form two asymmetries:

$$N^{\text{no osc}} - N^{\text{osc}} \propto e^{-\Gamma t} \cos\left(\Delta m_{B_q^0} t\right) \tag{2.126}$$

$$\frac{N^{\text{no osc}} - N^{\text{osc}}}{N^{\text{no osc}} + N^{\text{osc}}} = \cos\left(\Delta m_{B_q^0} t\right) \tag{2.127}$$

In principle, a simple histogram containing the relative difference could reveal the oscillation frequency. In a real experimental situation, however, resolution and background effects will make this more difficult, especially for the rapid $\Delta m_{B_s^0}$ oscillations.

2.5.4 Lifetime Differences in the B-System

As was seen in the neutral kaon system, the lifetime difference was generated by (the approximate) *CP* conservation. The fact that it was so large was just a kinematic 'accident', however. In the neutral B meson the lifetime difference is expected to be much smaller since there are many final states [1] and the available phase space is large.

If we now only consider the B_s^0 meson system, it can be showed [29] that the dominant quark subprocess leading to final states common to both flavor eigenstates is

$$b(\bar{s}) \longrightarrow c\bar{c}s(\bar{s})$$

which is predominantly *CP* even, i.e. it is the *CP* even state which has the shortest lifetime. It can be calculated [29] that

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} = 0.16^{+0.11}_{-0.09}$$

The largest contribution to the uncertainty comes from the uncertainty of the hadronic matrix elements. This estimate is quite similar to one obtained by summing over exclusive decay modes into ground state mesons [30].

The same argument can also be used for B_d^0 mesons, but in that case the corresponding decay is $b(\bar{d}) \longrightarrow c\bar{c}d(\bar{d})$, i.e. it would include the CKM matrix element V_{cd} instead of V_{cs}. This would already reduce the dominant decay by a factor ~5 (which subsequently is squared). The subprocess $b(\bar{d}) \longrightarrow u\bar{u}d(\bar{d})$ is even more suppressed.

One interesting relation is that the decay width difference is proportional to the mass difference. From reference [29] we have for the B_s^0 meson

$$\frac{\Delta\Gamma_{B_s^0}}{\Delta m_{B_s^0}} = (5.6 \pm 2.6)\, 10^{-3}$$

This means that if the oscillation frequency is too high to be measured experimentally the lifetime difference would also be large and could maybe be accessible. One could then use this relation to extract $\Delta m_{B_s^0}$. Unfortunately, the uncertainty on this ratio is still very high with most of it coming from non-perturbative effects. It is not obvious that this will improve sufficiently in the near future.

2.6 Theoretical Calculation of $\Delta m_{B_q^0}$

The mass difference is calculated to be [31]

$$\Delta m_{B_q^0} = \frac{G_F^2}{6\pi^2}\, \eta_B\, m_{B_q}\, m_W^2\, f_B^2 B_{B_q}\, |V_{tq}|^2\, |V_{tb}|^2\, S(x_t) \tag{2.128}$$

where:

- G_F is the Fermi coupling constant.
- $m_{B_q^0}$ is the B meson mass.
- $S(x_t)$, where $x_t = \frac{m_t^2}{m_W^2}$, is a smooth function coming from the evaluation of the box diagram without QCD corrections.
- η_B is a calculable QCD correction factor.
- m_t is the running top mass.
- f_B^2 is the B meson *decay constant*, a measure of the overlap of the quark wave functions in the meson, and contains all the QCD effects between the quarks. In leptonic decays of a pseudoscalar meson, like the B meson, one can show that only the axial part of the current contributes and that the hadronic matrix elements (containing all the QCD effects) is proportional to p^μ, i.e. $< 0|\bar{b}\,\gamma^\mu\,\gamma^5\,q|B(p)> = \; i f_B p^\mu$. The constant of proportionality is called the decay constant.
- B_{B_q} is the *bag factor.* All the non-perturbative QCD effects responsible for the particle-antiparticle mixing are contained in the $\Delta B = 2$ operator

  $$O^{|\Delta B=2|} = \bar{b}\gamma^\mu(1-\gamma^5)q\,\bar{b}\gamma_\mu(1-\gamma^5)q \tag{2.129}$$

  Ideally, we can neglect soft gluon emission between initial and final quarks. This is called the *vacuum saturation approximation.* Even though this approximation is better the heavier the meson mass, it is not perfect, and we, therefore, have to add a correction factor called the *Bag factor.*

  In the kaon system one can measure the kaon decay constant. In the B meson system this does not work. The problem is that the leptonic decay also involves a CKM matrix element and while V_{us} is known to a reasonable precision the precision on V_{ub} is far from satisfactory. This means that both the decay constant and the bag factor have to be evaluated with other means, notably lattice QCD. For the moment these two quantities are calculated with large uncertainties. We have [32]:

$$\begin{aligned} f_{B_d} &= (170 \pm 35)\ \mathrm{GeV} \\ f_{B_s} &= (195 \pm 35)\ \mathrm{GeV} \\ B_{B_d} &= 1.4\,(1) \end{aligned}$$

The large theoretical uncertainties make it difficult to extract V_{td} from the measurement of the mass difference $\Delta m_{B_d^0}$. However, one can reduce these by taking the ratio

$$\frac{\Delta m_{B_s^0}}{\Delta m_{B_d^0}} = \frac{m_{B_s^0}}{m_{B_d^0}} \frac{B_{B_s^0} f_{B_s^0}^2}{B_{B_d^0} f_{B_d^0}^2} \left|\frac{V_{ts}}{V_{td}}\right| \tag{2.130}$$

Up to SU(3) breaking effects the ratio $(\sqrt{B_{B_s^0}}\, f_{B_s^0})/(\sqrt{B_{B_d^0}}\, f_{B_d^0}) = 1$, i.e. the uncertainty on this ratio will be reduced with respect the the uncertainty on the individual terms. We have [32]

$$\frac{\sqrt{B_{B_s^0}}\, f_{B_s^0}}{\sqrt{B_{B_d^0}}\, f_{B_d^0}} = 1.14\,(8)$$

This means that by measuring $\Delta m_{B_d^0}$ and $\Delta m_{B_s^0}$ one can get the ratio of the CKM matrix elements $|V_{ts}/V_{td}|$ with a reasonable precision.

$\Delta m_{B_d^0}$ has already been measured to a high precision. This leaves $\Delta m_{B_s^0}$.

2.7 Numerical Values

$B_d^0 - \overline{B_d^0}$ System

Oscillations in the $B_d^0 - \overline{B_d^0}$ system have been observed in several experiments. Both time integrated and time dependent measurements have been obtained [1]:

$$\begin{aligned} \Delta m_{B_d^0} &= (0.464\ \pm\ 0.018)\ \mathrm{ps}^{-1} \\ \chi_{B_d^0} &= 0.172\ \pm\ 0.010 \end{aligned}$$

$B_s^0 - \overline{B_s^0}$ System

As already mentioned, a time integrated measurement is expected to be difficult to make. The current limit is [1]

$$\chi_{B_s^0} > 0.4975,\ \ \mathrm{C.L.} = 95\%$$

No time dependent measurement exists for the moment, only lower limits from OPAL [33], [34]; ALEPH [35], [36], [37], [38]; DELPHI [39], [40], [41], [42], [43]; CDF [44] and SLD [45]. Combining the results from the five experiments one obtains [46]

$$\Delta m_{B_s^0} > 12.4\,\mathrm{ps}^{-1}$$

There exist some upper limits on the relative decay width difference [43], [47], [48], but they are still far from the expected value.

Figure 2.5 shows the oscillating and non-oscillating components for oscillation frequencies of 0.5 ps^{-1} (B_d^0-type) and 15.0 ps^{-1} (B_s^0-type). The difference between the two cases is clearly seen. Experimental proper time resolutions, background etc will have the effect of washing out and damping the oscillations, making the observation of rapid oscillations even more difficult.

2.8 Constraints on the Unitarity Triangle

We see in figure 2.3 that the side given by the point A and (1,0) depends on the matrix element V_{td}. This means that a measurement of V_{td} will constrain the position of the point A to a circle centered at the point (1,0). A lower limit on the mass difference $\Delta m_{B_s^0}$ will (using the measured value of $\Delta m_{B_d^0}$ and the ratio of the two mass differences) limit the point A to inside the circle. A detailed discussion can be found in reference [25].

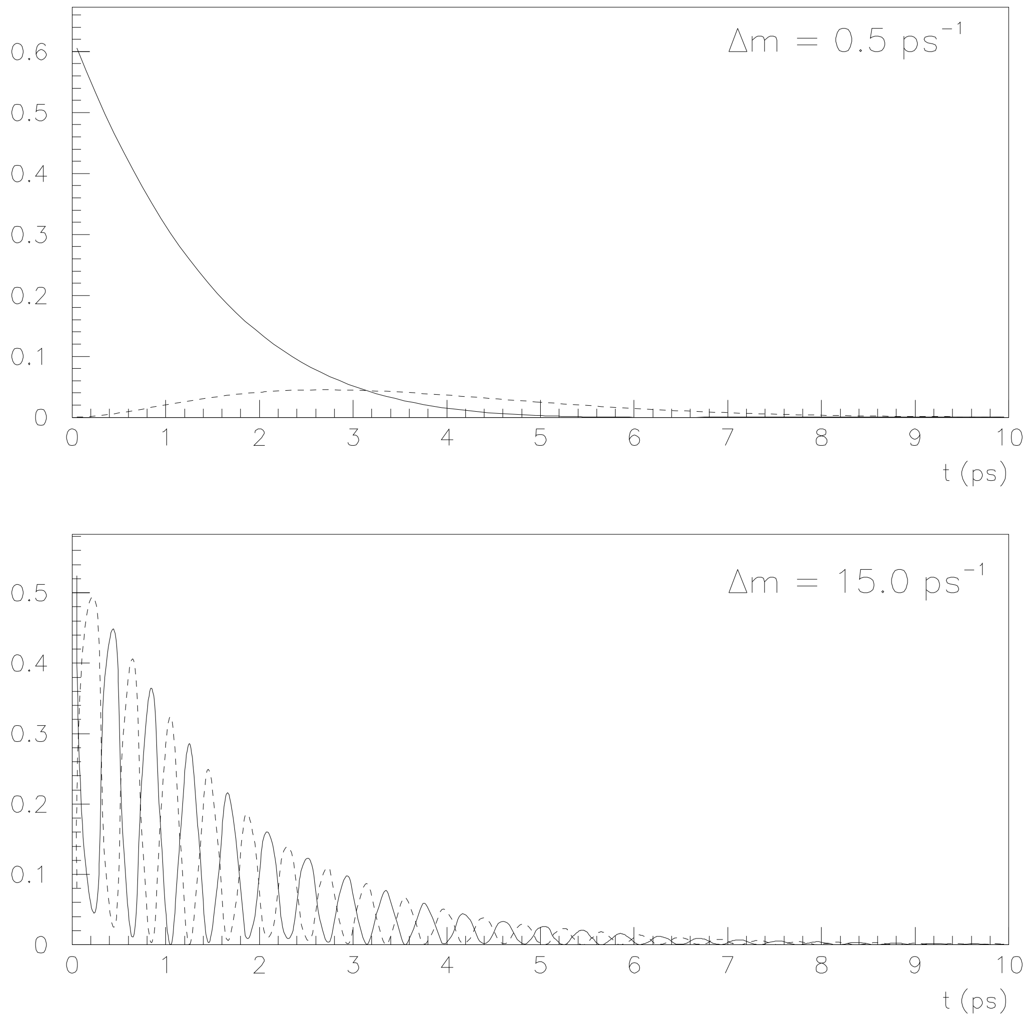


Figure 2.5: *Non-oscillating (continuous line) and oscillating (dashed line) components for the two frequencies 0.5* ps^{-1} *(B_d^0-type, upper plot) and 15.0* ps^{-1} *(B_s^0-type, lower plot).*

Chapter 3

The DELPHI Experiment

3.1 Detector Description

DELPHI (DEtector with Lepton, Photon and Hadron Identification) is a general purpose, high granularity detector with powerful particle identification capabilities over the whole solid angle. It is one of the four detectors installed at the Large Electron Positron (LEP) collider which during the years 1989-1995 operated at the Z^0 resonance.

This analysis depends on precisely reconstructed decay vertices, highest possible track reconstruction efficiency and good lepton and hadron identification. The relevant subdetectors will be described in the following. Due to the necessity of being inside the vertex detector coverage in general, and inside the coverage of the barrel electromagnetic calorimeter HPC for electron identification, the following detector description is limited to the barrel region (approximately $40^0 < \theta < 135^0$). Additional details can be found in references [49] and [50]. Unless stated otherwise, all quoted numbers are from these two references. Figure 3.1 gives a schematic layout of the detector.

The DELPHI Detector Setups

In 1994 the vertex detector was upgraded to give z information for two of the three layers. The same year the full ring imaging Cherenkov detector became operational. This means that the DELPHI detector can be divided into two fairly separate detector setups corresponding to 1989-1993 and 1994-1995. Starting in 1995 DELPHI has had major upgrades to prepare for the LEP2 phase. This has especially concerned the forward region. These improvements will not be described here.

The DELPHI Coordinate System

In the DELPHI coordinate system the z axis is along the electron direction, the x axis points towards the center of LEP while the y axis points upwards. The polar angle with respect to the z axis is called θ while ϕ is the azimuthal angle around the z axis. The radial coordinate is $R = \sqrt{x^2 + y^2}$.

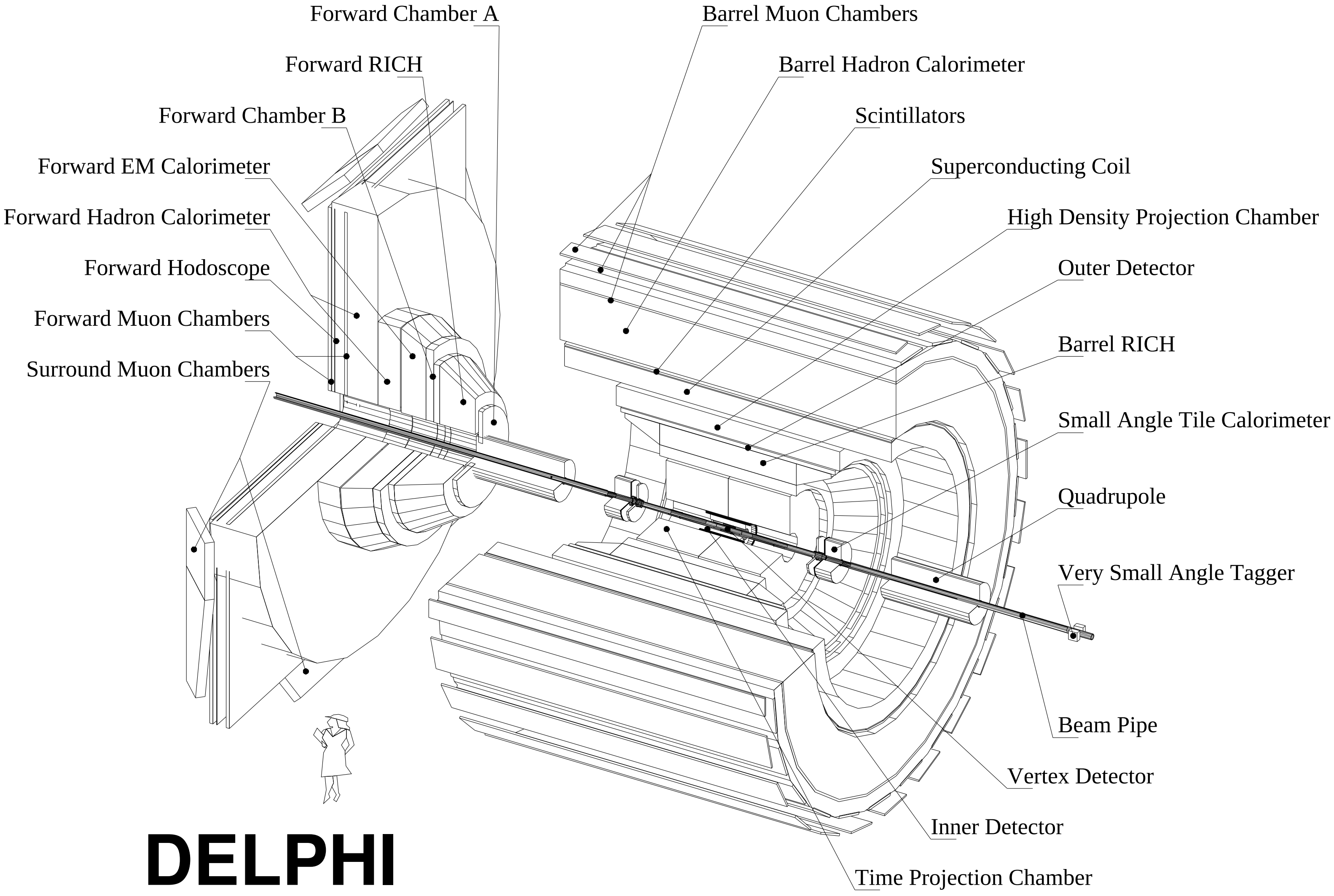


Figure 3.1: *Schematic layout of the DELPHI detector.*

3.1.1 Tracking Detectors

The tracking volume in DELPHI is placed inside a uniform magnetic field of 1.23 T provided by a superconducting solenoid. The field is parallel to the z axis. The main tracking detector is the time projection chamber. In the barrel region additional tracking informations is given by the vertex detector, the inner detector and the outer detector.

Vertex Detector VD

The VD has three cylindrical layers of silicon strip detectors at average radii of 6.3 cm, 9.0 cm and 10.9 cm. The full azimuth angle is covered by each layer with 24 partially overlapping sectors. For angles $44^0 < \theta < 136^0$ a particle crosses all three layers. In 1994 the layers at 6.3 cm and 10.9 cm were equipped with double-sided silicon detectors giving z information. The single hit precision in $R\phi$ is 7.6 μm for one layer. In z the single hit precision depends on the incidence angle of the track. It reaches a minimum of 9 μm.

Inner Detector ID

The ID is made of an inner drift chamber with a jet-chamber geometry made of 24 azimuthal sectors. Each sector provides a maximum of 24 $R\phi$ points per track between a radius of 12 cm and 23 cm. Surrounding the jet chamber there are five Multi Wire Proportional Chamber (MWPC) trigger layers which provide additional Rz information via a cathode readout. The $R\phi$ measurements are not only used in the trigger, but can also solve the left-right ambiguities in the jet chamber. The polar angle coverage is $30^0 < \theta < 150^0$. The ID was upgraded in 1995 with a new and longer version that increases the polar angle coverage down to 15^0. Five cylindrical layers of straw tube detectors were installed instead of the MWPC. They do not give any z information.

Time Projection Chamber TPC

The TPC is the main tracking detector in DELPHI and is placed just outside the ID. The overall shape is that of a cylinder of length 2×1.5 m and radius 120 cm. The total drift length is 1.34 m, and the drift gas is a mixture of 80% Ar and 20% CH_4 operating in 1 atm and 29^0C. The drift field is 187 V cm^{-1}, giving a drift velocity for primary electrons of 7 cm μs^{-1}. Each endplate is divided into 6 sectors in the azimuthal angle. Each sector has 16 circular pad rows, thus giving 16 space points per particle trajectory. The full polar angle coverage is $39^0 < \theta < 141^0$. The single point resolution for tracks from $Z \longrightarrow \mu^+\mu^-$ decays is 250 μm in $R\phi$ and 880 μm in Rz. Due to distortions the precision on track elements is limited to 150 μm and 600 μm in $R\phi$ and z, respectively.

Outer Detector OD

The OD is made of 5 layers of drift tubes which operate in the limited streamer mode. They are located between radii of 197 cm and 206 cm. It gives full coverage in the azimuthal angle and a polar angle coverage of $42^0 < \theta < 138^0$. Z coordinate information can be obtained by using the timing of the signals at the ends of the anode wires. This only concerns three of the layers. The single hit precision in $R\phi$ is 110 μm while the precision in z is 3.5 cm.

3.1.2 Momentum Resolution

In DELPHI track reconstruction starts locally with independent pattern recognition and subsequent formation of track elements in the different tracking subdetectors. These are then merged at the global level. The main tracking algorithm in the barrel starts in the TPC and then extrapolates inwards to the ID and outwards to the OD to build up the candidate strings. Because of dead zones in the TPC leading to lost tracks, DELPHI developed new algorithms in 1994-1995 that try to recover the tracks not seen in the TPC. The 1992-1995 data set was subsequently reprocessed using these new algorithms. The final track fit is based on Kalman filter techniques [51] and takes into account multiple scattering and energy loss in the material between measurements. It is only at this stage that successful track candidates are associated to the VD hits and the track parameters are refitted.

The final track reconstruction precision depends on the alignment of the detector. The starting point is optical and mechanical surveys of the OD. Then beam energy constrained tracks from the decay $Z \longrightarrow \mu^+\mu^-$ are used to align the VD. The ID and the TPC are then aligned by the same procedure. The final momentum resolution determined from 45.6 GeV/c muons seen by all four tracking detectors in the barrel is

$$\sigma(1/p) = 0.06\%$$

where the momentum p is in GeV/c. The corresponding resolution when the track is not seen in the OD is 0.17%. Figure 3.2 shows the obtained precision on the different track parameters as obtained from simulated hadronic Z^0 decays.

3.2 Particle Identification

This section describes the particle identification which is of particular interest for this analysis, namely the electron, muon, kaon and pion identification. It also gives a short description of the relevant detectors. Additional details and alternative identification methods can be found in [50].

3.2.1 Electron Identification

Electron identification in DELPHI is complicated by the large amount of matter in front of the electromagnetic calorimeters. To ensure a good electron identification only the angular region covered by the barrel electromagnetic calorimeter (HPC) was used. Even here the amount of material corresponds to $0.8/\sin\theta$ radiation lengths.

The Electromagnetic Calorimeter HPC

The High density Projection Chamber is the TPC principle applied to calorimetry to obtain high granularity. The HPC is a lead/gas sampling calorimeter made of 144 modules in six rings inside the magnetic field and is placed outside the OD. Each module is a small TPC detector with lead wires in the gas giving detailed three dimensional information about the charge distribution. This enables the HPC to distinguish high energetic photon pairs coming from π^0s from single photons. The wires provide the drift field and serve as converters. The total converter thickness is $18/\sin\theta$ radiation lengths. The angular precision is ±1.0 mrad in ϕ and ±1.7 mrad in θ. The energy resolution can be parameterized as

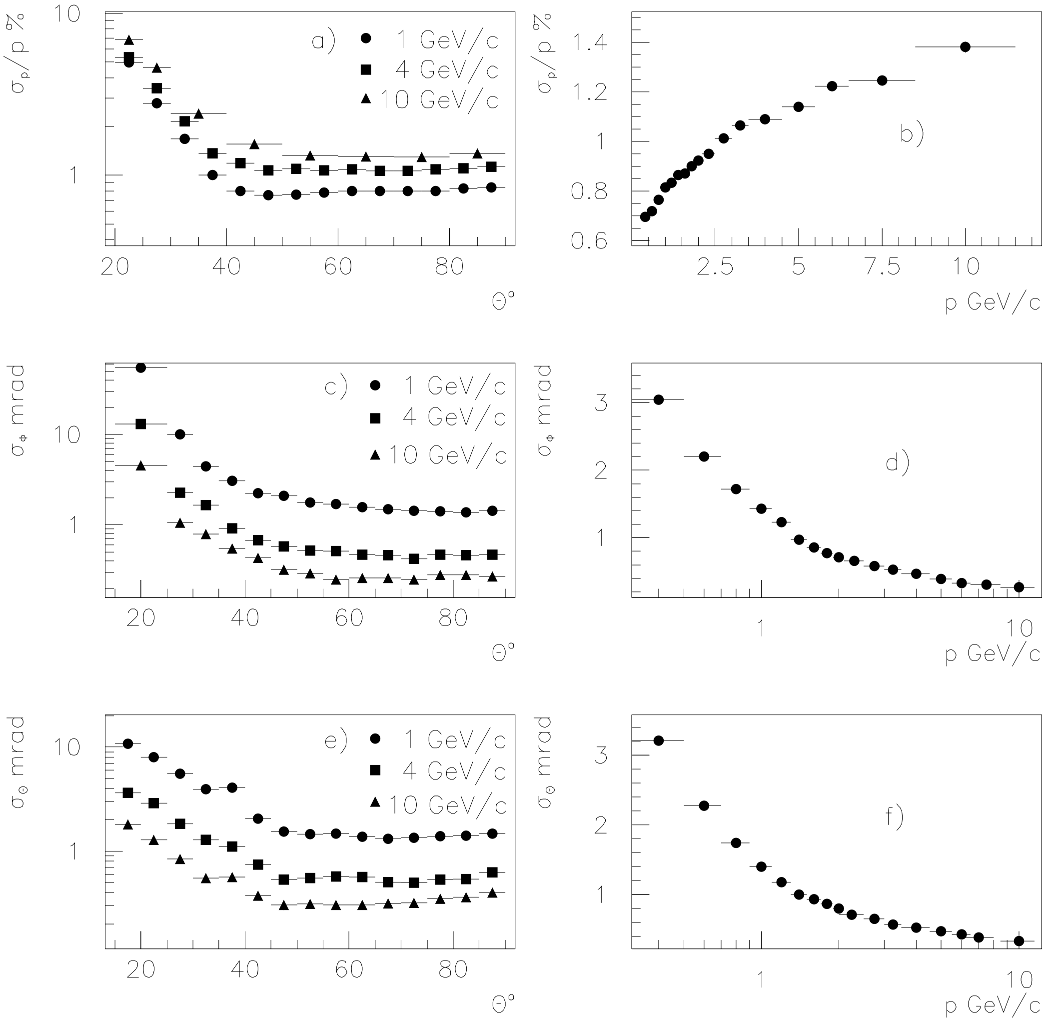


Figure 3.2: *Track parameter precision [50]: Figures a), c) and e) show the momentum precision and the precision on the the azimuthal angle ϕ and polar angle θ, respectively, versus the polar angle. Figures b), d) and f) show the same three quantities versus the track momentum p for barrel tracks.*

$$\sigma_E/E = 0.043 \otimes 0.32/\sqrt{E}$$

where the energy E is in GeV.

Electron Identification

Electron identification is based on two independent measurements. The energy deposit in the HPC gives four contributions:

- The ratio $E_{HPC}/p_{\mathrm{track}}$: By measuring the energy in the HPC and, independently, the track momentum from the tracking detectors, one can form the ratio of the two quantities. It is expected to be close to unity for electrons.
- The difference Z(extrapolation) - Z(shower): The track is extrapolated from the TPC to the HPC and the mismatch in z between this extrapolation and the shower position measured in the HPC separates electrons from pions.
- The difference ϕ(extrapolation) - ϕ(shower): Same principle as for the z mismatch.
- The HPC shower shape: The shower shape in the HPC can be parameterized by a Γ-distribution giving a χ^2-like variable that can quantify the electromagnetic character of a shower [52].

The last contribution comes from the dE/dX measurement in the TPC. For momenta below 4.5 GeV/c this gives a 3σ separation between electrons and pions in jets. This decreases to 2σ for momenta between 4.5 GeV/c and 20 GeV/c.

For all five contributions a probability for the electron hypothesis is derived [52]. Pions tend to be concentrated near zero while the electrons are distributed closer to one. They can thus be used to separate electrons from pions. The most important contributions come from the fit to the HPC shower shape and the dE/dX measurement.

Three different tagging levels are given [50]: *loose*, *standard* and *tight*, corresponding to different requirements on efficiency and purity. The identification efficiencies and (pion) misidentification probabilities for electrons with momentum above 3 GeV/c in multihadronic events are given in table 3.1. It should be noted that in an average event there are about 21 tracks of which 17 are pions. Figure 3.3 shows the electron identification efficiency and purity for the standard tag.

Tagging level	Efficiency (in %)	Misidentification probability (in %)
Loose	80	≈ 1.6
Standard	55	≈ 0.4
Tight	45	≈ 0.2

Table 3.1: *Electron identification and (pion) misidentification probabilities for electrons with momentum above 3 GeV/c in multihadronic events.*

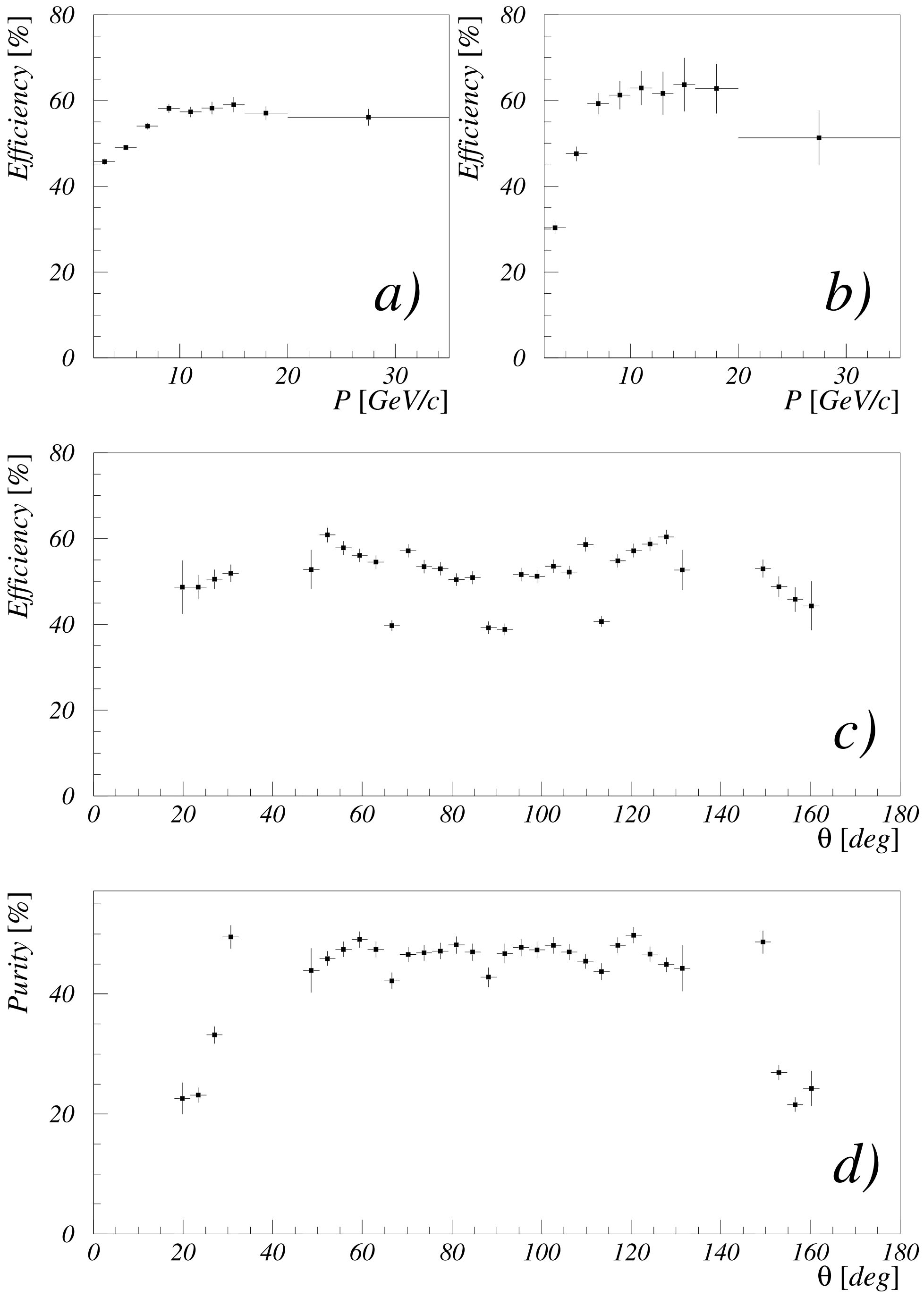


Figure 3.3: *Electron identification efficiency and purity for the standard tag [50]: Plots a) and b) show the electron identification efficiency for the* standard *tag versus the track momentum in the barrel region and the forward region, respectively. Plots c) and d) show the efficiency and purity, respectively, versus the polar angle θ.*

3.2.2 Muon Identification

The first level of the muon identification is given by the iron of the hadron calorimeter (HCAL). This stops most of the hadrons while nearly all muons with momentum above 2 GeV/c enter the muon chambers. However, due to punch-throughs from hadronic showers, decays in flight and particles going through the HCAL sector boundaries, HCAL information alone is not enough to identify muons. To improve the discrimination, the DELPHI muon identification uses the extrapolation of the candidate track and the signals in the muon chambers. A χ^2 comparison is then used to assign a muon tag level to the candidate [50]: *very loose*, *loose*, *standard* and *tight* . Table 3.2 gives the muon identification efficiencies and misidentification probabilities for the different tags. The numbers have been determined using tracks with momentum above 3 GeV/c from $Z \longrightarrow \mu^+\mu^-$ and $Z \longrightarrow \tau^+\tau^-$, $\tau \longrightarrow 3\pi\nu_\tau$ events. Figure 3.4 shows the muon identification efficiency for the standard tag.

Tagging level	Efficiency (in %)	Misidentification probability (in %)
Very loose	95.9 ± 0.1	5.4 ± 0.2
Loose	94.8 ± 0.1	1.5 ± 0.1
Standard	86.1 ± 0.2	0.7 ± 0.1
Tight	76 ± 0.2	0.4 ± 0.1

Table 3.2: *Muon identification and misidentification probabilities for muons with momentum above 3 GeV/c. The numbers have been obtained using tracks from $Z \longrightarrow \mu^+\mu^-$ and $Z \longrightarrow \tau^+\tau^-$, $\tau \longrightarrow 3\pi\nu_\tau$ events.*

3.2.3 Charged Hadron Identification

Charged hadrons in DELPHI are identified by the RICH detectors and dE/dX in the TPC.

The Barrel Ring Imaging Cherenkov Detector BRICH

Charged particles traversing a medium with a velocity that is larger than the local velocity of light emit photons in a cone with half-angle θ_c [1] (see figure 3.5). The angle θ_c is called the Cherenkov angle. We have the following relations:

$$\cos\theta_c = \frac{1}{\beta\, n} \tag{3.1}$$

$$N_{\text{photons}} = N_0 \sin^2\theta_c \tag{3.2}$$

where β is the velocity of the track, n the refractive index of the medium and N the number of observed photons. By using the momentum as given by the tracking detectors in combination with the velocity found from the Cherenkov angle, the mass of the particle can be found, thus revealing its identity.

In DELPHI the BRICH is placed between the TPC and the OD (thus contributing a large fraction of the material which makes electron identification difficult). As can be seen in figure 3.6, with increasing momentum all particles tend to a saturation value of the Cherenkov angle

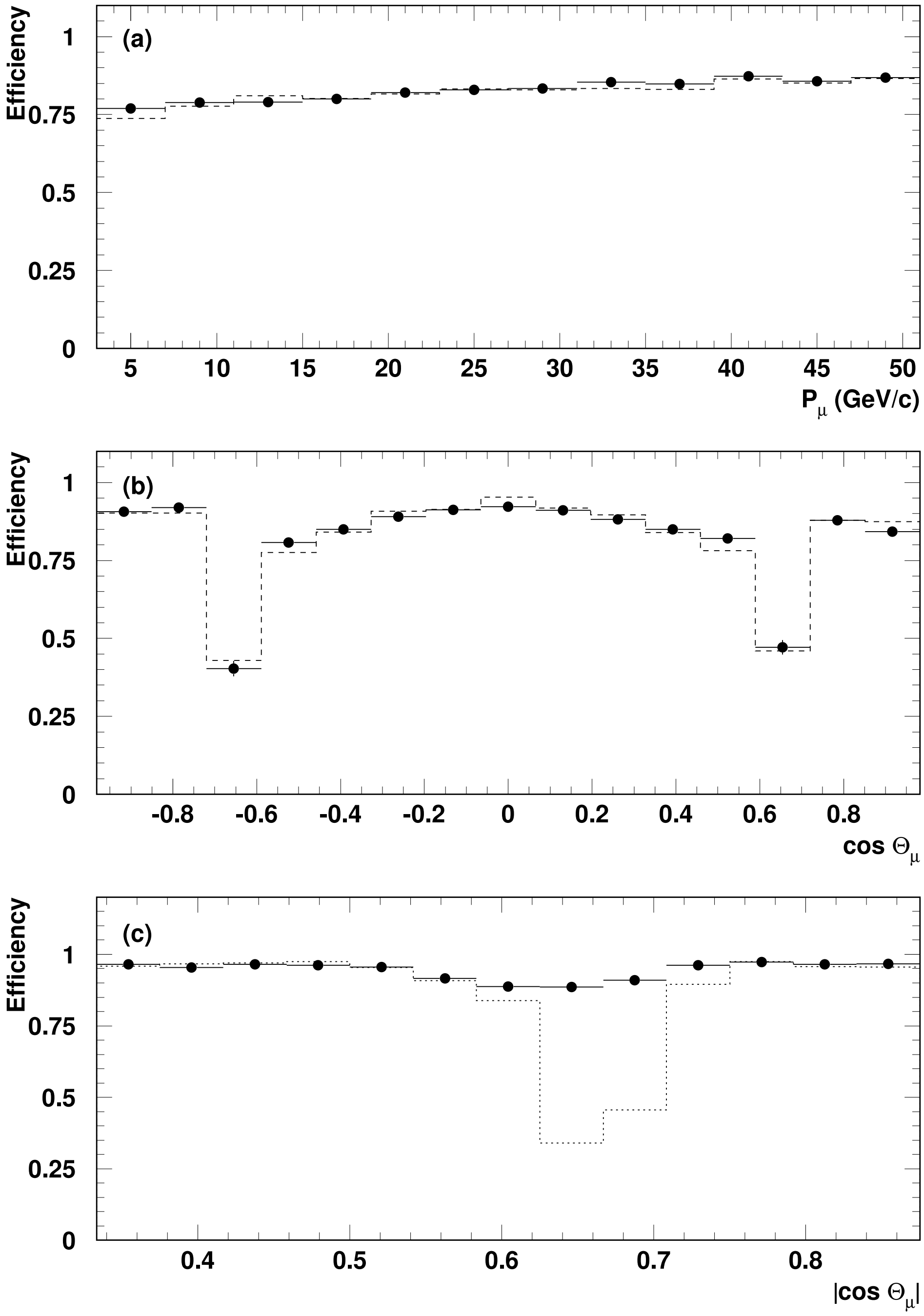


Figure 3.4: *Muon identification efficiency for the* standard *tag [50]: Plots a) and b) show the efficiency versus the muon momentum and polar angle, respectively. The points represent real data and the dashed line Monte Carlo simulations. Plot (c) shows the performance of the* very loose *tag in the surround muon chambers which is not used in this analysis.*

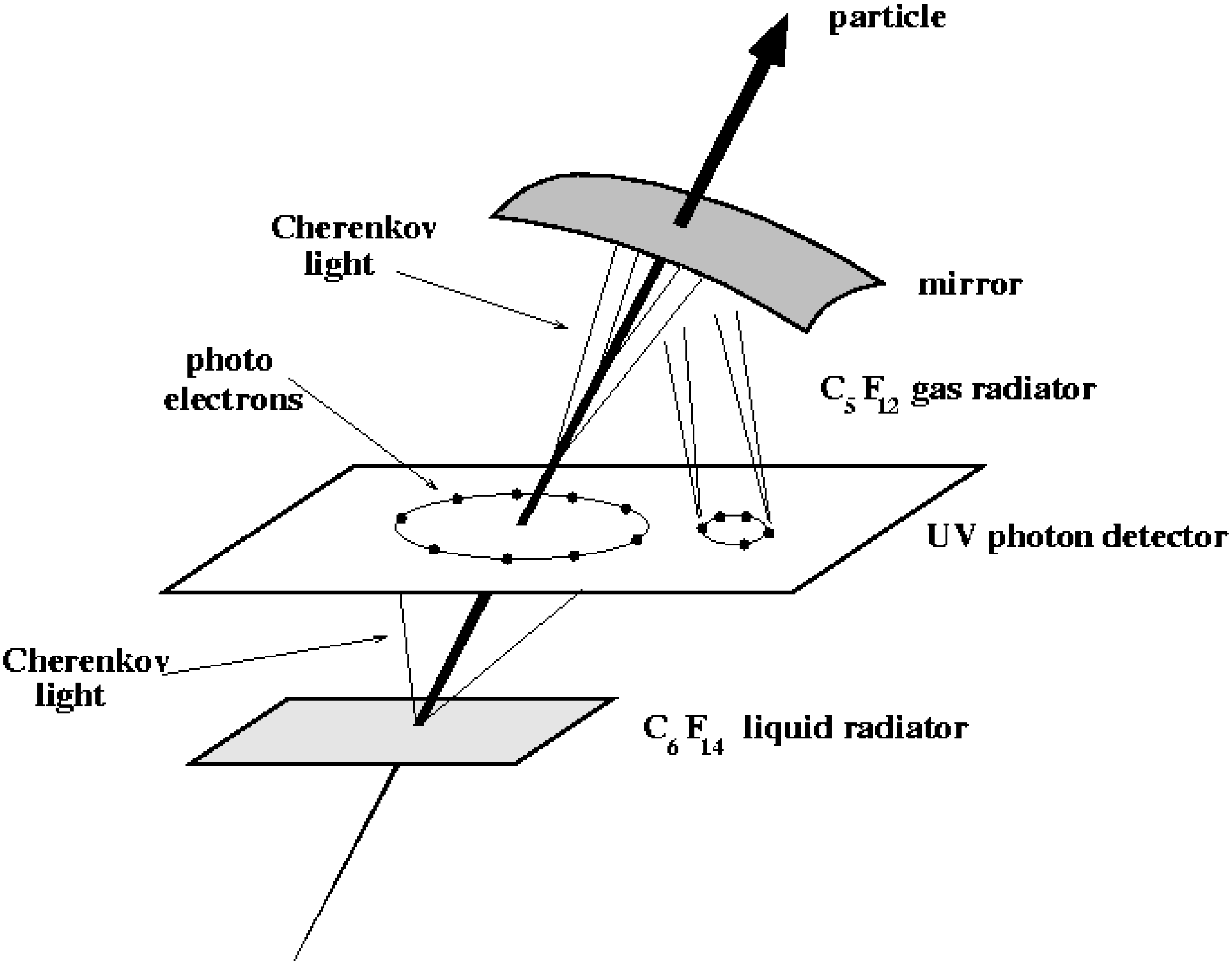
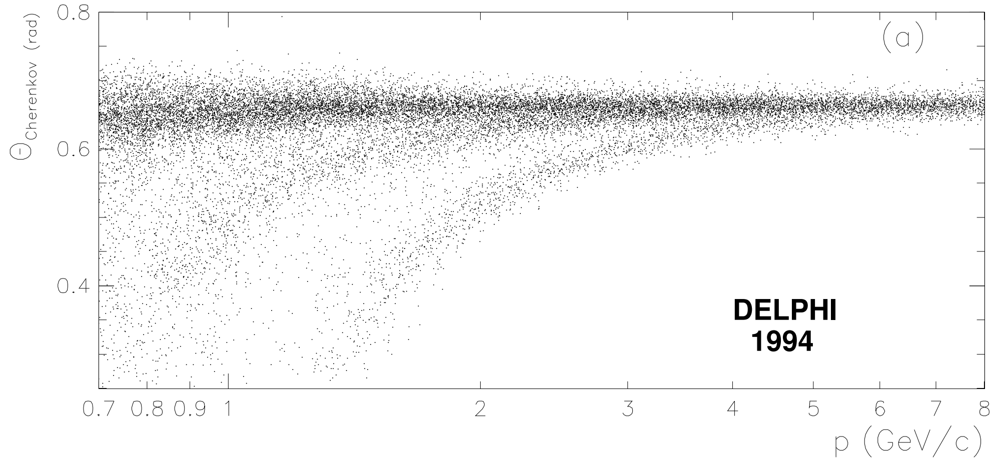


Figure 3.5: *Cherenkov radiation in the RICH detector: In the DELPHI RICH the track first crosses the liquid radiator before going through the gas radiator. Photons from the liquid radiator are recovered as a ring in the photon detector. Photons from the gas radiator are reflected by spherical mirrors onto the same photon detector.*



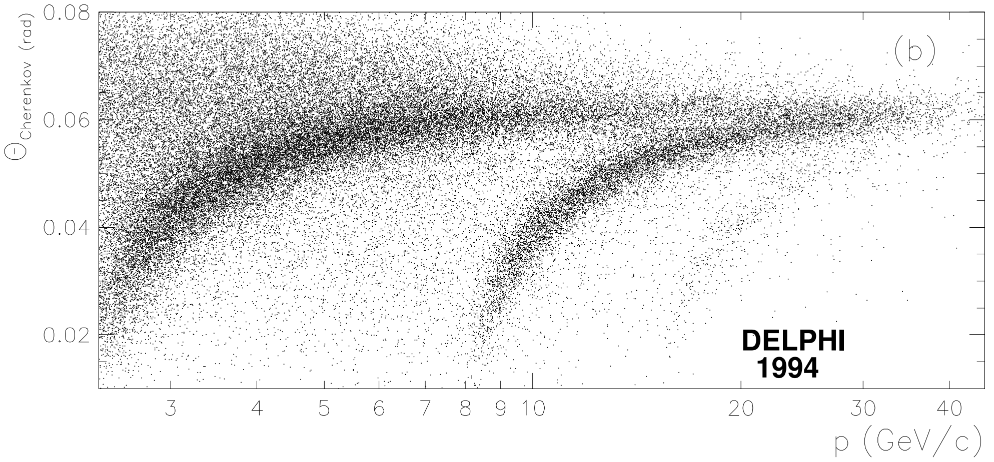


Figure 3.6: *DELPHI particle identification - RICH [50]: The figures show the average Cherenkov angle per track versus the momentum in multihadronic events in the Barrel RICH for the liquid (top) and gas (bottom) radiators. The three bands in both plots correspond to pions (uppermost band), kaons (middle band) and protons (lowest band) for the 1994 data set.*

which is independent of the particle mass. DELPHI, therefore, has two different radiators to better explore the whole momentum range.

The particle first traverses the liquid RICH, a 1 cm C_6F_{14} liquid radiator enclosed in transparent quartz. It then goes through the gas RICH, a 40 cm C_5F_{12} gas radiator. In average, a track gives 14 photons in the liquid radiator and 8 in the gas radiator. As figure 3.5 shows, the photons from the liquid RICH are seen as a ring in the photon detector. The photons from the gas RICH are reflected by spherical mirrors back onto the same photon detector as for the liquid RICH. In the photon detector the photons are converted to photoelectrons by a photoconvertive substance called TMAE which is mixed into the drift gas. The photon detector is a small TPC. By using all the available information the conversion point can be reconstructed in three dimensions. The photon can thus be associated with a track and its Cherenkov angle can be calculated. Table 3.3 gives some of the characteristics of the Barrel RICH.

Barrel RICH characteristics	Barrel liquid	Barrel gas
Number of photoelectrons per track	14	8
Cherenkov angle (mrad)	666	62.3
Angular precision per photoelectron (mrad)	13.2	4.3
Angular precision per track (mrad)	5.2	1.5

Table 3.3: *Barrel RICH characteristics for* $Z \longrightarrow \mu^+\mu^-$ *events.*

The RICH detector can be used in the *ring* mode which means that a signal is observed. The liquid RICH is useful for particle identification in the momentum range 0.7 GeV/c - 8 GeV/c while the gas RICH works from 2.5 GeV/c - 25 GeV/c. It is also possible to exploit the fact that there is no signal in the RICH. If a hadron with momentum below the kaon threshold for Cherenkov light goes through the RICH without leaving any signal (and the RICH was operational), it would mean, with a high probability, that it was a pion. This is called the *veto* mode.

There are several possible approaches to hadron identification depending on what one needs (high efficiency, high purity). The one [53] used in this analysis is based on an approach which in the ring mode uses a clustering algorithm on single photons which have similar Cherenkov angles. The photons are given weights depending on the expected background and clusters are formed. The cluster with the largest number of weighted photons is kept, with its mean weighted Cherenkov angle, its error and the number of photoelectrons. The expected error σ_{ring} is a function of the momentum p of the track, the expected error per photon σ_{photon} and the observed number of photons N [53]:

$$\sigma_{\text{ring}}^{\text{gas}} = \sigma_{\text{photon}}\ (1 + \frac{0.8}{2+p})\ \sqrt{1/N}$$

$$\sigma_{\text{ring}}^{\text{liquid}} = 0.85\ \sigma_{\text{photon}}\ (1 + \frac{0.8}{2+p})\ \sqrt{1/N\ +\ 1/12}$$

The expected error per photon σ_{photon} is given by [54]

$$\sigma^2_{\text{photon}} = \sigma^2_{\text{chromatic}} + \sigma^2_{\text{bend}} + \sigma^2_{\text{position}}$$

where $\sigma_{\text{chromatic}}$ is due to the variation of the refractive index with the photon energy, σ_{bend} is due to the bending of the track in the magnetic field and σ_{position} is due to the finite spatial resolution of the photon detector. For the liquid radiator, the term σ_{bend} is replaced by a contribution which is due to the finite depth of the liquid radiator.

To extract a probability for a certain particle hypothesis one uses a Gaussian part for the Cherenkov angle and a Poissonian part for the observed number of photons:

$$P_{\text{Cherenkov}} = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-(\theta_c - \theta_c^{\text{expected}})^2/2\sigma^2} \tag{3.3}$$

$$P_{\text{Photons}} = N^N_{\text{expected}}\, e^{-N_{\text{expected}}} \,/\, N! \tag{3.4}$$

where θ_c is the measured Cherenkov angle, σ is the expected error of the ring, $\theta_c^{\text{expected}}$ is the expected Cherenkov angle for that particle hypothesis, N is the observed number of photons and N_{expected} is the expected number of photons for that particle hypothesis. The total probability is then obtained by multiplying the two partial probabilities.

The actual situation is more complicated because the status of the RICH has to be taken into account. This is especially important in the veto mode. In addition, track quality selection criteria have an important influence on the final result. Requiring that the track has been seen in the OD reduces the background, but also affects efficiencies. Just like for the leptons, three tags (*loose*, *standard* and *tight*), with different efficiencies and purities, are provided [50]. Figure 3.7 shows the performance of the *tight* tag.

The TPC

Besides being the main tracking detector in DELPHI, the TPC's 192 sense wires measure the specific energy loss dE/dX as a 80% truncated mean of the amplitudes. The minimum requirement is 30 wires. The dE/dX measurement is available for 75% of the charged tracks in hadronic jets and provides a 1σ separation for pions and kaons with momenta above 1 GeV/c (see figure 3.8). Using the measured dE/dX, the expected error and the number of wires, probabilities for the different hypothesis can be constructed. Just like in the previous cases, three different tags (*loose*, *standard* and *tight*) are provided [50].

Combined RICH and TPC tag

The RICH and the dE/dX probabilities for the different particle hypothesis can be combined to give a combined RICH and dE/dX tag (*loose*, *standard*, *tight*) [50]. The *standard* tag identifies a charged kaon with 70% efficiency averaged over the whole momentum range above 0.7 GeV/c with a contamination of 30%.

While a combined tag seems *a priori* to be the best approach, caution must be shown. It is the RICH which gives the main contribution to the combined probability. Hits in the TPC which are too close in time are not correctly separated and are, therefore, not used in the dE/dX calculation. Tracks must be separated by at least 2 cm in the drift direction z. This is not always the case for tracks in a hadronic jet. For kaons coming from a ϕ this can be a

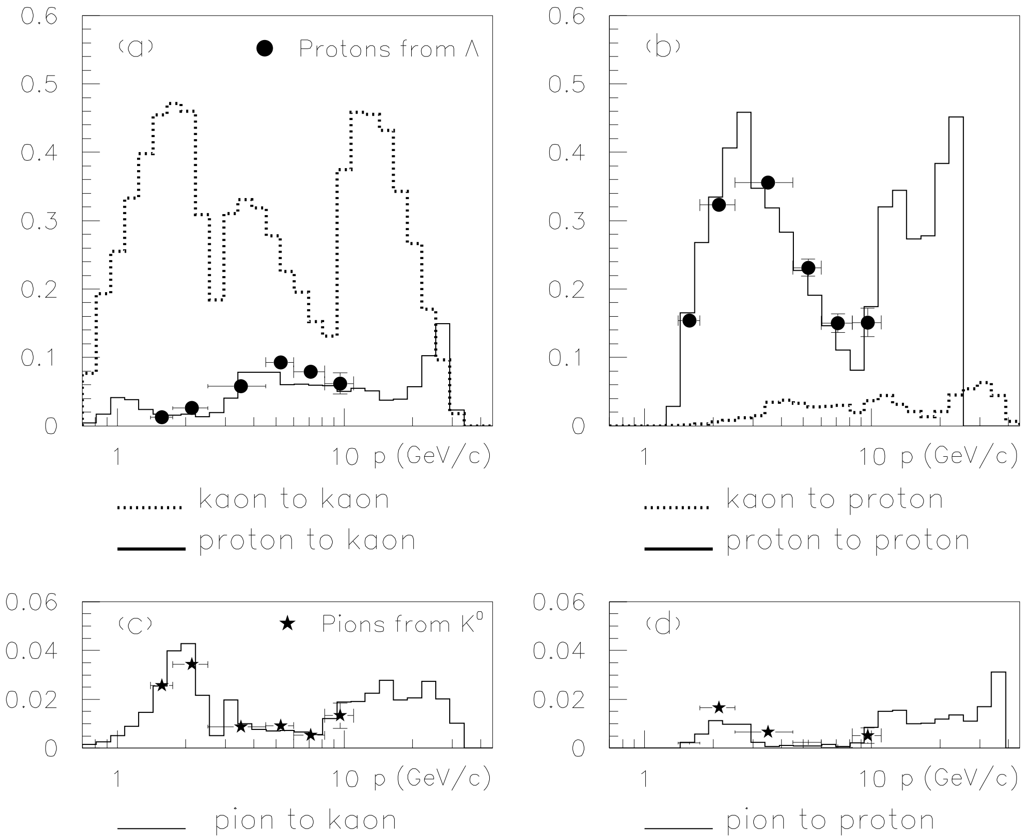


Figure 3.7: *Performance of the* tight *RICH tag [50]: Plot a) shows the fractions of kaons (dashed line) and protons (solid line) tagged as kaons in simulated hadronic decays versus the momentum. The circles are values obtained from protons from Λ^0 decays in real data. Plot b) shows the same as plot a), but for particles tagged as protons. Plot c) shows the fraction of pions (solid line) tagged as kaons in simulated hadronic decays. The circles are values obtained for pions from K^0 decays in real data. Plot d) shows the same as plot c), but for particles tagged as protons.*

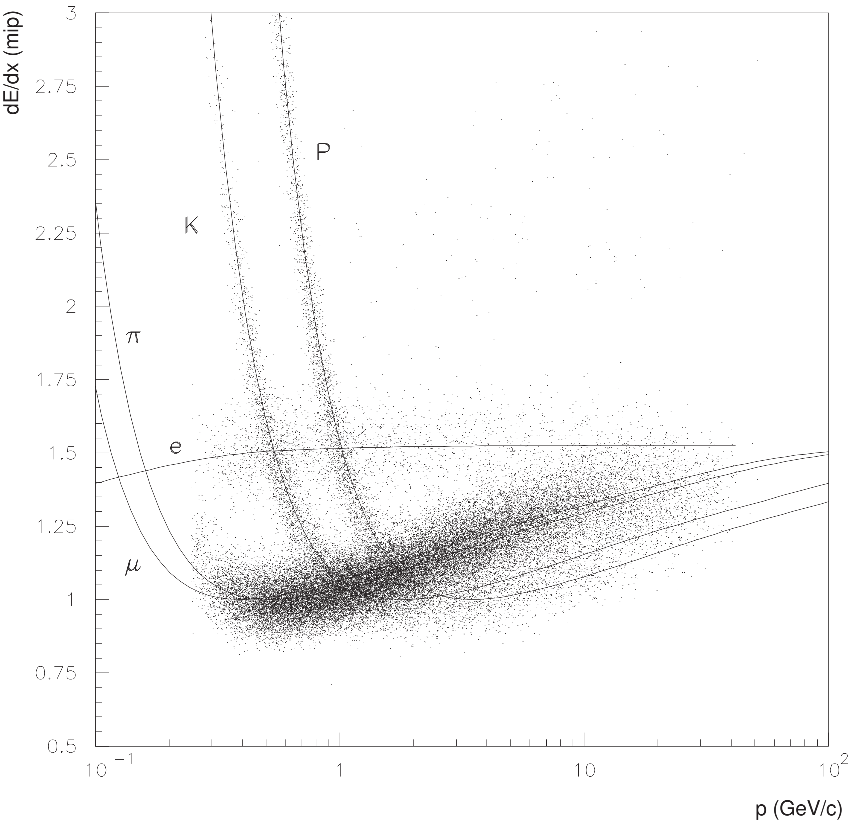


Figure 3.8: *DELPHI particle identification - dE/dX: The figure shows the specific energy loss dE/dX in the TPC versus the momentum (data) [50].*

serious problem since they have a very small opening angle. These kaons will quite often have no dE/dX information or, even worse, the dE/dX measurements from the two tracks can be mixed up. They may, however, be correctly identified in the RICH. It is not obvious that combining the RICH information with the dE/dX informations is the best approach in this case. It will not add anything to the RICH tag and may actually have a negative (momentum dependent) effect. For this reason, only the RICH information, instead of the combined tag, is sometimes used in certain modes.

Chapter 4

Experimental Motivation

4.1 Experimental Constraints

The time integrated mixing probability χ_q defined in equation (2.121), and plotted in figure 4.1, quickly saturates at 0.5 for high oscillation frequencies, This means that for the expected rapid oscillations in the $B_s^0 - \overline{B_s^0}$ system *a time independent* measurement will not be possible.

In the actual experimental situation the theoretical time distribution of the signal $P(t)$ is smeared out by the proper time resolution. Technically, this means that the signal is convoluted by a resolution function $g(t_m - t, \sigma)$ where t_m is the measured proper time and t the true proper time (which is integrated out). To see what kind of effect the resolution has on the observed proper time distribution, the Fourier transformation (FT) is used.

The Fourier transform of the convolution of two functions is just the product of the Fourier transform of the two functions. Assuming the resolution function to be a Gaussian with constant standard deviation (it will later be shown that the constant term in the proper time resolution is related to the decay length resolution and it is, therefore, written σ_l) it follows that

$$FT\left[(g(t, \sigma_l))\right](\Delta m_q) \;=\; e^{-\frac{\sigma_l^2}{2}\Delta m_q^2} \tag{4.1}$$

This means that the proper time resolution gives a *damping factor* which reduces the oscillations. The damping factor becomes more important for high oscillation frequencies, i.e. it is much more important for the rapid $B_s^0 - \overline{B_s^0}$ oscillations than for the relatively slow $B_d^0 - \overline{B_d^0}$ oscillations. The damping factor also depends on the decay length resolution. The momentum resolution part of the proper time resolution introduces a time dependent part, $\sigma_t^2 = \sigma_l^2 + (\sigma_p t)^2$, which complicates the calculations, but the general feature stays the same [55]. This means that only short proper times really contribute to the measurement. As the proper time increases, the damping factor grows and effectively kills any observation of the oscillations. This induces very stringent constraints on any analysis which tries to measure the $B_s^0 - \overline{B_s^0}$ oscillation frequency. It will need excellent proper time resolution.

4.2 B_s^0 Decay Modes and Possible Analyses

The proper time resolution is one important point of this analysis. Two others are the number of signal events and the fraction of background in the sample. There are several

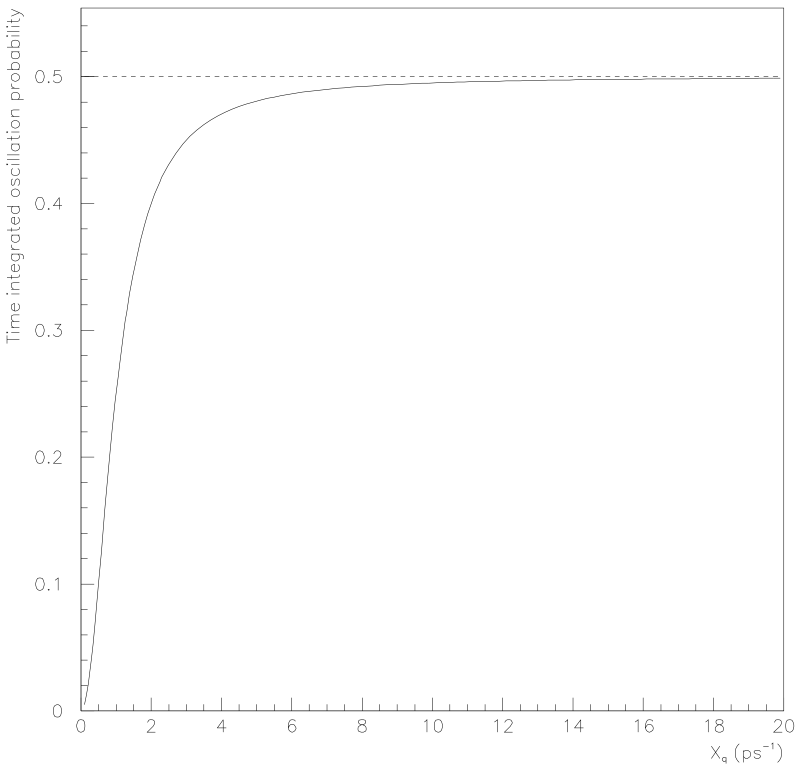


Figure 4.1: *Time integrated mixing probability versus* $x_q = \Delta m_{B_q^0}\,\tau_{B_q^0}$.

possible experimental approaches.

By doing an inclusive analysis one gets high statistics, but suffers from large time resolutions and high backgrounds. The exact composition of the sample is difficult to obtain and Monte Carlo simulations have to be used extensively. Such analyses are using b-enriched events with a high transverse momentum lepton in one hemisphere [34], [35], [36], [40], and events with one such lepton in each hemisphere [33].

The opposite approach is to completely reconstruct the B_s^0 meson. This gives an excellent proper time resolution, but yields only a handful of events [41].

A semi-exclusive approach, therefore, seemed to be the best solution. There are several possibilities. Most B_s^0 mesons decay into a D_s^- meson. One can either use events with a reconstructed $D_s^{\mp}$ meson [37] or look at events with a reconstructed $D_s^{\mp}$ meson in combination with a charged hadron $D_s^{\mp}h^{\pm}$ [42]. The advantage of this approach is that it includes 75% of the total B_s^0 decays [1]. Potential problems are high backgrounds from several sources and a proper time resolution which for parts of the sample is too high for this kind of analysis.

One alternative approach is to do an inclusive ϕl analysis. Since many of the D_s^{mp} decay modes involve a ϕ meson which decays into K^+K^- and DELPHI has powerful kaon identification with the RICH detector, this analysis seemed to be promising. It suffers, however, from the usual inclusive problems, i.e. high proper time resolution and high backgrounds.

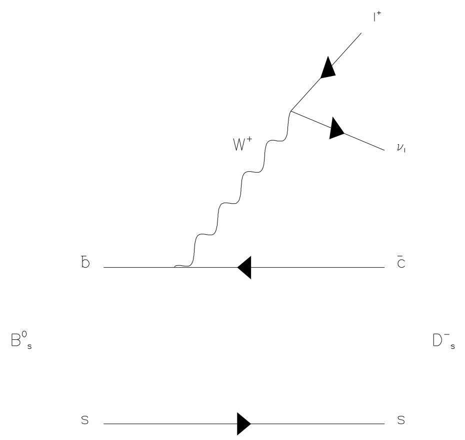


Figure 4.2: $B_s^0 \longrightarrow D_s^- l^+ \nu_l$ *Feynman diagram.*

This approach was therefore abandoned. The analysis has since been done in reference [56].

A third approach is to use the semileptonic decays of the B_s^0 meson in combination with a fully reconstructed $D_s^{\mp}$ meson [38], [39]. The corresponding Feynman diagram is shown in figure 4.2. This has the advantage of a low proper time resolution even when taking the missing neutrino into account. Because of the presence of a high transverse momentum lepton and a fully reconstructed $D_s^{\mp}$ meson, the background is low. In particular, it does not suffer from a high B_d^0 background which is the case for all the other inclusive and semi-exclusive analyses. This is especially important for the measurement of the B_s^0 lifetime since the B_d^0 and B_s^0 lifetimes are quite similar [1].

Only 16.2% of B_s^0 mesons decay semileptonically to an electron or muon [1] (the tau is not used). In addition, the D_s has to be reconstructed, and most of its decay modes are very small, between 1% and 4%. To get satisfactory statistics it is, therefore, necessary to reconstruct the D_s in as many modes as possible. To complicate things further, some of these modes involve up to six tracks in the final state. With ordinary Fortran based analysis programs this is a very tedious, time consuming and error prone task. It became obvious that an alternative programming approach had to be found.

4.3 Kinematic Analysis Language

Doing physics analysis with a multipurpose detector like DELPHI operating at the Z^0 resonance can be a rather complex task. The event multiplicity is large, a lot of different quantities for each particle are measured and one sometimes needs sophisticated selection criteria to distinguish the signal from the background. In the case of B-physics, one has additional problems with a huge combinatorial background and excellent book keeping is necessary to avoid double counting. The situation is made even more complicated by the fact that one has intermediate resonances which also have to be reconstructed.

At the same time there are huge similarities from mode to mode in the sense that the overall structure is the same. One combines tracks to intermediate resonances which themselves are combined with other tracks to form new intermediate resonances etc. This continues until one has (partially) reconstructed the B meson. During each step one has to find decay vertices, access information and keep track of each of the constituent particles, apply different selection criteria and access any Monte Carlo generator information. Eventually, the informations has to be stored for further analysis, usually in Ntuples, and an easy to use interface to a graphical analysis package like PAW [57] is necessary.

Kinematic Analysis Language KAL is a modern tool for data analysis that satisfies all these criteria (and even more). Originally developed by H. Albrecht for use at the ARGUS experiment [58], it was made modular, and thus easy to port to other experiments, by B. Spaan in CLEO. KAL was, therefore, adapted to the DELPHI experiment. A special effort was done to incorporate it completely into the DELPHI software environment, and make all of the detector information available to the user. A detailed description of its features and its use in DELPHI can be found in reference [59].

In KAL, the user no longer cares about the underlying data structures, but writes her analysis in the form of a KAL script. An example is given in appendix A.

Chapter 5

B_s^0 Candidate Selection

5.1 Hadronic Event Selection and Track Quality Criteria

In DELPHI there are several different sets of selection criteria for selecting hadronic events [50]. The one used in this analysis is meant to give a high, reliable and well understood efficiency and was originally made for the measurement of the hadronic cross section

$$\mathrm{e^+e^-} \rightarrow \mathrm{Z^0} \rightarrow \mathrm{hadrons}$$

The selection efficiency is more than 95% with a background of 0.7%, mainly from $\tau^+\tau^-$ pairs and $\gamma\gamma$ collisions.

The hadronic Z^0 events are selected by requiring:

- More than four charged tracks with:
  - momentum: $p > 400$ MeV/c
  - polar angle: $20^0 < \theta < 160^0$
  - track length of a least 30 cm in the Time Projection Chamber (TPC)
- The total energy of all charged tracks must be larger than 12% of $E_{\mathrm{Center\,of\,mass}}$
- The thrust axis must be clear of the beam pipe: $|\cos\theta_{\mathrm{thrust}}| < 0.95$

Monte Carlo simulation studies show that over 96% of the signal events are selected with these criteria:

$$\begin{array}{lllll}
\mathrm{e^+e^-} & \longrightarrow & \mathrm{Z^0} & \longrightarrow & \mathrm{b\bar{b}} \\
 & & & & \quad \hookrightarrow \mathrm{B_s^0} \\
 & & & & \qquad\quad \hookrightarrow \mathrm{D_s^{(*)-}}\ \mathrm{l^+}\ \nu_\mathrm{l}
\end{array}$$

where l means an electron or a muon. Unless stated otherwise, the charge conjugated processes are always implied.

To improve the track quality some additional requirements were placed on the charged tracks:

- $I_{R\phi} < 4.0$ cm
- $I_z < 10.0$ cm
- $|\cos\theta_{\text{track}}| < 0.937$
- $\frac{\sigma_p}{p} < 100\%$

$I_{R\phi}$ and I_z are the impact parameters with respect to the origin of DELPHI in $R\phi$ and z, respectively. These requirements help to remove tracks that do not come from the primary interaction point, like cosmic rays. The polar angle of the track, θ_{track}, is restricted to be above 20^0. This value reflects the angular coverage of the tracking detectors in DELPHI during LEP1. The uncertainty on the momentum, σ_p, is used to remove badly reconstructed tracks with a very small curvature and hence unrealistically high momenta (much larger than the beam energy).

There are also special software programs [50] reconstructing V0s like:

$$\begin{aligned} K_s^0 &\longrightarrow \pi^+\pi^- \\ \Lambda &\longrightarrow p\pi^- \end{aligned}$$

and photon conversions:

$$\gamma \longrightarrow e^+e^-$$

In these cases it is always the incoming (i.e. reconstructed) particle which is used, not the outgoing tracks.

The selection criteria are not only motivated by the underlying physics processes, i.e. hadronic Z^0 decays, but they also reflect the detector geometry and track reconstruction software. The θ_{track} criterium explained above is one example of this. Another example is the requirement of at least 30 cm track length in the TPC. The TPC is the main tracking detector in DELPHI and all track reconstruction started traditionally from the TPC and was then extrapolated inwards and outwards. The TPC criteria, therefore, made sense to get well reconstructed tracks.

During the reprocessing of the LEP1 data the DELPHI track reconstruction software underwent major changes to take into account changes in the detector hardware and to incorporate new ideas for combining information from different tracking detectors. This meant that tracks were allowed to originate in other tracking detectors than the TPC, like the Vertex Detector (VD), Inner Detector (ID) and Outer Detector (OD). Tracks composed of only VD, ID and/or OD track elements, i.e. not seen at all in the TPC, are now also reconstructed and accepted. A TPC track length criteria does of course not make sense for these new tracks.

The influence of theses changes on the final result was tested by putting the 1994-1995 data sample through the analysis criteria as described later in this chapter for the decay mode $D_s^- \longrightarrow \phi\pi^-$, but with the track and hadronic event selection turned off. The difference was found to be negligible. This is according to expectations since reconstructing a B_s^0 with a well defined decay vertex in the selected modes requires at least four high-quality tracks. It should also be added that the average charged track multiplicity of weakly decaying B hadrons is 4.97 [60], so most $b\bar{b}$ events will contain more than four charged tracks.

5.2 Data and Simulation Samples

This analysis is based on data recorded with the DELPHI detector at energies on and off the Z^0 resonance peak between 1991 and 1995. In total, 4 million hadronic Z^0 decays were collected in this period [50]. Table 5.1 gives the total number of hadronic Z^0 decays recorded by DELPHI during LEP1.

Starting at the end of 1995 the LEP energy has been gradually increased with the aim to achieve center-of-mass energies of 200 GeV. This phase is called LEP2 and will last until 2000. Even though some time is spent on the Z^0 resonance peak every year also during LEP2, the total number of Z^0s is rather low and the calibration and detector alignment that results is not precise enough for this analysis. This was checked explicitly by running on the 1996-1997 data set in the mode $D_s^- \longrightarrow \phi\pi^-$, a mode where only a minimal particle identification is used. No signal was observed.

Year:	1989	1990	1991	1992	1993	1994	1995	Total
Off peak:	4K	30K	52K	-	243K	-	236K	717K
Total:	13K	125K	275K	751K	755K	1484K	750K	4153K

Table 5.1: *Number of hadronic Z^0 decays recorded by the DELPHI detector during LEP1 (1K = 1000 hadronic Z^0 decays).*

Simulated events were generated with the JETSET 7.3 generator [61] with a special DELPHI parameter tuning based on experimental (DELPHI) measurements [62]. The generated events were then put through a full simulation of the DELPHI detector [50] and the simulated raw data was then put through the same reconstruction and analysis programs as the data.

5.3 Event Topology

The JETSET routine LUTHRU [61] was used to calculate the thrust axis of the event. The thrust is defined by

$$T = \max_{|\vec{n}|=1} \frac{\sum_i |\vec{n} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$$

where the sum is over all particles in the event, $\vec{p}$ is the momentum of the particle and $\vec{n}$ is a unit vector. The thrust axis is given by the vector $\vec{n}$ which maximizes the thrust. The thrust can have the values $1/2 \leq T \leq 1$. A two jet event corresponds to $T \approx 1$ and an isotropic event to $T \approx 1/2$.

The plane transverse to the thrust axis was used to divide the event into two hemispheres. The JETSET clustering routine LUCLUS [61], using both tracks[1] and neutral electromagnetic and hadronic clusters, were used to define jets.

The transverse momentum of each lepton candidate was calculated with respect to the jet it belonged to, after having removed the track from the jet and recomputed the jet axis.

[1]By convention, only charged particles are called "tracks".

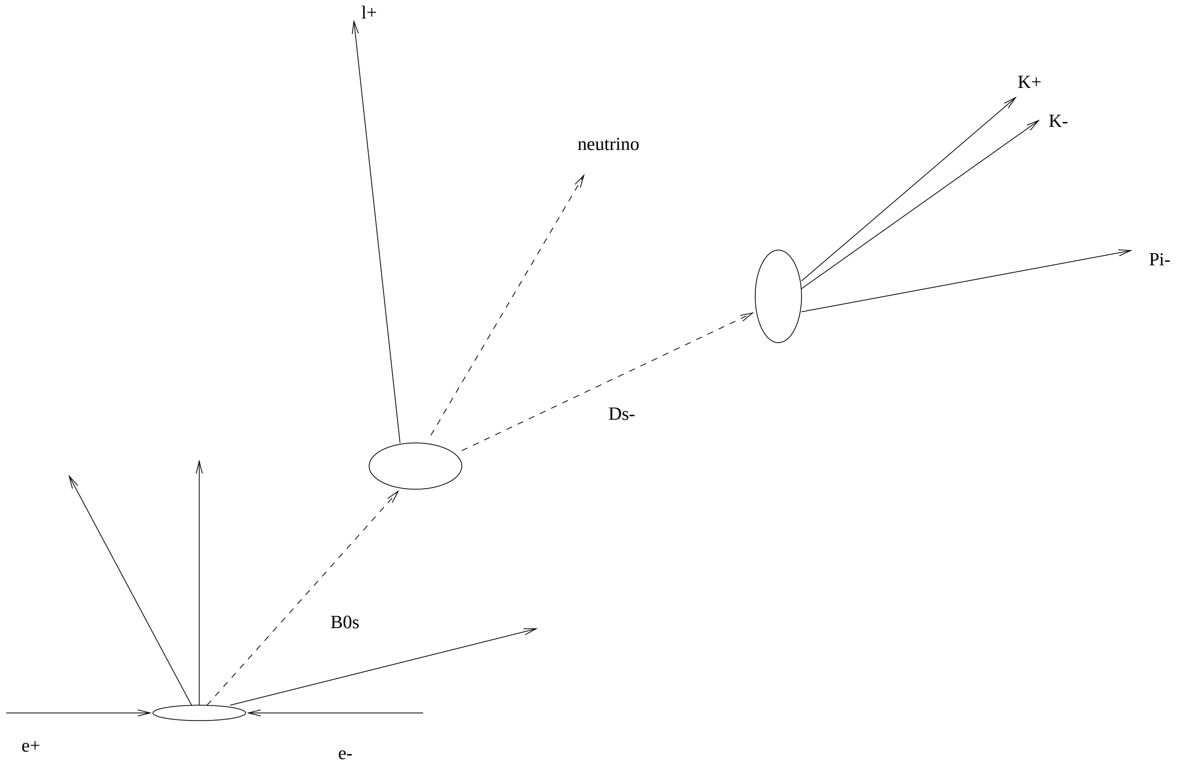


Figure 5.1: *Signal event topology (not in scale): The B_s^0 meson decay vertex will in average be around 3 mm from the primary interaction point. The D_s^- meson's average flight will be around 1 mm. The neutrino escapes unseen by the detector.*

This was done to maximize the difference between leptons from b-decays and other processes, especially from the cascade decay $b \longrightarrow c \longrightarrow l$ [63].

Figure 5.1 shows the general signal event topology. In all the modes the charged D_s^- decay products were fitted to a common vertex in three dimensions. A D_s^- pseudotrack was reconstructed and the extrapolation of this track was fitted to a common three dimensional vertex with the lepton candidate to give the B_s^0 decay vertex. The B_s^0 decay length was then determined by the distance between the B_s^0 decay vertex and the reconstructed three dimensional primary interaction point. On average, the B_s^0 meson decay vertex will be 3 mm from the primary interaction point while the D_s^- meson decay vertex will be 1 mm from the B_s^0 meson decay vertex. The neutrino escapes the detector unseen.

5.4 D_s^- Selection

5.4.1 General Remarks

The D_s^- candidate was reconstructed by combining all tracks in the same hemisphere compatible with the decay mode. All the tracks were required to belong to the same jet. Both momentum and particle identification selection criteria were used. In some modes only a veto particle identification was necessary, usually that a kaon candidate should not be tagged as a pion. In other modes a positive particle identification was necessary, both to reduce the combinatorial background and to reduce the contribution from background processes that could imitate the signal. Mass constraints from intermediate resonances were used to reduce the combinatorial background. The tracks were then fitted to a common three dimensional vertex and the corresponding $\chi^2_{D_s^-}$ of the vertex fit was used as a selection criterium. This reduces the combinatorial background and makes sure that the decay vertex is well defined. For the same reason, some of the tracks were required to have hits in the vertex detector.

After the vertex fit the track momenta were refitted with the reconstructed vertex point as a constraint and the D_s^- mass was recalculated. This could not be done for the 1991-1993 data set because of lack of z-information from the vertex detector during that period. Using the refitted tracks improves the mass resolution slightly. The track momenta are initially measured with respect to the origin of the detector. Refitting the tracks from a known secondary or tertiary vertex changes the momentum components and thus the invariant D_s^- mass. The total track momenta, however, do not change significantly.

The actual values of the different selection criteria were chosen either to suppress specific background processes or from maximizing the ratio $Signal/\sqrt{Background}$. This was done separately for each mode and each data set.

Most of the data was taken in 1994-1995. In addition, the detector did not change during that period. All Monte Carlo simulation plots, numbers and comparisons with data are, therefore, unless stated otherwise, for the 1994C processing, and for the decay mode[2] $D_s^- \longrightarrow \phi\pi^-$.

5.4.2 D_s^- Decay Modes

The D_s^- was reconstructed in the following five hadronic modes:

$$\begin{aligned} D_s^- &\longrightarrow \phi\pi^- \\ D_s^- &\longrightarrow K^{*0}K^- \\ D_s^- &\longrightarrow \phi\pi^+\pi^-\pi^- \\ D_s^- &\longrightarrow \phi\pi^-\pi^0 \\ D_s^- &\longrightarrow K^{*0}K^{*-} \end{aligned}$$

where:

$$\begin{aligned} \phi &\longrightarrow K^+K^- \\ K^{*0} &\longrightarrow K^+\pi^- \\ K^{*-} &\longrightarrow K_s^0\pi^-, \;\; K_s^0 \longrightarrow \pi^+\pi^- \end{aligned}$$

For the modes: $D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$, $D_s^- \longrightarrow \phi\pi^-\pi^0$ and $D_s^- \longrightarrow K^{*0}K^{*-}$, only the 1994-1995 data set was used.

In addition, the D_s^- was reconstructed in the two semileptonic modes:

$$\begin{aligned} D_s^- &\longrightarrow \phi e^-\bar{\nu}_e \\ D_s^- &\longrightarrow \phi\mu^-\bar{\nu}_\mu \end{aligned}$$

[2]This is to be understood as $B_s^0 \longrightarrow D_s^{(*)-}l^+\nu_l$, $(D_s^{*-} \longrightarrow D_s^-\gamma)$, $D_s^- \longrightarrow \phi\pi^-$, $\phi \longrightarrow K^+K^-$.

Attempts were also made to use the modes:

$$\begin{aligned} D_s^- &\longrightarrow \pi^+\pi^-\pi^- \\ D_s^- &\longrightarrow f_0(980)\,\pi^-, \;\; f_0(980) \longrightarrow \pi^+\pi^- \end{aligned}$$

but no clear signal was seen due to the large combinatorial background. The average numbers of pions and kaons in $Z^0 \longrightarrow b\bar{b}$ events are 18.5 and 2.5, respectively [64].

The mode

$$D_s^- \longrightarrow K_s^0 K^-$$

was not used. Due to the large measurement errors on the K_s^0, and the subsequent high B_s^0 decay length resolution, its contribution to B_s^0 mixing is negligible.

The next sections give a detailed description of the selection criteria used for the different modes.

5.4.3 $D_s^- \longrightarrow \phi\pi^-$

The ϕ was reconstructed in the mode $\phi \longrightarrow K^+K^-$ by taking all possible pairs of oppositely charged tracks that had a K^+K^- invariant mass within 13 MeV/c^2 of the nominal [1] ϕ mass.

To reconstruct the D_s^- candidate, a third track, the pion candidate, was added. Because of the narrow ϕ resonance the background is rather low and only a weak particle identification was necessary. The kaon candidates should not be tagged as pions by the RICH detector, neither should the pion candidate be tagged as a kaon. All three tracks were required to have at least one vertex detector hit and they should all be compatible with a single vertex. In addition, the following kinematic selection criteria were applied:

- $p_{K^\pm} > 2.0$ GeV/c
- $p_{\pi^-} > 1.0$ GeV/c
- $p_\phi > 4.5$ GeV/c
- $p_{D_s^-} > 6.0$ GeV/c
- $\chi^2_{D_s^-} < 50.0$

The reason for the relative high value of the $\chi^2_{D_s^-}$ is that the small opening angle of an energetic ϕ causes some problems for the vertex fit. The final decay length resolution is, as will be shown later, not affected, however.

The decay of the D_s^- is of the type

$$P \longrightarrow V\,P$$

where P and V mean a *pseudoscalar* and a *vector* meson, respectively. The *helicity* angle Ψ, i.e. the angle between the K^+ and the π^- in the ϕ rest frame, then follows a $\cos^2\Psi$ distribution for the signal due to helicity conservation, while it is flat for the background. Requiring $|\cos\Psi| > 0.40$ removes 40% of the background while keeping 96% of the signal.

For the 1991-1993 data set the dE/dX information was used instead of the RICH. Because of this, the veto requirements were less efficient and the fraction of combinatorial background in the sample increased.

Figures 5.2 and 5.3 show the K^+K^- invariant mass distribution for signal events (i.e. events also satisfying the B_s^0 selection criteria detailed in the next section) before any mass requirement for the 1994-1995 and 1991-1993 data sets, respectively.

5.4.4 $D_s^- \longrightarrow K^{*0}K^-$

The K^{*0} was reconstructed in the mode $K^{*0} \longrightarrow K^+\pi^-$ by taking all possible pairs of oppositely charged tracks that had a $K^+\pi^-$ invariant mass within 75 MeV/c^2 of the nominal K^{*0} mass. A third track, with electric charge opposite to the K^+ candidate, was then added.

Due to the large width of the K^{*0} resonance, the combinatorial background was more important than in the previous mode, thus requiring a tighter particle identification. At least one of the two kaon candidates had to be tagged as a *loose* kaon and neither kaon should be tagged as a pion. The pion should not be tagged as a kaon. These rather strict requirements also helped to reduce contributions from the two kinematic *reflections*

$$\begin{aligned} \mathrm{B_d^0} \quad \longrightarrow \quad & \mathrm{D^{(*)-}l^+\nu_l} \\ & \hookrightarrow \mathrm{K^{*0}(K^+\pi^-)}\ \pi^- \\ & \hookrightarrow \mathrm{K^+\pi^-\pi^-}\ \text{(non resonant)} \end{aligned}$$

By misidentifying a pion for a kaon, these decay modes can give a signature that imitates the signal. Further details will be given in section 5.8.4.

All three tracks were required to have at least one vertex detector hit and they should be compatible with a single vertex. In addition, the following kinematic selection criteria were applied:

- $p_{K^+} > 2.5$ GeV/c
- $p_{K^-} > 4.0$ GeV/c
- $p_{K^{*0}} > 4.0$ GeV/c
- $p_{D_s^-} > 8.0$ GeV/c
- $\chi^2_{D_s^-} < 20$

Because of the larger K^{*0} opening angle, the χ^2 requirement on the D_s^- vertex fit could be loosened up compared to the previous mode.

This decay is also of the type $P \longrightarrow V\ P$, with the same angular dependence as in the mode $D_s^- \longrightarrow \phi\pi^-$. Due to the higher combinatorial background, the requirement on the helicity angle was tightened to $|\cos\Psi| > 0.45$.

For the 1991-1993 data set one of the two kaons were required not to be tagged as a pion by the dE/dX. In order to reduce the increased combinatorial background due to less efficient particle identification, the selection criteria on the invariant mass was tightened to 50

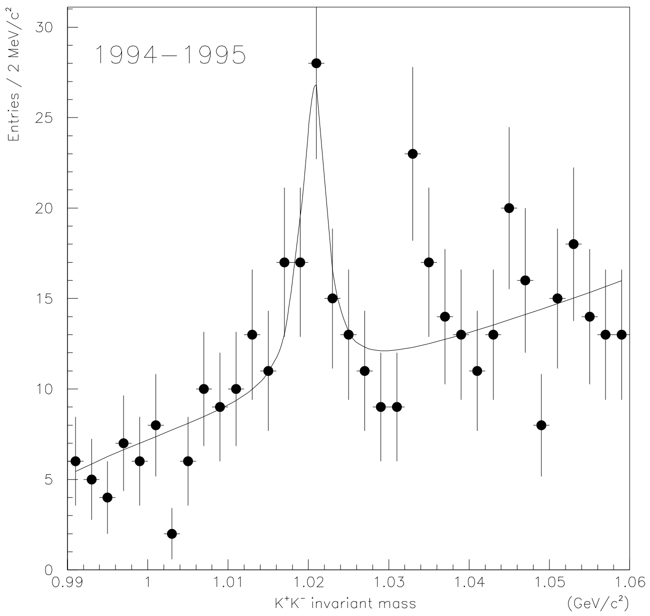


Figure 5.2: K^+K^- *invariant mass distribution for the* $D_s^- \longrightarrow \phi\pi^-$ *mode for the 1994-1995 data set.*

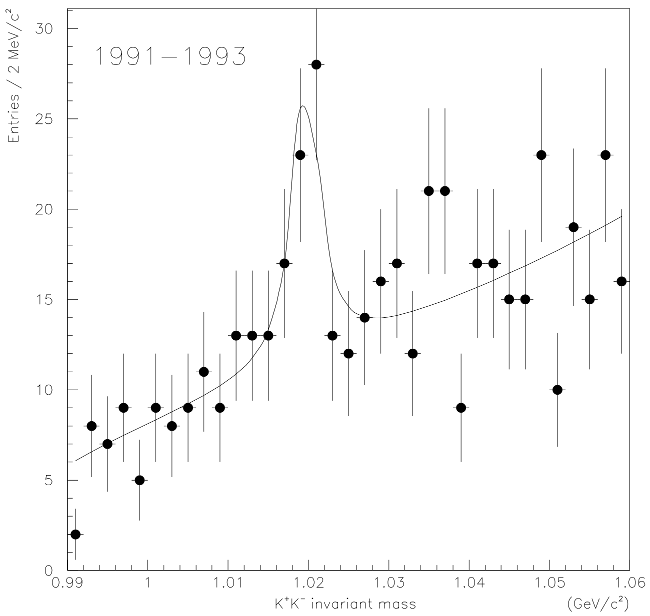


Figure 5.3: K^+K^- *invariant mass distribution for the* $D_s^- \longrightarrow \phi\pi^-$ *mode for the 1991-1993 data set.*

MeV/c^2, and the requirement on the helicity angle increased to $|\cos \Psi| > 0.50$. In addition, the following kinematic selection criteria were tightened:

- $p_{K^-} > 4.5$ GeV/c
- $p_{K^{*0}} > 4.5$ GeV/c
- $p_{D_s^-} > 10.0$ GeV/c

5.4.5 $D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$

The ϕ was reconstructed in the mode $\phi \longrightarrow K^+K^-$ by taking all possible pairs of oppositely charged tracks that had a K^+K^- invariant mass within 13 MeV/c^2 of the nominal ϕ mass. Neither kaon candidate should be tagged as a pion. Then three tracks, all compatible with the pion hypothesis, were added. Three of the five tracks were required to have at least one vertex detector hit and all five tracks should be compatible with a single vertex. No χ^2 selection was made on the corresponding D_s^- vertex, however. The following kinematic selection criteria were applied:

- $p_{K^\pm} > 2.0$ GeV/c
- $p_{\pi^\pm} > 0.7$ GeV/c
- At least two pions with $p_{\pi^\pm} > 1.2$ GeV/c
- $p_{D_s^-} > 9.0$ GeV/c

5.4.6 $D_s^- \longrightarrow \phi\pi^-\pi^0$

The ϕ was reconstructed in the mode $\phi \longrightarrow K^+K^-$ by considering all possible pairs of oppositely charged tracks with a K^+K^- invariant mass within 13 MeV/c^2 of the nominal ϕ mass. A third track, the pion candidate, was then added. Neither kaon candidate should be tagged by the RICH as a pion, neither should the pion candidate be tagged as a kaon. All three tracks were required to have at least one vertex detector hit and they should all be compatible with coming from a single vertex. A π^0 candidate was then added to the three tracks to make a D_s^-.

The π^0s were reconstructed by using a kinematic fit with the nominal π^0 mass as a constraint. All $\gamma\gamma$ combinations whose invariant mass was within 20 MeV/c^2 of the nominal π^0 mass, and whose fit probability was above 0.01, were kept. The π^0 was not used in the D_s^- vertex fit. In addition, the following kinematic selection criteria were applied:

- $p_{K^\pm} > 2.5$ GeV/c
- $p_{\pi^-} > 1.0$ GeV/c
- $p_{D_s^-} > 10.0$ GeV/c
- $\chi^2_{D_s^-} < 20$

Because of the presence of a π^0 in the final state, the ϕ is less energetic than in the mode $D_s^- \longrightarrow \phi\pi^-$, and the opening angle of the ϕ is larger. This is reflected in the stricter χ^2 requirement on the D_s^- vertex fit.

5.4.7 $D_s^- \longrightarrow K^{*0}K^{*-}$

The K^{*0} was reconstructed in the mode $K^{*0} \longrightarrow K^+\pi^-$ by considering all possible pairs of oppositely charged tracks with a $K^+\pi^-$ invariant mass within 75 MeV/c^2 of the nominal K^{*0} mass. The K^{*-} was reconstructed in the mode $K^{*-} \longrightarrow K_s^0\pi^-$ by combining all K_s^0 candidates[3] with a negatively charged track. The invariant $K_s^0\pi^-$ mass was required to be within 95 MeV/c^2 of the nominal K^{*-} mass.

Since both resonances have large widths, the same tracks can be used in both of them, i.e. the same track can both appear as a π^- from the K^{*0} in one D_s^- candidate, and as a π^- from the K^{*-} in another one. In this case, the $K^{*0}K^{*-}$ combination that minimized the squared mass differences

$$\sqrt{(M_{K^+\pi^-} - M_{K^{*0}})^2 + (M_{K_s^0\pi^-} - M_{K^{*-}})^2}$$

was chosen.

All three tracks were required to have hits in the vertex detector. The charged tracks were then fitted to a common vertex. In addition, the following kinematic selection criteria were applied:

- $p_{\pi^-} > 1.0$ GeV/c
- $p_{K^+} > 1.5$ GeV/c
- $p_{K^{*0}} > 3.5$ GeV/c
- $p_{K^{*-}} > 4.0$ GeV/c
- $p_{D_s^-} > 11.0$ GeV/c
- $\chi^2_{D_s^-} < 30.0$

5.4.8 $D_s^- \longrightarrow \phi l^- \bar{\nu}_l$

The ϕ was reconstructed in the mode $\phi \longrightarrow K^+K^-$ by combining all possible pairs of oppositely charged tracks with a K^+K^- invariant mass between 0.99 GeV/c^2 and 1.06 GeV/c^2. The two kaon candidates were required to have a momentum above 1.5 GeV/c while the ϕ momentum had to be above 3.5 GeV/c. Neither kaon should be tagged as a pion.

A *loosely* identified lepton with momentum above 1.75 GeV/c was then added. There is no muon identification below 3 GeV/c, but electrons can still be identified in this range. Electron candidates were restricted to be within the barrel, thus ensuring a good identification in the electromagnetic calorimeter HPC. If the two leptons were of the same kind it was verified that they were not compatible with a J/Ψ decay.

There are now two leptons, from the B_s^0 and the D_s^-, which means that two neutrinos are missing. To ensure a correct lepton assignment, the reconstructed D_s^- invariant mass had to be smaller than the nominal D_s^- mass. In addition, the invariant mass resulting from

[3] The selection for $K_s^0 \longrightarrow \pi^+\pi^-$ has an efficiency averaged over the whole momentum spectrum of about 36%, with a background of 3% [50].

combining the lepton candidate from the B_s^0 with the two kaons had to be larger than the nominal D_s^- mass.

The total and visible energy in the signal hemisphere were calculated according to the procedure in section 6.2.1, and the missing energy was defined as

$$E_{\text{missing}} = E_{\text{total}} - E_{\text{visible}} \tag{5.1}$$

The missing energy was required to be positive.

Two of the three tracks were required to have hits in the vertex detector and all three tracks were fitted to a common vertex with $\chi^2_{D_s^-} < 25.0$.

For the 1991-1993 data set the electron identification was tightened to *standard* and the momentum criterium was increased to 2.25 GeV/c. In addition the two kaons were required to have momenta above 1.75 GeV/c and the ϕ momentum had to be larger than 4.75 GeV/c.

5.4.9 D_s^- Selection Criteria for All Decay Modes

Tables 5.2 and 5.3 give the different selection criteria for all the D_s^- decay modes.

D_s^- decay mode	Mass selection criterium (MeV/c^2)	D_s^- vertex fit	Helicity angle
$\phi\pi^-$	$\Delta M_\phi < 13$	$\chi^2_{D_s^-} < 50.0$	$\lvert\cos\Psi\rvert > 0.4$
$K^{0*}K^-$	$\Delta M_{K^{*0}} < 75$ (50)	$\chi^2_{D_s^-} < 20.0$	$\lvert\cos\Psi\rvert > 0.45$ (0.50)
$\phi\pi^+\pi^-\pi^-$	$\Delta M_\phi < 13$		
$\phi\pi^-\pi^0$	$\Delta M_\phi < 13$	$\chi^2_{D_s^-} < 20.0$	
$K^{0*}K^{*-}$	$\Delta M_{K^{*0}} < 75$ $\Delta M_{K^{*-}} < 95$	$\chi^2_{D_s^-} < 30.0$	
$\phi l^-\bar{\nu}_l$		$\chi^2_{D_s^-} < 25.0$	

Table 5.2: D_s^- *selection criteria for the 1994-1995 (1991-1993) data set.*

$\phi\pi^-$	$K^{0*}K^-$	$\phi\pi^+\pi^-\pi^-$	$\phi\pi^-\pi^0$	$K^{0*}K^{*-}$	$\phi l^-\bar{\nu}_l$
$P_{D_s^-} > 6.0$	$P_{D_s^-} > 8.0$ (10.0)	$P_{D_s^-} > 9.0$	$P_{D_s^-} > 10.0$	$P_{D_s^-} > 11.0$	
$P_\phi > 4.5$	$P_{K^{0*}} > 4.0$ (4.5)			$P_{K^{0*}} > 3.5$ $P_{K^{*-}} > 4.0$	$P_\phi > 3.5$ (4.75)
$P_{K^\pm} > 2.0$	$P_{K^-} > 4.0$ (4.5)	$P_{K^\pm} > 1.75$	$P_{K^\pm} > 2.5$	$P_{K^+} > 1.5$	$P_{K^\pm} > 1.5$ (1.75)
$P_{\pi^-} > 1.0$	$P_{K^+} > 2.5$	$P_{\pi^\pm} > 0.7$	$P_{\pi^-} > 1.0$	$P_{\pi^-} > 1.0$	$P_{lep} > 2.0$ (2.25)

Table 5.3: D_s^- *kinematic selection criteria for the 1994-1995 (1991-1993) data set.*

5.5 B_s^0 Selection

The D_s^- candidate tracks were used to form a three-dimensional vertex. The D_s^- pseudotrack was extrapolated backwards and intersected with a *loosely* tagged high transverse momentum lepton of opposite charge. The lepton was not required to be in the same jet as the D_s^- tracks. Requiring a large transverse momentum reduces contributions from other processes, especially the *cascade* decay

$$B \longrightarrow D^{(*)}(lep)\ D_s^{(*)}X$$

Here $D^{(*)}(lep)$ means the charmed meson will eventually give a lepton (in combination with other particles) in the final state. The distribution of the lepton transverse momentum is shown in figure 5.4 for signal and *cascade* events.

The lepton candidate was required to have a momentum above 3.0 GeV/c, transverse momentum p_t above 1.2 GeV/c and at least one hit in the vertex detector. Electron candidates were restricted to be within the barrel, thus ensuring a good identification in the HPC.

The D_s^- and the lepton candidates were fitted to a common three-dimensional vertex and the lepton momentum components were refitted with the vertex point as a constraint (1994-1995 data set only). In addition, the following kinematic selection criteria were applied:

- $\chi^2_{D_s^- l^+} < 20$
- $p_{D_s^- l^+} > 14$ GeV/c
- 3.0 GeV/c^2 $< M_{D_s^- l^+} <$ 5.5 GeV/c^2

The lower limit on the $D_s^- l^+$ invariant mass was changed to 2.5 GeV/c^2 for the semileptonic decays of the D_s^- due to the presence of a second neutrino. Requiring a rather high $D_s^- l^+$ invariant mass removes parts of the *cascade* background, as can be seen in figure 5.5.

Depending on the decay mode, sometimes more than one B_s^0 candidate in an event satisfied all the selection criteria. In that case, only the candidate with the highest lepton transverse momentum was kept. If the same lepton was attached to different D_s^- candidates, the candidate with the highest D_s^- momentum was kept.

5.6 Method for Choosing the Selection Criteria

As was previously mentioned, the selection criteria were chosen by maximizing the ratio $Signal/\sqrt{Background}$. This was done separately for each criterium and, therefore, neglects any correlation between the different criteria. To check the impact of these correlations, it was verified that optimizing the ratio $Signal/\sqrt{Background}$ for the lepton momentum and transverse momentum simultaneously did not change the already obtained numerical values for the two selection criteria.

Another approach for including all the correlations is the use of a *Linear discriminant* [65]. In this approach, all the different variables are combined into one discriminant variable which then contains all the discriminant power. The variable is optimized by maximizing $Signal/\sqrt{Background}$. This approach was tested on the $D_s^- \longrightarrow \phi\pi^-$ mode. No improvement was seen compared to the conventional approach.

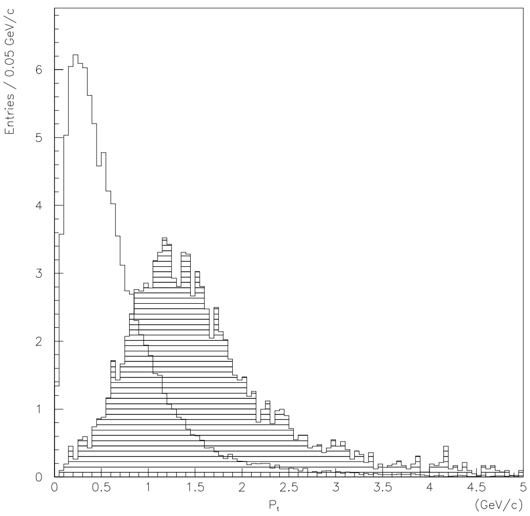


Figure 5.4: *Generated lepton transverse momentum with respect to the thrust axis for the signal (hatched) and* cascade *background.*

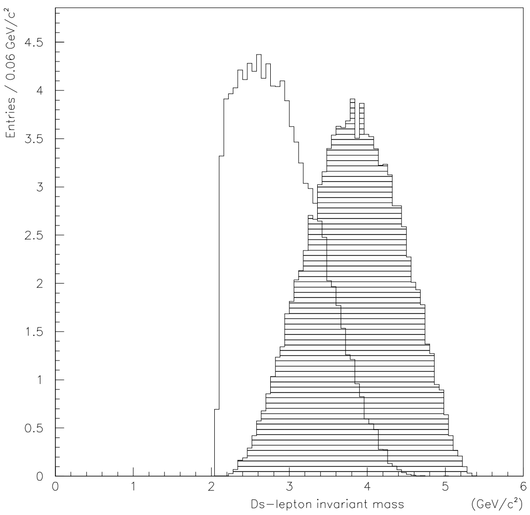


Figure 5.5: *Generated $D_s^- l^+$ invariant mass for the signal (hatched) and* cascade *background.*

It is, however, not obvious that maximizing the signal to the (square root of the) total number of background events is the best solution. The combinatorial background has, as will be shown later, a flying and a non-flying component. These two components will have a different influence on, for example, the lifetime measurement. On the other hand, it is not obvious how to take these different background components into account when choosing the selection criteria, especially when doing several different analyses like the B_s^0 lifetime, lifetime difference and mixing. Since it is useful to use the lifetime fit as a check that the sample composition is well understood, the same selection criteria must be used in all three analyses. It was, therefore, decided to maximize the signal to the square root of the total number of background events.

5.7 Signal Acceptance Function

The selection criteria applied to the signal may *a priori* introduce a bias with respect to the proper time of the signal. For instance, it may be more difficult to reconstruct a B_s^0 with a short lifetime, i.e. decaying close to the primary vertex, than it is to reconstruct one which flies several mm before decaying. Any bias will have to be corrected for with an *acceptance* function. Figure 5.6 (top) shows the reconstruction efficiency, here defined as the generated proper time distribution for all B_s^0 passing the selection criteria divided by the generated proper time distribution for all B_s^0, versus the generated proper time. The distribution is uniform.

For long proper times, however, the statistics are low. To obtain reliable results in this case, it is useful to choose the following representation of the acceptance. The generated signal proper time follows an exponential distribution, in this case with a generated lifetime of 1.6 ps. Therefore, the distribution of the variable $u = e^{-t_{gen}/\tau}$ should be uniform. The statistics should also be the same in each bin of u. Figure 5.6 (middle) shows the distribution of u for all the generated B_s^0 and it is uniform. The lower plot shows the corresponding distribution for those passing the selection criteria. Within the statistical fluctuations it is uniform. A line fit through the data gives a slope which is compatible with 0 within 1.2 sigmas.

Thus, the selection criteria do not introduce any bias with respect to the proper time and there is no need for any *acceptance* function.

5.8 Sample Composition

5.8.1 Signal

In order to determine the relative contributions of the signal and the background processes, the D_s^- mass spectrum was fitted with a Gaussian for the signal and a second Gaussian for contributions coming from the Cabibbo-suppressed hadronic decays of the D^-. A linear function was used for the combinatorial background, except for the mode $D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$ where an exponential was used.

The ϕ mass spectrum was fitted using a Breit-Wigner for the signal, and a combination of a linear function and a square root for the background.

For the signal, both the mean value and the width were left as free parameters. For the D^-, the mean was fixed to the nominal D^- mass and the width was fixed to be the same as

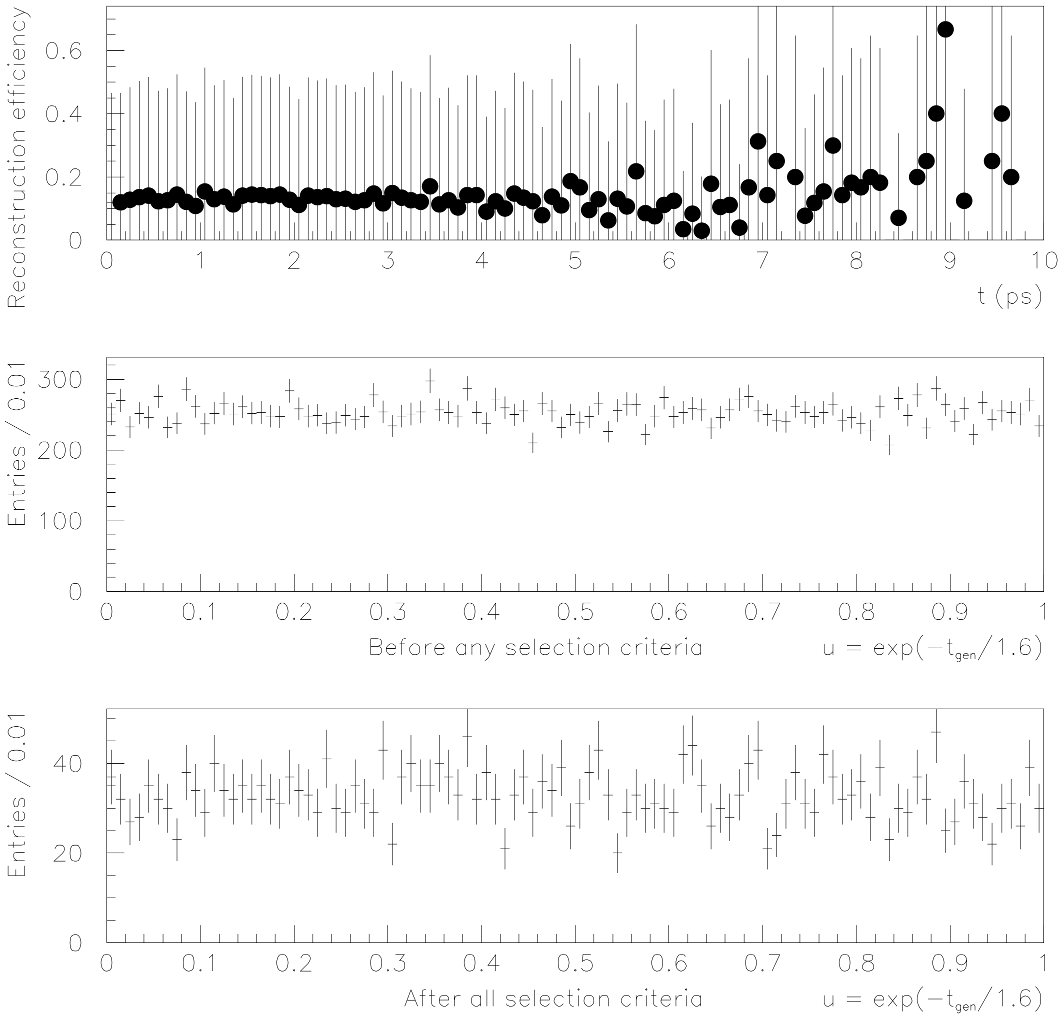


Figure 5.6: *Top: Signal acceptance defined as the generated proper time distribution for all B_s^0 passing the selection criteria divided by the generated proper time distribution for all B_s^0 as a function of the generated proper time. Middle:* $u = e^{-t_{gen}/\tau}$ *for all generated* B_s^0*. Bottom:* $u = e^{-t_{gen}/\tau}$ *for all* B_s^0 *passing the selection criteria.*

the D_s^- width. The resulting mass spectra are shown in figures 5.7-5.11.

For the five hadronic modes the fit yields:

- Number of $D_s^{\mp}$ candidates: 181 $\pm$ 20
- D_s mass: (1.968 $\pm$ 0.002) GeV/c^2
- D_s width: (15.9 $\pm$ 2.0) MeV

For the two semileptonic modes we get:

- Number of candidates: 78 $\pm$ 18
- ϕ mass: (1.021 $\pm$ 0.001) GeV/c^2
- ϕ width: (8.2 $\pm$ 2.4) MeV

Tables 5.4 - 5.6 show the contributions from the different modes and for the different data sets. Note that the estimated signal is the total as given by the fit to the D_s^- mass and, therefore, contains events from the signal itself, *cascade* decays and (in the case of $D_s^- \longrightarrow K^{*0}K^-$) kinematic *reflections*. Tables 5.7 and 5.8 show the reconstruction efficiencies and expected number of B_s^0 signal events estimated from Monte Carlo simulations.

The expected numbers agree with the data when taking all the uncertainties into account. The only exception is the mode $D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$, where one event was expected and 12$\pm$7 are observed. No events were observed in this mode in the 1991-1993 data. The signal in the 1994-1995 data set, therefore, seems to be due to a statistical fluctuation.

The fractions of electrons and muons in the data are shown in table 5.9. Taking the low statistics for 1991-1993 into account, the fraction is consistent with the expected difference in electron and muon identification efficiency.

The D_s mass resolutions in the modes $D_s^- \longrightarrow \phi\pi^-$ and $D_s^- \longrightarrow K^{*0}K^-$ are higher for the 1994-1995 data set than for the 1991-1993 data set while the opposite behaviour is expected from Monte Carlo simulations. They are, however, all compatible, within $1\sigma - 1.5\sigma$, with the individual values expected from Monte Carlo simulations.

In the analysis, only B_s^0 candidates with a D_s^- (ϕ) mass within $\pm 2\sigma (\pm 1.5\Gamma)$ of the measured mean mass were used.

5.8.2 Like-sign Candidates

The figures also include mass spectra for the like-sign candidates $D_s^{\pm}l^{\pm}$. If there is a problem with hadrons being misidentified as leptons, this would show up as a peak in the D_s mass region for the like-sign candidates since the electric charge of the hadron misidentified as a lepton is uncorrelated to the charge of the real D_s. There is no such structure in the complete data set, as can be seen in figure 5.7. In the mode $D_s^- \longrightarrow K^{*0}K^-$ for the 1994-1995 data set there seems to be some kind of peak structure in the D_s mass region, but since this is not seen in any of the other modes, and it is absent in the 1991-1993 data set for the same mode, it is interpreted as a purely statistical fluctuation. There is, therefore, no problem with hadrons being misidentified as leptons.

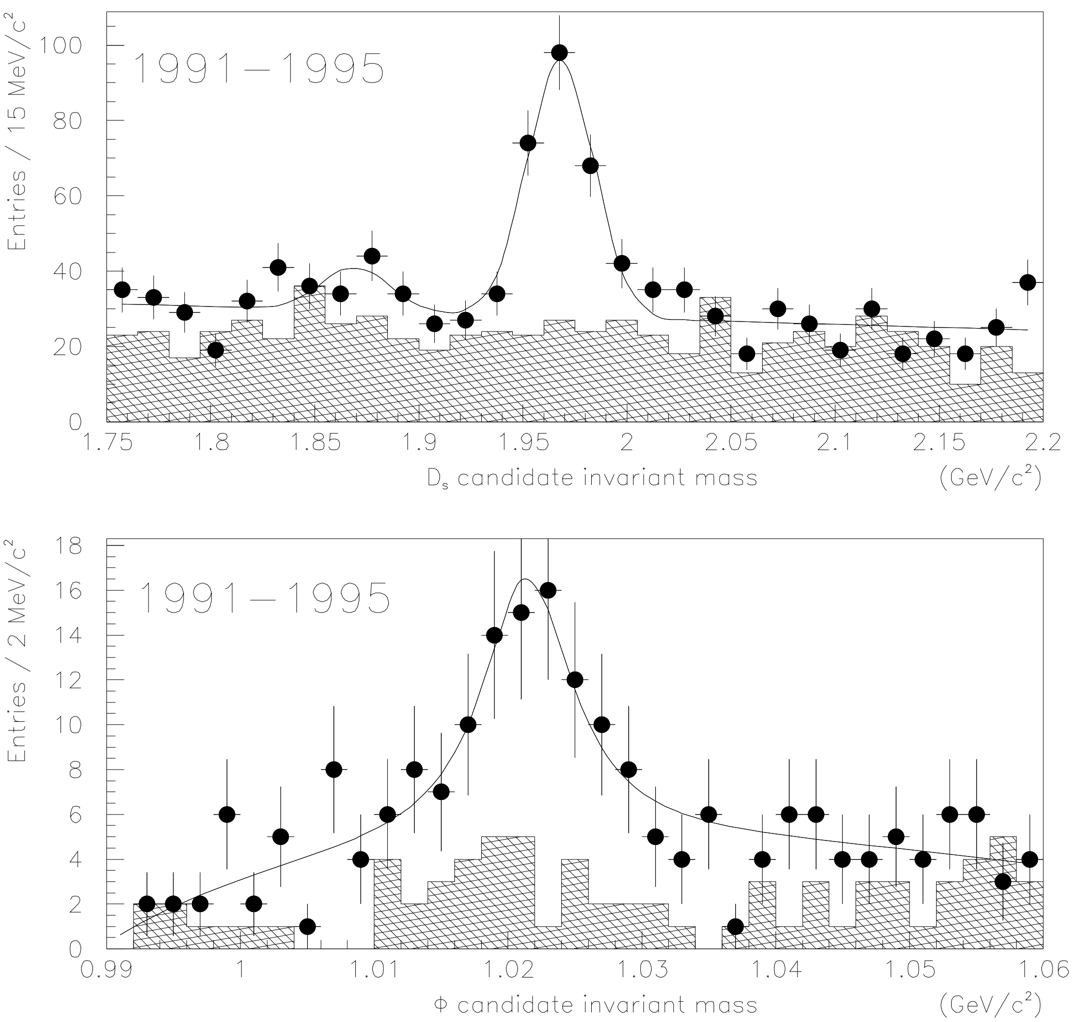


Figure 5.7: *Invariant mass distributions for* $D_s^{\mp}$ *(all hadronic decay modes) and* ϕ *candidates accompanied by an oppositely charged lepton in the same hemisphere for the entire data set. The dots show the data points while the solid lines correspond to the fits described in the text. The shaded histograms show the like-sign combinations.*

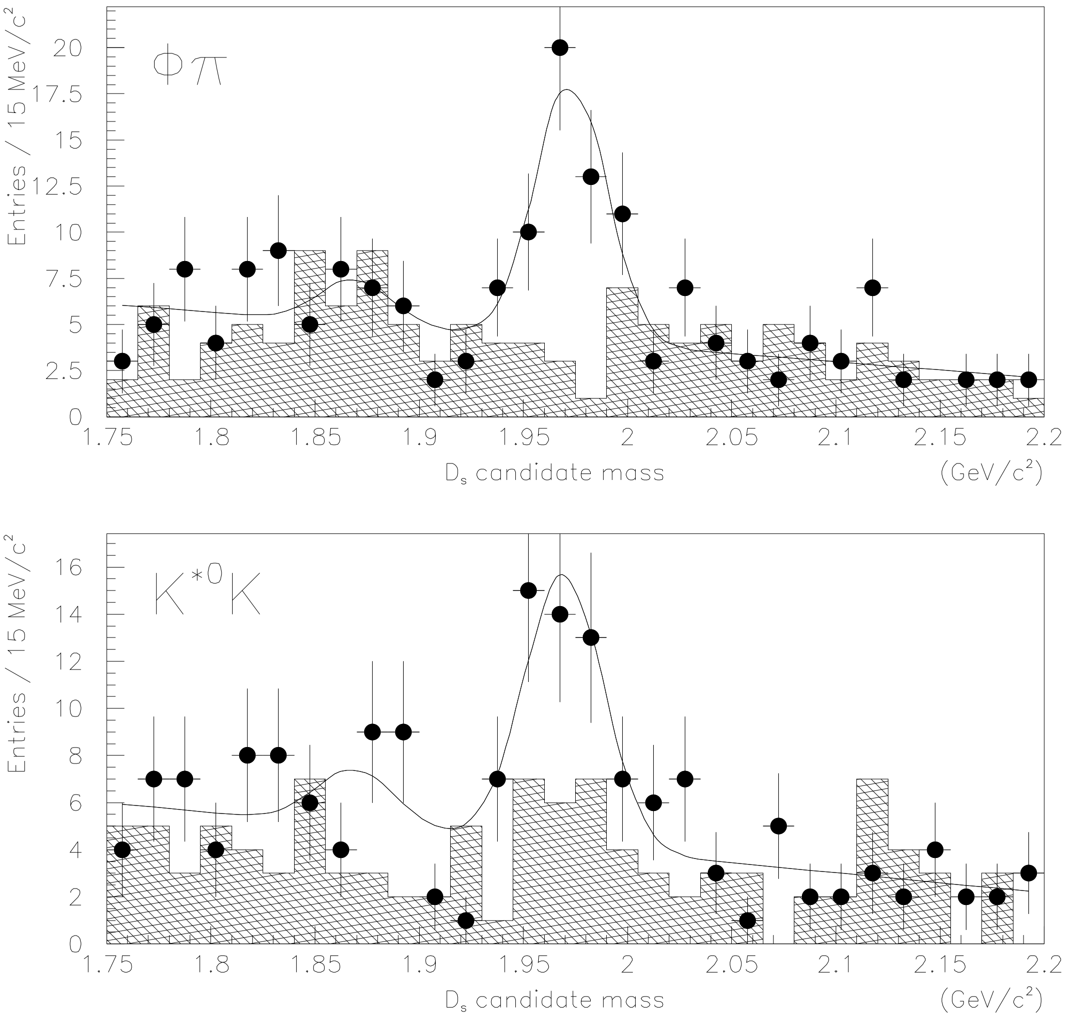


Figure 5.8: *Invariant mass distributions for $D_s^{\mp}$ candidates in the $\phi\pi^{\mp}$ and $K^{*0}K^{\mp}$ modes accompanied by an oppositely charged lepton in the same hemisphere for the 1994-1995 data set. The dots show the data points while the solid lines correspond to the fits described in the text. The shaded histograms show the like-sign combinations.*

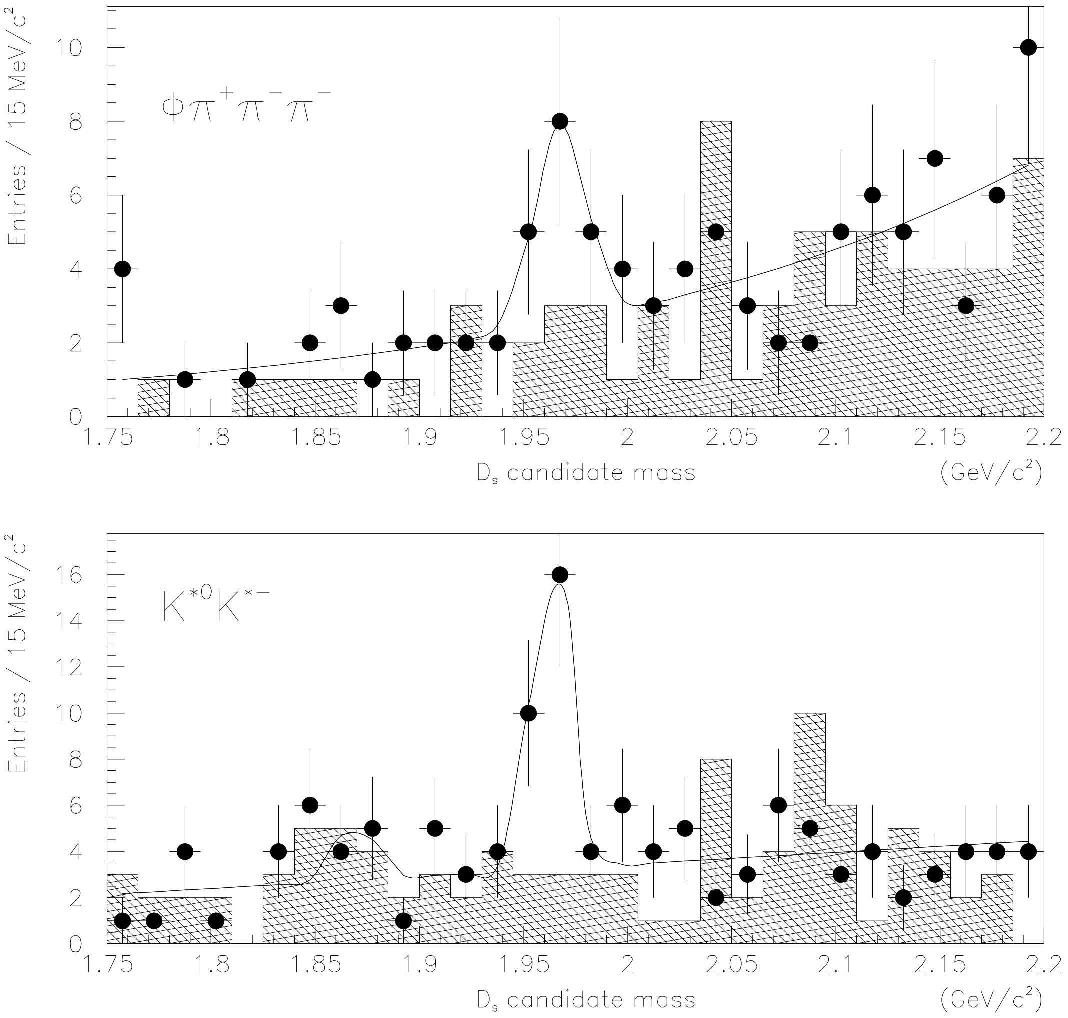


Figure 5.9: *Invariant mass distributions for $D_s^{\mp}$ candidates in the $\phi\pi^{\pm}\pi^{\mp}\pi^{\mp}$ and $K^{*0}K^{*\mp}$ modes accompanied by an oppositely charged lepton in the same hemisphere for the 1994-1995 data set. The dots show the data points while the solid lines correspond to the fits described in the text. The shaded histograms show the like-sign combinations.*

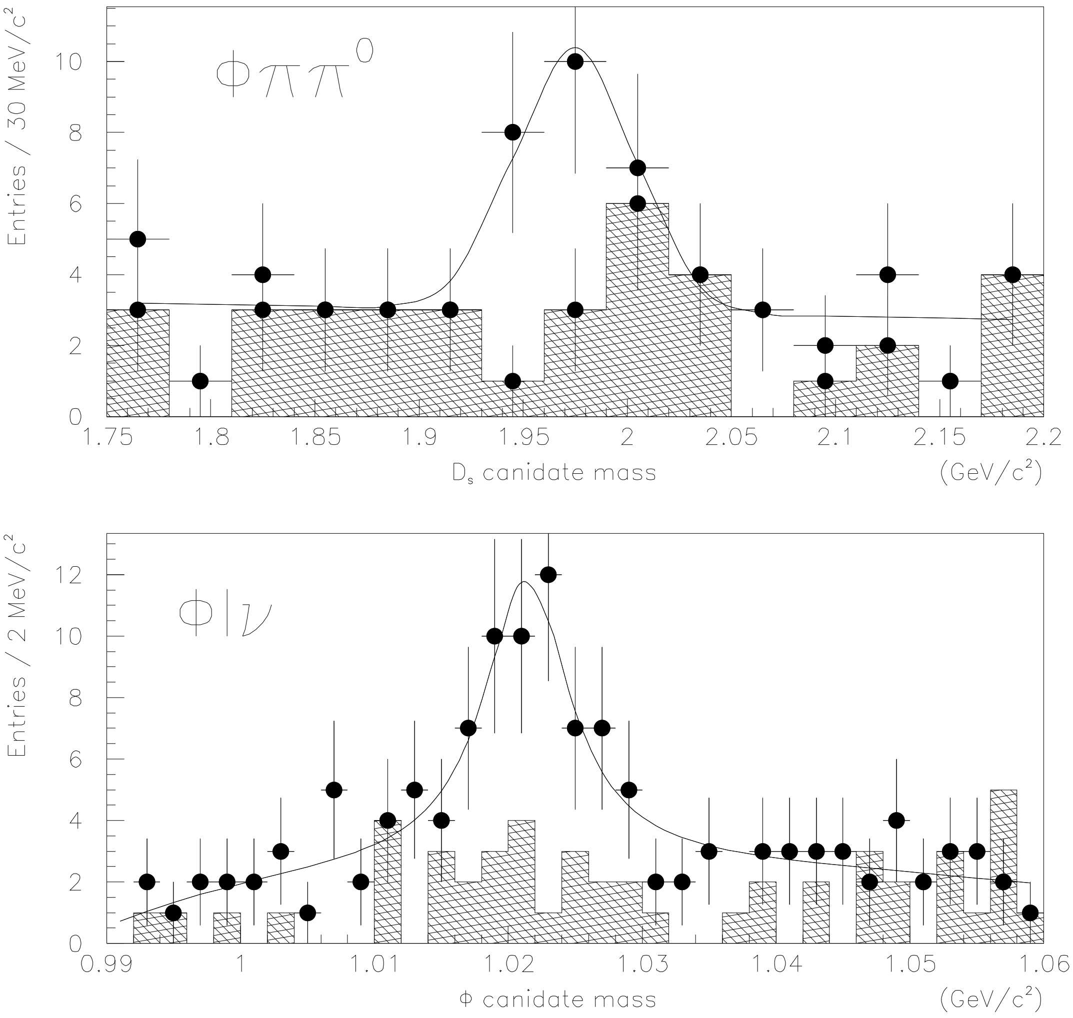


Figure 5.10: *Invariant mass distributions for $D_s^{\mp}$ and ϕ candidates in the $\phi\pi^{\mp}\pi^0$ and $\phi l^{\mp}\bar{\nu}_l$ modes accompanied by an oppositely charged lepton in the same hemisphere for the 1994-1995 data set. The dots show the data points while the solid lines correspond to the fits described in the text. The shaded histograms show the like-sign combinations.*

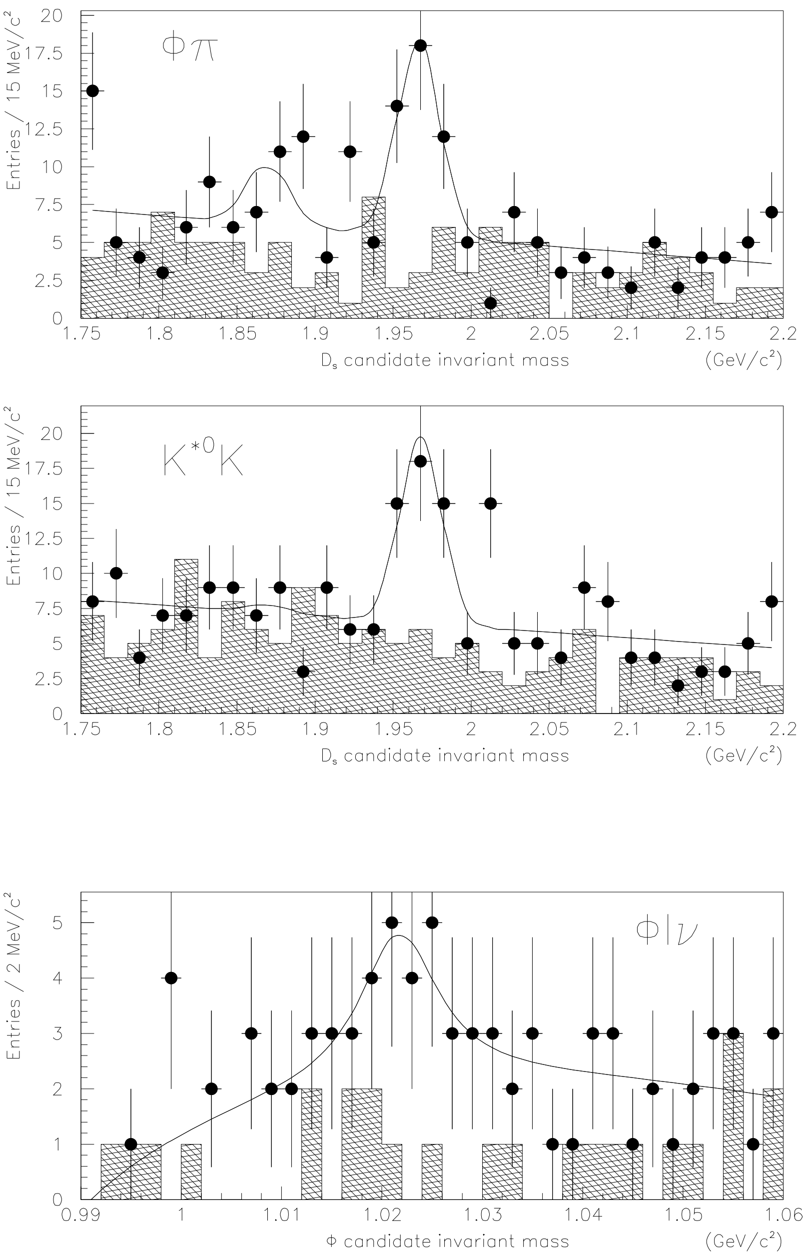


Figure 5.11: *Invariant mass distributions for $D_s^{\mp}$ and ϕ candidates accompanied by an oppositely charged lepton in the same hemisphere for the 1991-1993 data set. The dots show the data points while the solid lines show the fits described in the text. The shaded histograms show the like-sign combinations.*

D_s^- decay mode	Estimated signal	Combinatorial background / total (%)	D_s (ϕ) width (MeV)
$D_s^- \longrightarrow \phi\pi^-$	67 ± 11	38 ± 6	15.1 ± 2.5
$D_s^- \longrightarrow K^{*0}K^-$	70 ± 17	43 ± 7	18.4 ± 7.8
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	11 ± 5	44 ± 15	12.4 ± 3.8
$D_s^- \longrightarrow K^{*0}K^{*-}$	21 ± 6	28 ± 8	8.9 ± 3.9
$D_s^- \longrightarrow \phi\pi^-\pi^0$	18 ± 6	40 ± 18	28.2 ± 9.7
All hadronic modes	181 ± 20	41 ± 4	15.9 ± 2.0
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	78 ± 18	47 ± 8	8.2 ± 2.4

Table 5.4: *Number of $D_s^{\mp}l^{\pm}$ signal events in the entire data set, ratio of the combinatorial background to the total number of candidates within $\pm 2\sigma$ ($\pm 1.5\Gamma$) of the D_s (ϕ) mass peak, and the D_s (ϕ) width.*

D_s^- decay mode	Estimated signal	Combinatorial background / total (%), f_{comb}	D_s (ϕ) width (MeV)
$D_s^- \longrightarrow \phi\pi^-$	40 ± 8	33 ± 7	17.1 ± 3.5
$D_s^- \longrightarrow K^{*0}K^-$	36 ± 8	37 ± 8	18.9 ± 4.7
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	11 ± 5	44 ± 15	12.4 ± 3.8
$D_s^- \longrightarrow K^{*0}K^{*-}$	21 ± 6	28 ± 8	8.9 ± 3.9
$D_s^- \longrightarrow \phi\pi^-\pi^0$	18 ± 6	40 ± 18	28.2 ± 9.7
All hadronic modes	126 ± 16	38 ± 5	17.6 ± 2.3
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	58 ± 14	38 ± 9	7.9 ± 2.4

Table 5.5: *Number of $D_s^{\mp}l^{\pm}$ signal events in the 1994-1995 data set, ratio of the combinatorial background to the total number of candidates within $\pm 2\sigma$ ($\pm 1.5\Gamma$) of the D_s (ϕ) mass peak, and the D_s (ϕ) width.*

D_s^- decay mode	Estimated signal	Combinatorial background / total (%), f_{comb}	D_s (ϕ) width (MeV)
$D_s^- \longrightarrow \phi\pi^-$	29 ± 7	41 ± 10	13.5 ± 3.3
$D_s^- \longrightarrow K^{*0}K^-$	29 ± 8	45 ± 8	13.2 ± 3.6
Both hadronic modes	58 ± 11	43 ± 6	12.9 ± 2.3
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	20 ± 10	64 ± 12	9.5 ± 5.7

Table 5.6: *Number of $D_s^{\mp}l^{pm}$ signal events in the 1991-1993 data set, ratio of the combinatorial background to the total number of candidates within $\pm 2\sigma$ ($\pm 1.5\Gamma$) of the D_s (ϕ) mass peak, and the D_s (ϕ) width.*

D_s^- decay mode	D_s (ϕ) mass resolution (MeV/c^2)	Reconstruction efficiency (%)	Expected B_s^0 signal in data
$D_s^- \longrightarrow \phi\pi^-$	13.8	12.0 $\pm$ 0.2	35 $\pm$ 12
$D_s^- \longrightarrow K^{*0}K^-$	14.9	6.5 $\pm$ 0.2	23 $\pm$ 8
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	7.9	1.0 $\pm$ 0.1	1 $\pm$ 0.6
$D_s^- \longrightarrow K^{*0}K^{*-}$	12.1	5.4 $\pm$ 0.2	15 $\pm$ 7
$D_s^- \longrightarrow \phi\pi^-\pi^0$	25.8	0.7 $\pm$ 0.1	5 $\pm$ 3
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	6.5	8.3 $\pm$ 0.1	27 $\pm$ 9

Table 5.7: *Expected $D_s(\phi)$ mass resolutions, signal reconstruction efficiencies and expected number of B_s^0 signal events in the 1994-1995 data set (from Monte Carlo simulations and [1]).*

D_s^- decay mode	D_s (ϕ) mass resolution (MeV/c^2)	Reconstruction efficiency (%)	Expected B_s^0 signal in data
$D_s^- \longrightarrow \phi\pi^-$	16.1	13.5 $\pm$ 0.2	30 $\pm$ 10
$D_s^- \longrightarrow K^{*0}K^-$	18.1	5.6 $\pm$ 0.2	15 $\pm$ 5
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	7.5	7.2 $\pm$ 0.3	18 $\pm$ 6

Table 5.8: *Expected $D_s(\phi)$ mass resolution, signal reconstruction efficiencies and expected number of B_s^0 signal events in the 1991-1993 data set (from Monte Carlo simulations and [1]).*

Lepton type	Fraction 1994-1995 (%)	Fraction 1991-1993 (%)
Electrons	43 $\pm$ 4	34 $\pm$ 5
Muons	57 $\pm$ 5	66 $\pm$ 7

Table 5.9: *Fractions of electrons and muons from the B_s^0 candidate in the two data sets.*

5.8.3 *Cascade* Decays

The signal $B_s^0 \longrightarrow D_s^{(*)-} l^+ \nu_l$ is reconstructed by combining a D_s^- candidate with an oppositely charged lepton. The following processes give the same signature and, therefore, look like the signal:

$$\overline{B_d^0} \longrightarrow D^{(*)+} D_s^-$$

$$B^- \longrightarrow D^{(*)0} D_s^-$$

where:

$$\begin{array}{ccc} D^{*+} & \longrightarrow & D^0 X,\ D^+ X \\ D^{*0} & \longrightarrow & D^0 X \end{array}$$

and:

$$D^+ \longrightarrow l^+ \nu_l X$$

$$D^0 \longrightarrow l^+ \nu_l X$$

i.e. B_d^0 and B^- can decay into two charmed mesons, with the D_s^- decaying in the considered signal mode and a lepton originating from the other charmed meson. For these two modes the charges are "flipped" compared to the signal, i.e. a negatively charged lepton does no longer indicate a b-quark, but a $\bar{b}$ (see figures 5.12 and 4.2).

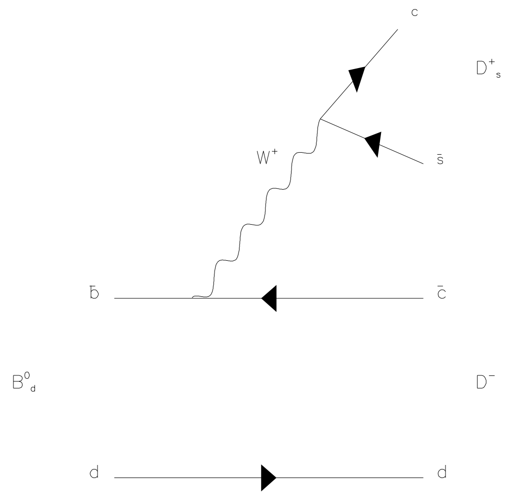


Figure 5.12: $B_d^0 \longrightarrow D^- D_s^+$ *Feynman diagram.*

Mode	Reconstruction efficiency (%)	Relative fraction with respect to the signal (%)
$\overline{B_d^0} \longrightarrow D^{(*)+} D_s^-$	1.37 ± 0.05	5.2 ± 2.3
$B^- \longrightarrow D^{(*)0} D_s^-$	1.30 ± 0.06	5.4 ± 2.3

Table 5.10: $\overline{B_d^0}$ *and* B^- cascade *decay reconstruction efficiencies and relative fractions with respect to the signal for the hadronic modes as expected from Monte Carlo simulations.*

Mode	Reconstruction efficiency (%)	Relative fraction with respect to the signal (%)
$\overline{B_d^0} \longrightarrow D^{(*)+} D_s^-$	1.14 ± 0.04	5.9 ± 2.6

Table 5.11: $\overline{B_d^0}$ *and* B^- cascade *decay reconstruction efficiencies and relative fractions with respect to the signal for the semileptonic modes as expected from Monte Carlo simulations.*

The reconstruction efficiencies for the modes are given in table 5.10. They were calculated for the mode $D_s^- \longrightarrow \phi\pi^-$ and are assumed to be the same for all the hadronic modes and the two data sets. This is reasonable since the selection criteria on the invariant mass and momentum of the $D_s^- l^+$ combinations only differ for the hadronic and semileptonic decay modes of the D_s^- meson and are the same for the two data sets.

Table 5.11 gives the corresponding numbers for the semileptonic modes. Based on what was seen for the hadronic modes, only the contribution from the B_d^0 was calculated; the contribution from the B^- was assumed to be identical.

The uncertainties take into account the statistical uncertainty coming from the Monte Carlo simulations and the uncertainties on the relevant branching fractions [1].

5.8.4 Kinematic *Reflections*

In addition to the mode independent *cascade* decay background, there are other physics processes which can look like the signal. They are called kinematic *reflections*, and they only concern the mode

$$D_s^- \longrightarrow K^{*0} K^-$$

If a pion is misidentified as a kaon we may mistakenly identify a D^- in the modes:

$$\begin{aligned} D^- &\longrightarrow K^{*0}(K^+\pi^-)\,\pi^- \\ D^- &\longrightarrow K^+\pi^-\pi^- \text{ (non resonant)} \end{aligned}$$

for a D_s^-.

D^- decay mode	Reconstruction efficiency (%)	Relative fraction with respect to the signal (%)
$D^- \longrightarrow K^{*0}(K^+\pi^-)\pi^-$	0.86 ± 0.08	12.3 ± 5.5
$D^- \longrightarrow K^+\pi^-\pi^-$	0.13 ± 0.03	12.4 ± 5.5

Table 5.12: Reflection *decay reconstruction efficiencies and relative fractions with respect to the signal decay mode* $D_s^- \longrightarrow K^{*0}K^-$ *as expected from from Monte Carlo simulations.*

Instead of reconstructing

$$B_s^0 \longrightarrow D_s^{(*)-} l^+ \nu_l$$

we reconstruct

$$B_d^0 \longrightarrow D^{(*)-} l^+ \nu_l$$

It was explicitly verified that the probability to reconstruct these two $K\pi\pi$ modes as a $D_s^- \longrightarrow \phi\pi^-$ was zero.

Even though the corresponding reconstructed D_s^- mass is shifted and becomes rather broad, a certain fraction of the events still fall within the selected mass range, as can be seen in figure 5.13. Note that for this background the charges are not "flipped" as with the *cascade* background.

The reconstruction efficiencies are given in table 5.12. The relative fraction with respect to the signal is 24.7%. The uncertainties take into account both the statistical uncertainty coming from the Monte Carlo simulations and the uncertainties on the relevant branching ratios [1].

5.9 Combinatorial Background Composition

Even though all the relevant parameters used to describe the combinatorial background will be taken from the data itself, it can be useful to know its actual composition. From a generic $q\bar{q}$ Monte Carlo simulation one finds that the combinatorial background almost completely comes from b-quarks, with the ratio of B^+, B_d^0 and B_s^0 mesons approximately following the production ratio [1].

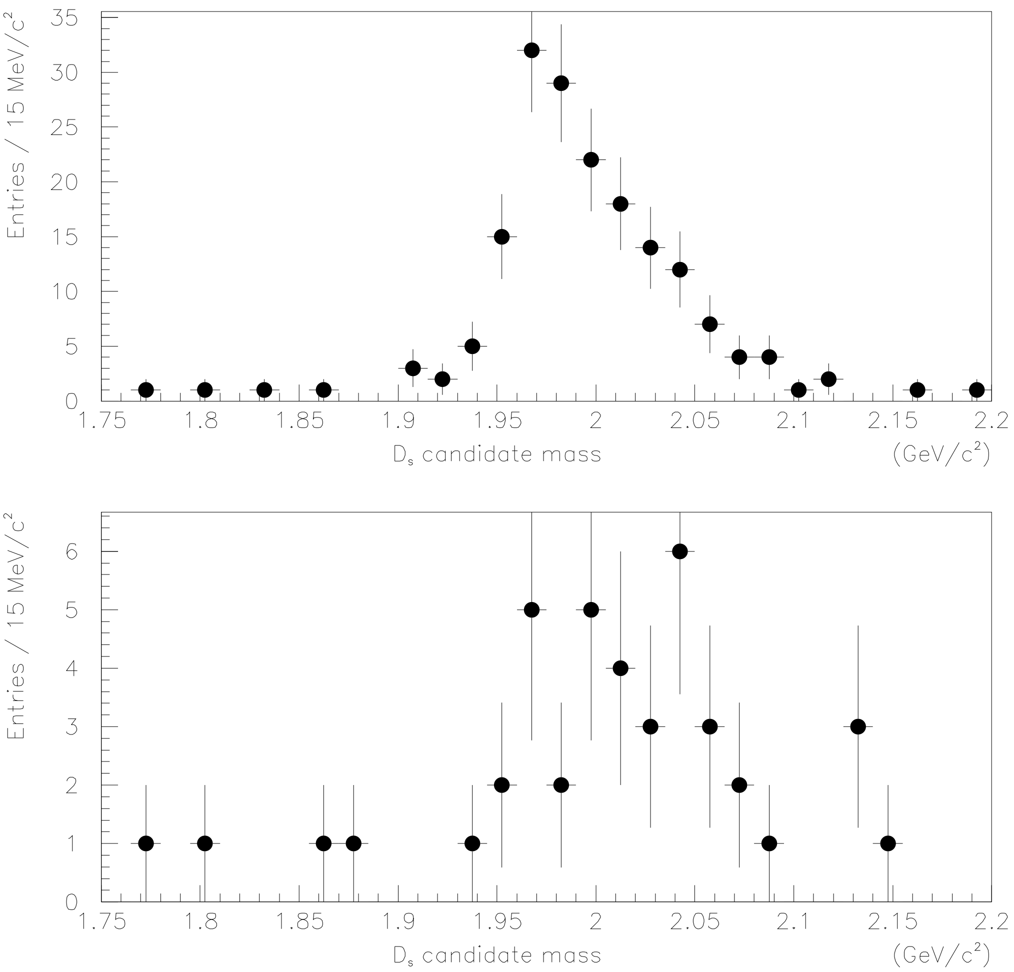


Figure 5.13: *Reconstructed* $K\pi\pi$ *mass for the kinematic* reflection *decays* $K^{*0}\pi^-$ *(top) and the non-resonant* $K^+\pi^-\pi^-$ *(bottom) after all selection criteria, as expected from Monte Carlo simulations.*

Chapter 6

Proper Time Reconstruction

6.1 General Remarks

The measured B_s^0 proper time is given by

$$t = \frac{m_{B_s^0} L_{B_s^0}}{p_{B_s^0}} \tag{6.1}$$

where $m_{B_s^0}$ is the nominal B_s^0 mass, $L_{B_s^0}$ is the measured B_s^0 decay length and $p_{B_s^0}$ is the estimated B_s^0 momentum.

Assuming no correlation between the uncertainty on the measured decay length and the estimated momentum, the uncertainty on the proper time is then given by standard error propagation, yielding

$$\sigma_t = \sqrt{\left(\frac{\sigma_{L_{B_s^0}} m_{B_s^0}}{p_{B_s^0}}\right)^2 + \left(\frac{\sigma_{p_{B_s^0}}}{p_{B_s^0}} t\right)^2} \tag{6.2}$$

This approximation is not always valid as it implies that terms of the type '$\sigma_L \sigma_P$' are neglected versus terms like '$L\sigma_P$'. This can be interpreted as a higher-order term in a Taylor expansion, or equivalently, as a correlation term between momentum and decay length.

Figure 6.1 shows that the decay length resolution does not depend significantly on the decay length. Fitting a line $\sigma_L = a + b \times L$ through the data points in figure 6.1 yields $a = (199 \pm 62)$ μm and $b = -0.002 \pm 0.022$, i.e. the slope is compatible with zero. For small decay lengths this means that $\sigma_L \sim L$ and the approximate formulae is no longer valid.

The proper time resolution is found using another method which will be detailed later. However, the approximate formula is useful since it gives the physical origin of the two contributions. The decay length resolution gives a *constant* (i.e. time independent) term to the proper time resolution while the relative momentum resolution introduces a *time dependent* term. The following two sections will give more details about these two contributions.

None of the following values for the decay length and relative momentum resolutions are used explicitly to find the proper time resolution. In particular, for several distributions a three-Gaussian fit would have been necessary to give a perfect description. The contribution of the third Gaussian is, however, small, and only one- and two-Gaussian fits have been used since they contain the main part of the relevant information. The final proper time distribution is found with an unbinned likelihood fit which takes all this into account.

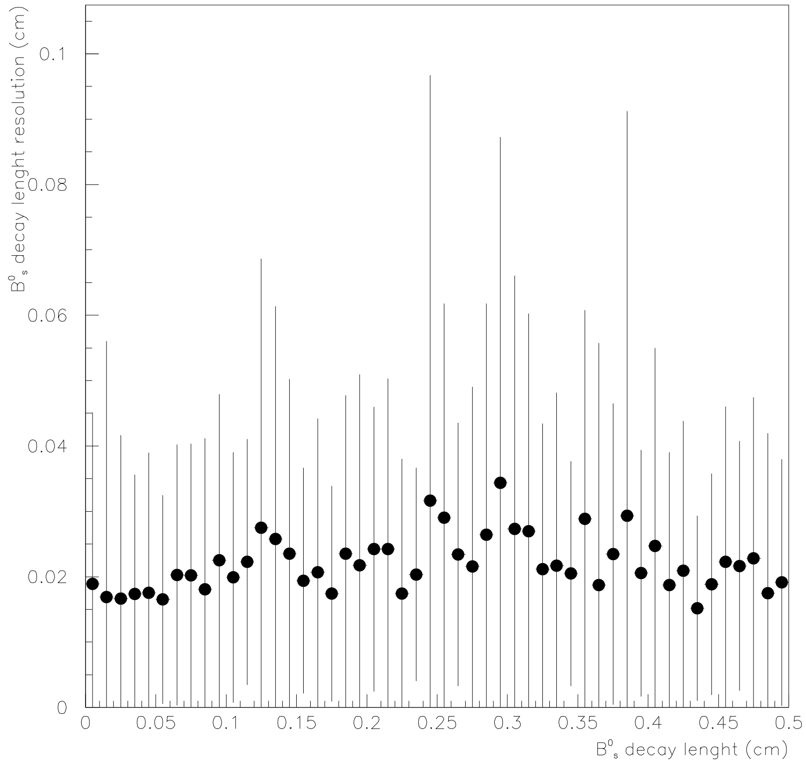


Figure 6.1: *B_s^0 decay length resolution versus the B_s^0 decay length.*

All numbers are found from B_s^0 signal candidates that have passed all the selection criteria. *Linked tracks* have not been used to avoid any possible bias in the linking process between the generated particles and the reconstructed tracks.

6.2 Momentum Resolution

6.2.1 Neutrino Momentum Estimate

The difficulty in estimating the B_s^0 momentum arises from the fact that it is reconstructed in the semileptonic decay modes and, therefore, involves a neutrino which escapes detection. For the two semileptonic modes of the D_s^- there is an additional neutrino. It is, therefore, necessary to estimate the neutrino energy(ies).

The starting point is the relation

$$p_{B_s^0}^2 = (E_{D_s^- l^+} + E_{\nu_l})^2 - m_{B_s^0}^2 \tag{6.3}$$

where the nominal value for $m_{B_s^0}$ is used. The neutrino energy is then estimated from the missing energy in the hemisphere of the B_s^0 candidate, i.e.

D_s decay mode	c_1 (GeV/c)$^{-1}$	c_2
Hadronic	-6.26 ± 0.03	0.115 ± 0.002
Semileptonic	-5.27 ± 0.08	0.090 ± 0.002

Table 6.1: *Linear fit constants for the estimate of the neutrino energy.*

$$E_{\text{missing}} = E_{\text{total}} - E_{\text{visible}} \tag{6.4}$$

The visible energy E_{visible} is taken to be the sum of the energies of all charged particles, using particle identification where possible and else assuming the tracks to be pions. Photons are defined to be all neutral energy clusters in the electromagnetic calorimeters not associated to a track. The information from the hadronic calorimeter is not precise enough to be useful here.

The total energy in the hemisphere, E_{total}, is found using total energy-momentum conservation in the event and can be written as

$$E_{\text{total}} = E_{\text{beam}} + \frac{M^2_{\text{same}} - M^2_{\text{opposite}}}{4E_{\text{beam}}} \tag{6.5}$$

where M^2_{same} and M^2_{opposite} are the invariant masses in the same and opposite hemispheres, respectively. The correction factor to the beam energy is small, as can be seen in figure 6.2a. However, its contribution increases for events with more than two jets, i.e. events where a large fraction of the center-of-mass energy may be carried away by hard gluon radiation (see figure 6.2b). The mean number of jets in B^0_s signal events is 2.8 (see figure 6.2c).

It is assumed that there is a linear relation between the neutrino energy and the $D_s^- l^+$ energy (see figure 6.3):

$$E_\nu = E_{\text{missing}} + c_1 + c_2 \times E_{D_s^- l^+} \tag{6.6}$$

The constants c_1 and c_2 were estimated from Monte Carlo simulations and are given in table 6.1. The coefficients are the same for all the D_s^- hadronic decays, but are different for the semileptonic decays because of the second neutrino. It was verified that the constants are the same for the two data sets.

The final neutrino energy resolution was

$$\sigma_{E_\nu} = E_{\text{reconstructed}} - E_{\text{generated}} = 2.55 \text{ GeV}$$

The neutrino energy in the D_s^- semileptonic decays was estimated in the same way, except that here E_ν represented the total energy of the two neutrinos: $E_\nu = E_{\nu_{B^0_s}} + E_{\nu_{D_s^-}}$. The final neutrino energy resolution was the same as for the hadronic modes.

Table 6.2 gives the corresponding numbers if two Gaussian distributions are used to fit the neutrino energy resolution. This gives a better idea about the core width and any non-Gaussian tails in the fitted distribution (see also figure 6.4).

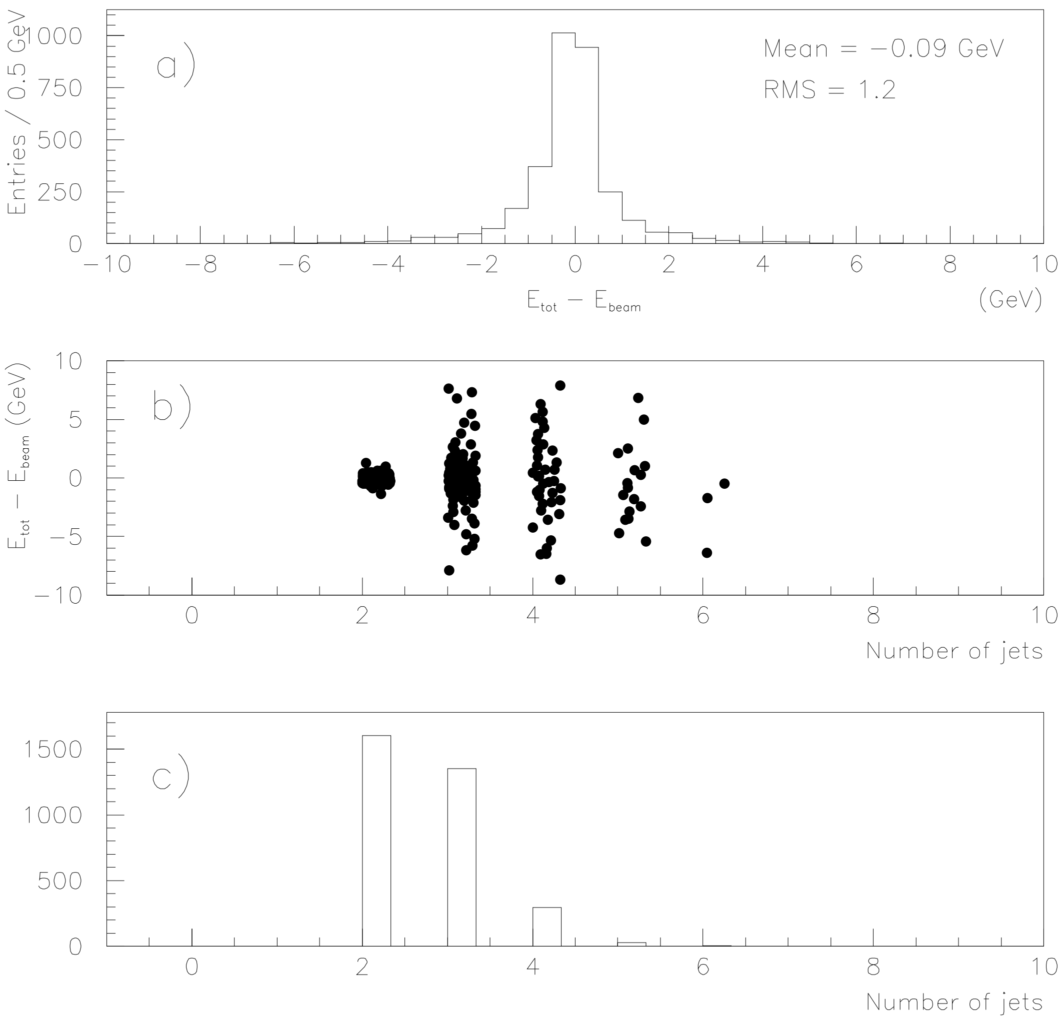


Figure 6.2: *a) The difference between the total hemisphere energy and the beam energy:* $E_{\rm tot} - E_{\rm beam}$*, b)* $E_{\rm tot} - E_{\rm beam}$ *versus the number of jets in* B_s^0 *signal events as expected from Monte Carlo simulations, c) Number of jets in* B_s^0 *signal events as expected from Monte Carlo simulations.*

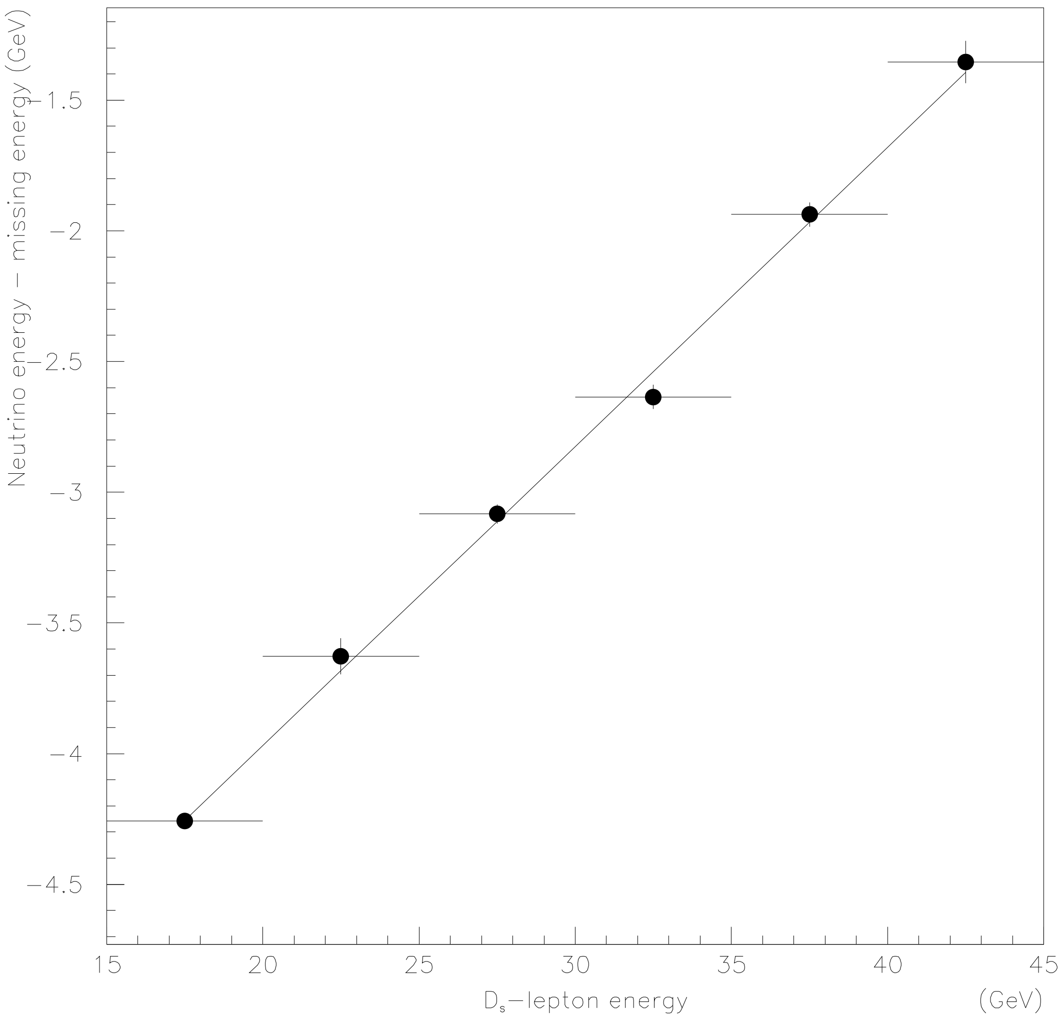


Figure 6.3: *The difference between the neutrino energy and the missing energy versus the* $D_s^- l^+$ *energy. The line is the fit described in the text.*

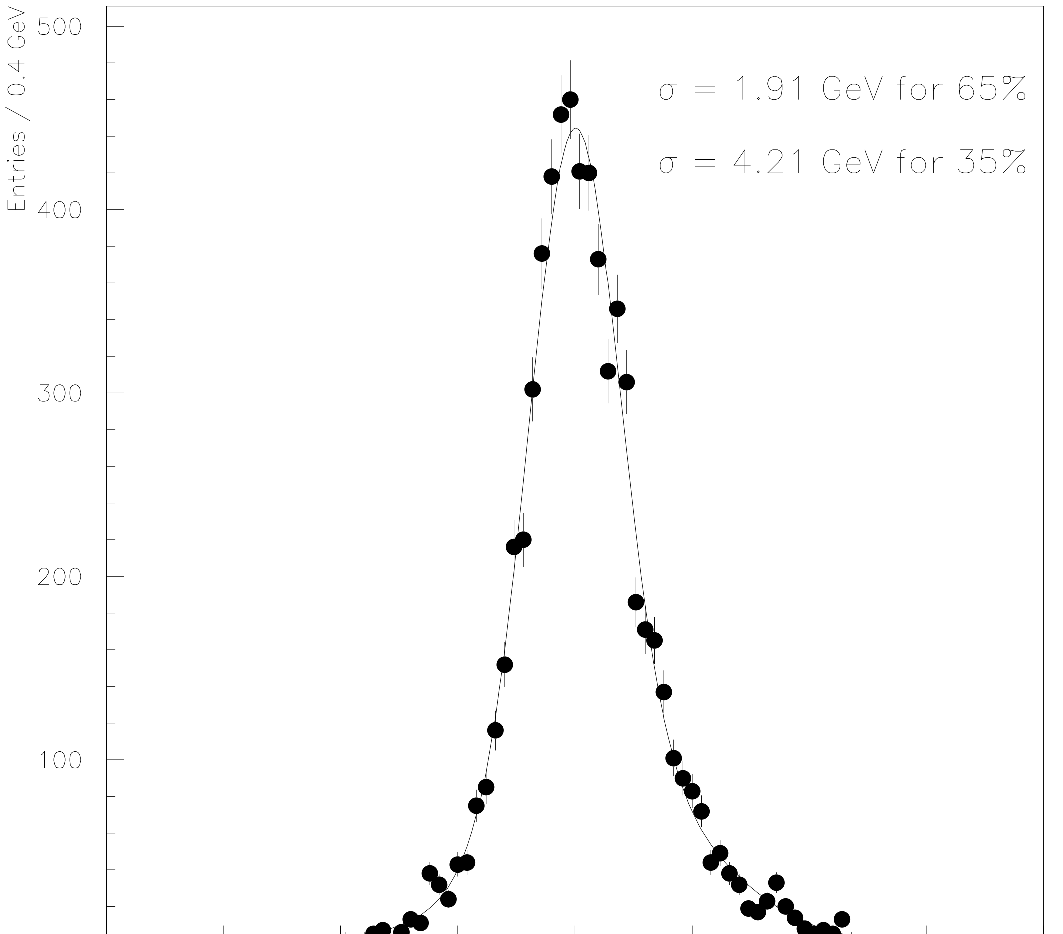


Figure 6.4: *Neutrino energy resolution: $E_{\nu_{\rm rec}} - E_{\nu_{\rm gen}}$. The solid line is the fit described in the text.*

Neutrino energy resolution σ_{E_ν} (GeV)	First Gaussian σ_{E_ν} (GeV)	First Gaussian fraction (%)	Second Gaussian σ_{E_ν} (GeV)
2.55 ± 0.03	1.91 ± 0.05	65	4.21 ± 0.11

Table 6.2: *Neutrino energy resolution $E_{\nu_{\rm rec}} - E_{\nu_{\rm gen}}$. Numbers are given for both a single Gaussian and a double Gaussian fit.*

6.2.2 B_s^0 Momentum Estimate

One additional complication in estimating the B_s^0 momentum comes from the fact that in a ratio of 3:1 (obtained from naive spin counting) the B_s^0 decays into a D_s^{*-} instead of a D_s^-, i.e.

$$B_s^0 \longrightarrow D_s^{*-} l^+ \nu_l$$

The small mass difference between the D_s^{*-} and the D_s^- implies that only the radiative decay

$$D_s^{*-} \longrightarrow D_s^- \gamma$$

and the hadronic decay

$$D_s^{*-} \longrightarrow D_s^- \pi^0$$

are kinematically allowed. The latter violates isospin, however, and is, therefore, suppressed, contributing at the 6% level [66]. It is neglected in the following.

The mean photon energy in the radiative decay is 1 GeV (see figure 6.5). To take this into account, the photon with the highest energy yielding a D_s^{*-} invariant mass within 200 MeV/c^2 of the nominal D_s^{*-} mass and lying within 30^0 of the D_s^- direction was added to the D_s^- candidate. A photon is thus assigned to 62% of the D_s^- candidates. The mean photon candidate energy was 1.44 GeV. This introduces a small bias in the final B_s^0 momentum in the sense that there is a small tendency to overestimate the B_s^0 momentum. This is reflected in the enhancement on the upper side of the relative B_s^0 momentum resolution (see figure 6.6 and also figures 6.7 and 6.8).

The final relative B_s^0 momentum resolution was

$$\frac{\sigma p}{p} = \frac{p_{\rm reconstructed} - p_{\rm generated}}{p_{\rm generated}} = 6.6\%$$

It can be seen in tables 6.3 and 6.4 that the relative momentum resolution does not change significantly from one mode to another, not even for the semileptonic modes. The only exception is the mode $D_s^- \longrightarrow \phi\pi^-\pi^0$ mode where the presence of the π^0 broadens the second Gaussian component. The relative momentum resolution is the same for the two data sets.

By not correcting for the photon from the D_s^{*-} the resolution increases from 5.4% to 6.2% for the narrow Gaussian in the $D_s^- \longrightarrow \phi\pi^-$ mode. In addition, it is centered at -0.026 instead of -0.004, i.e. the photon-correction both improves the overall relative momentum resolution

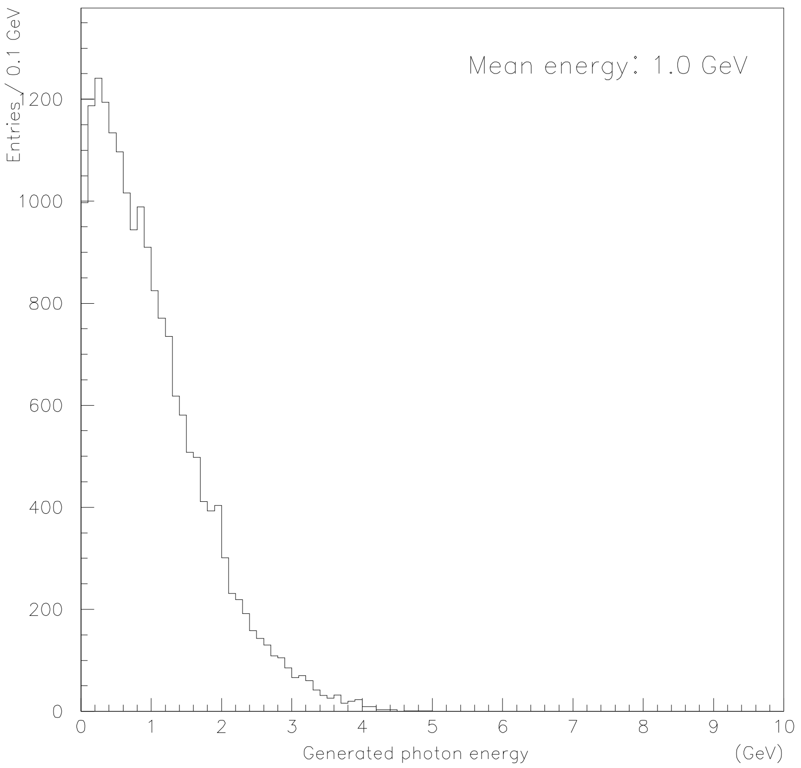


Figure 6.5: *Generated energy for photons from the decay $D_s^{*-} \longrightarrow D_s^- \gamma$ in $B_s^0 \longrightarrow D^{*-} l^+ \nu_l$ events.*

and reduces any bias in the momentum reconstruction. Figure 6.7 shows the improvement on the D_s^- momentum distribution by adding a photon. Figure 6.8 shows the reconstructed D_s^{*-} mass obtained from Monte Carlo simulations.

The relative B_s^0 momentum resolution versus the B_s^0 momentum is given in figure 6.10. It is a decreasing function of the B_s^0 momentum:

$$\frac{\sigma_p}{p} = a - b \times p \tag{6.7}$$

The fit gives $a = 0.316 \pm 0.008$ and $b = 0.0060 \pm 0.0002$.

Figure 6.9 shows the comparison between the signal momentum distributions for data and Monte Carlo simulations. The estimate from real data has been obtained by subtracting the momentum distribution for events in the side band and those for like-sign events from the momentum distribution of events in the signal region. The result is corrected for *cascade* contribution obtained from Monte Carlo simulations.

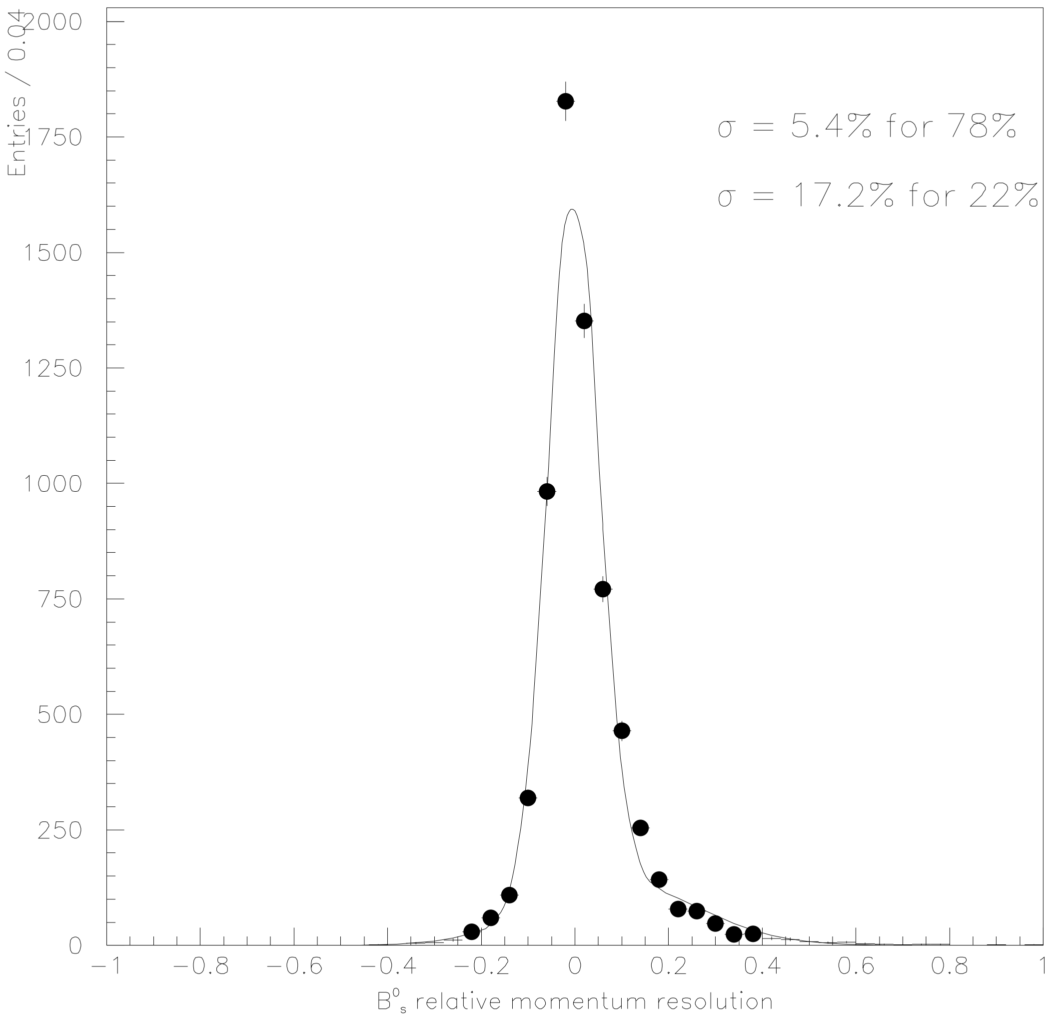


Figure 6.6: *Relative B_s^0 momentum resolution.*

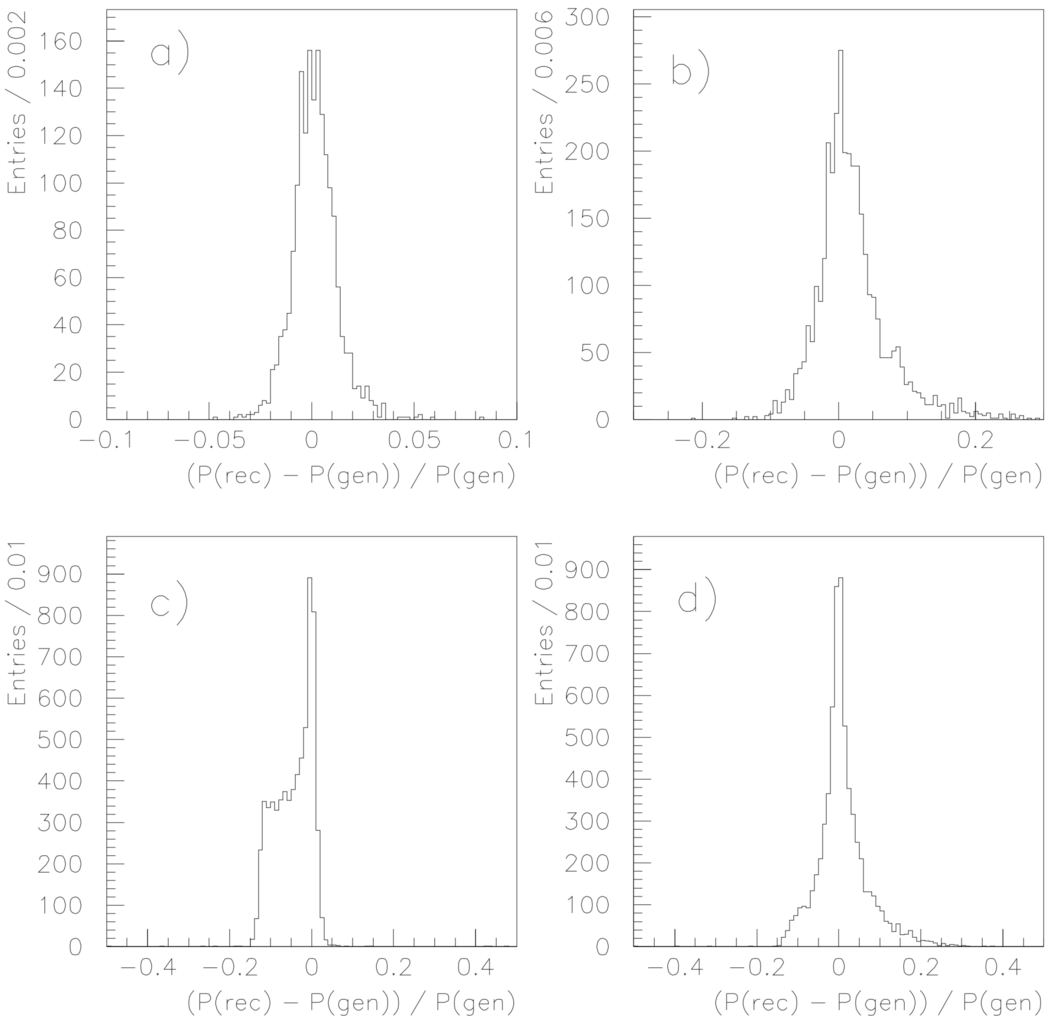


Figure 6.7: *Relative $D_s^{(*)-}$ momentum resolution of a) a D_s^- which was reconstructed as such, b) a D_s^{*-} which was reconstructed as such, c) the overall $D_s^{(*)-}$ momentum resolution without any photon correction and d) the overall $D_s^{(*)-}$ momentum resolution with the photon correction.*

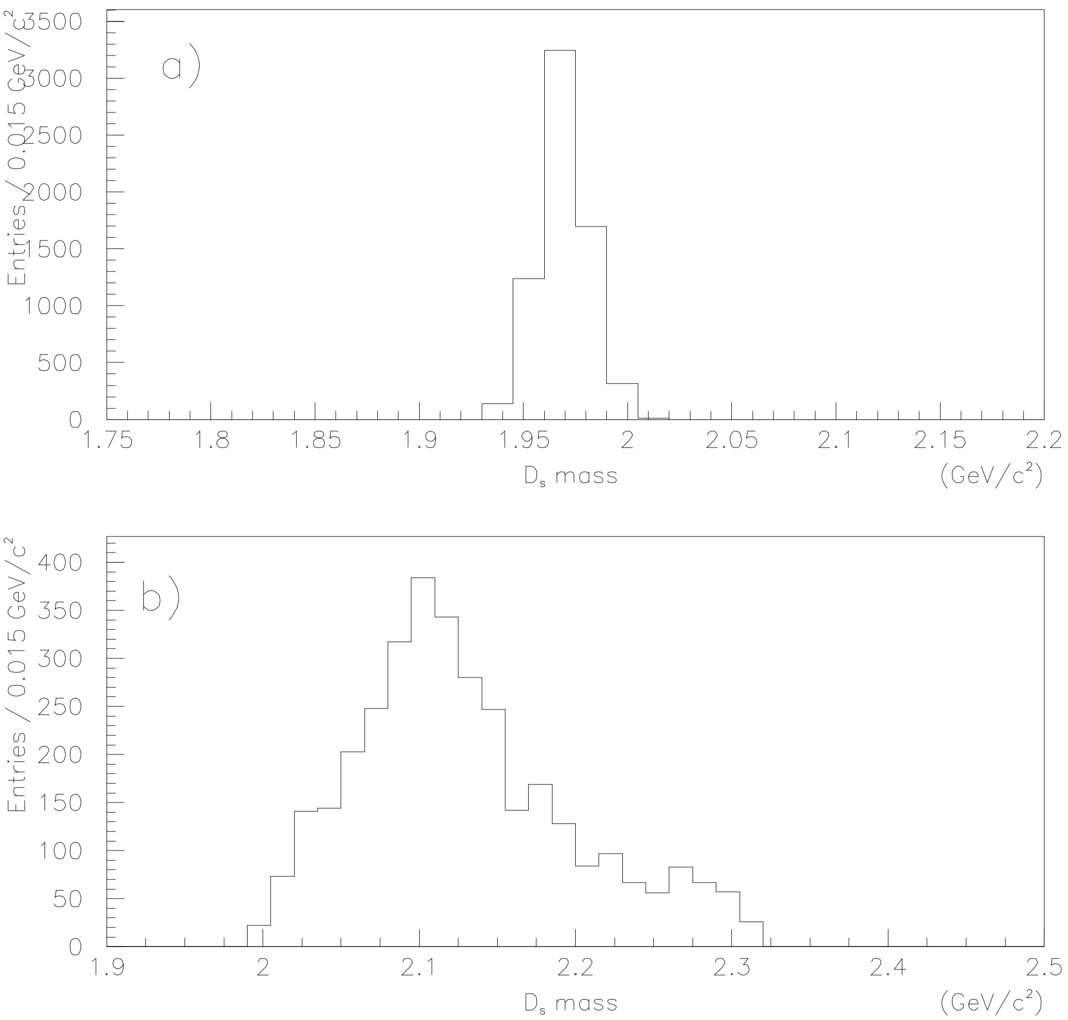


Figure 6.8: *Reconstructed D_s^- (a) and D_s^{*-} mass (b) as expected from Monte Carlo simulations. The fact that the added photon is in average too energetic is seen as an enhancement on the upper side of the D_s^{*-} mass distribution.*

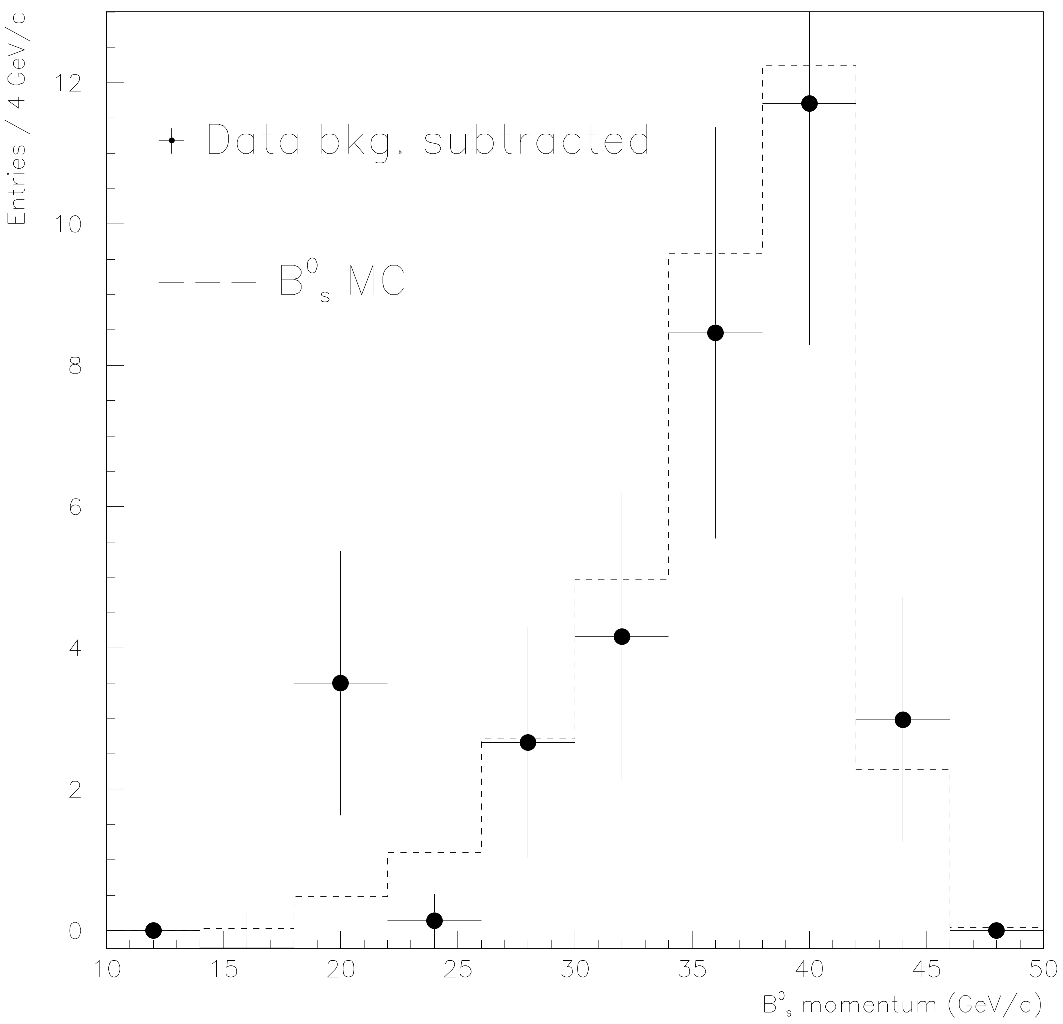


Figure 6.9: *Comparison between the signal momentum distributions for data and Monte Carlo simulations.*

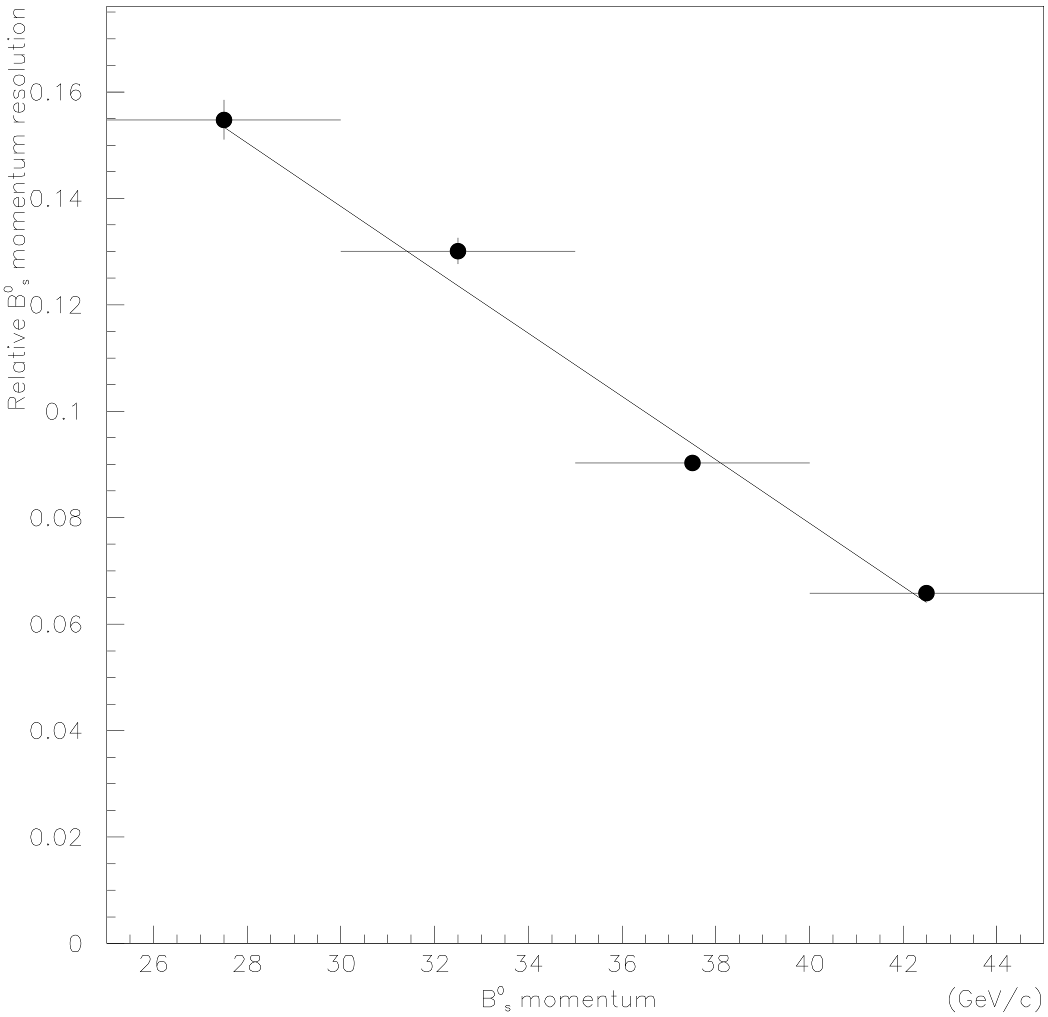


Figure 6.10: *Relative B_s^0 momentum resolution versus the B_s^0 momentum. The line is the fit described in the text.*

D_s^- decay mode	Single Gaussian σ_p/p (%)	First Gaussian σ_p/p (%)	First Gaussian fraction (%)	Second Gaussian σ_p/p (%)
$D_s^- \longrightarrow \phi\pi^-$	6.6 ± 0.1	5.4 ± 0.1	78	17.2 ± 0.4
$D_s^- \longrightarrow K^{*0}K^-$	6.5 ± 0.2	5.3 ± 0.3	78	14.1 ± 0.8
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	7.6 ± 0.4	6.2 ± 0.5	79	16.4 ± 1.5
$D_s^- \longrightarrow K^{*0}K^{*-}$	7.2 ± 0.3	6.0 ± 0.3	79	17.3 ± 1.2
$D_s^- \longrightarrow \phi\pi^-\pi^0$	7.7 ± 0.5	5.2 ± 0.5	58	15.0 ± 1.1
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	6.3 ± 0.1	5.1 ± 0.1	75	17.8 ± 0.4

Table 6.3: *Relative B_s^0 momentum resolution for the 1994-1995 data set.*

D_s^- decay mode	Single Gaussian σ_p/p (%)	First Gaussian σ_p/p (%)	First Gaussian fraction (%)	Second Gaussian σ_p/p (%)
$D_s^- \longrightarrow \phi\pi^-$	6.3 ± 0.1	5.1 ± 0.1	76	16.1 ± 0.5
$D_s^- \longrightarrow K^{*0}K^-$	5.7 ± 0.2	5.1 ± 0.2	82	15.5 ± 1.1
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	6.4 ± 0.3	5.2 ± 0.4	80	15.8 ± 1.4

Table 6.4: *Relative B_s^0 momentum resolution for the 1991-1993 data set.*

6.3 Decay Length Resolution

6.3.1 Decay Length Estimate

A three-dimensional primary interaction vertex was fitted in each event with an iterative procedure [67] using all tracks with at least one hit in the Vertex Detector. The tracks associated with the B_s^0-candidate were excluded from the fit to avoid any bias in the vertex reconstruction. The interaction region [68] was used as a constraint in the fit. In the few cases where the primary vertex fit failed, the interaction region itself was taken as the primary vertex.

The interaction region coordinates are known with different precision (see figure 6.11). The y coordinate is well known and this good precision is reflected in the corresponding primary vertex y coordinate (figure 6.12). Even though the interaction region z coordinate is poorly known, the precision on the primary vertex z coordinate is as good as that of the x coordinate. This shows the importance of having tracking detectors providing good z track-information such as double sided silicon vertex detectors. This is further demonstrated by the corresponding primary vertex coordinates resolutions for the 1991-1993 data set (figure 6.13). For this data set, only a single sided vertex detector was available. The precision on the x and y coordinates are quite similar to those for the 1994-1995 data set while the precision on the z coordinate is orders of magnitude worse.

The decay length is defined as the distance between the primary interaction point and the $D_s^- l^+$ vertex projected onto the $D_s^- l^+$ momentum. This procedure avoids any bias in the decay length [56]. Since the missing neutrino does not alter the direction significantly,

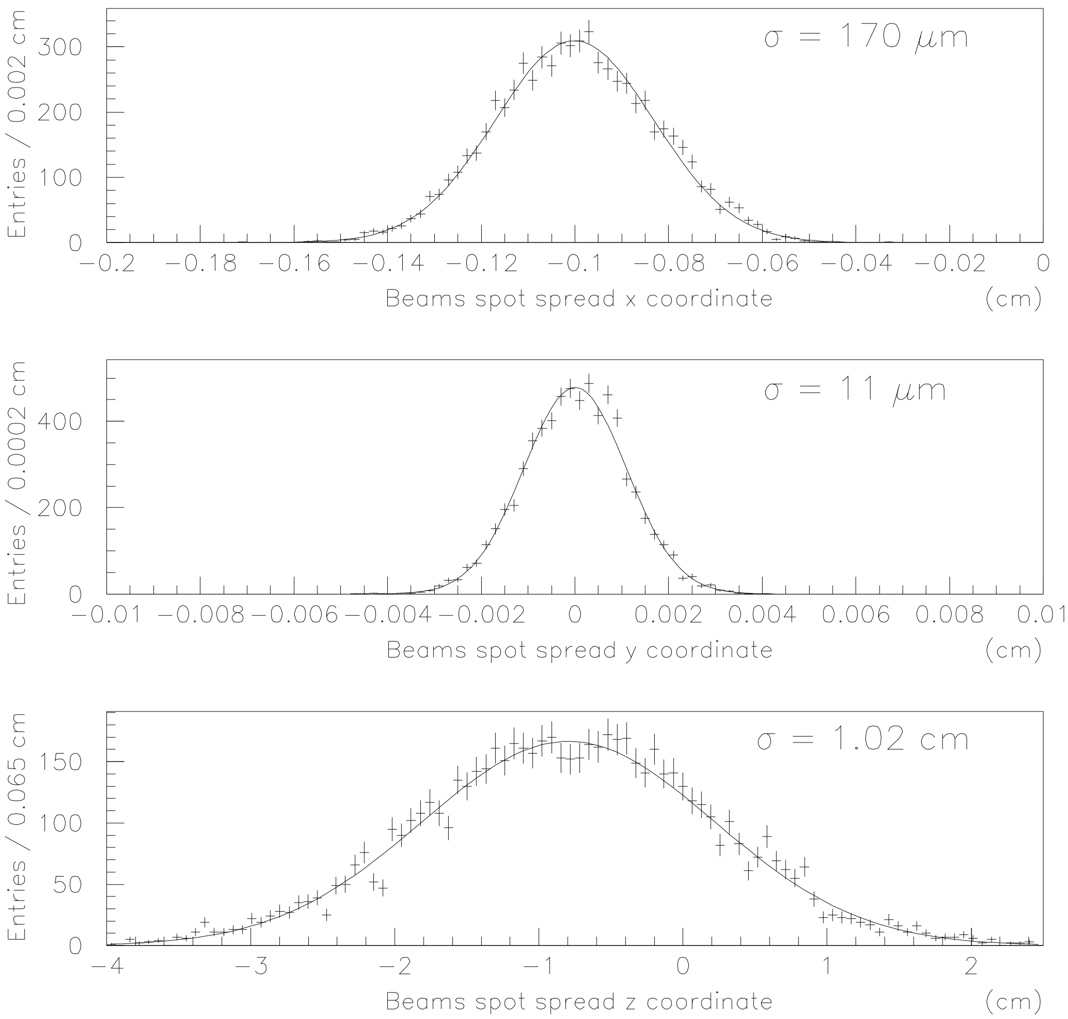


Figure 6.11: *Interaction region coordinates spread for the 1994-1995 data set as expected from Monte Carlo simulations. The solid line is a Gaussian fit.*

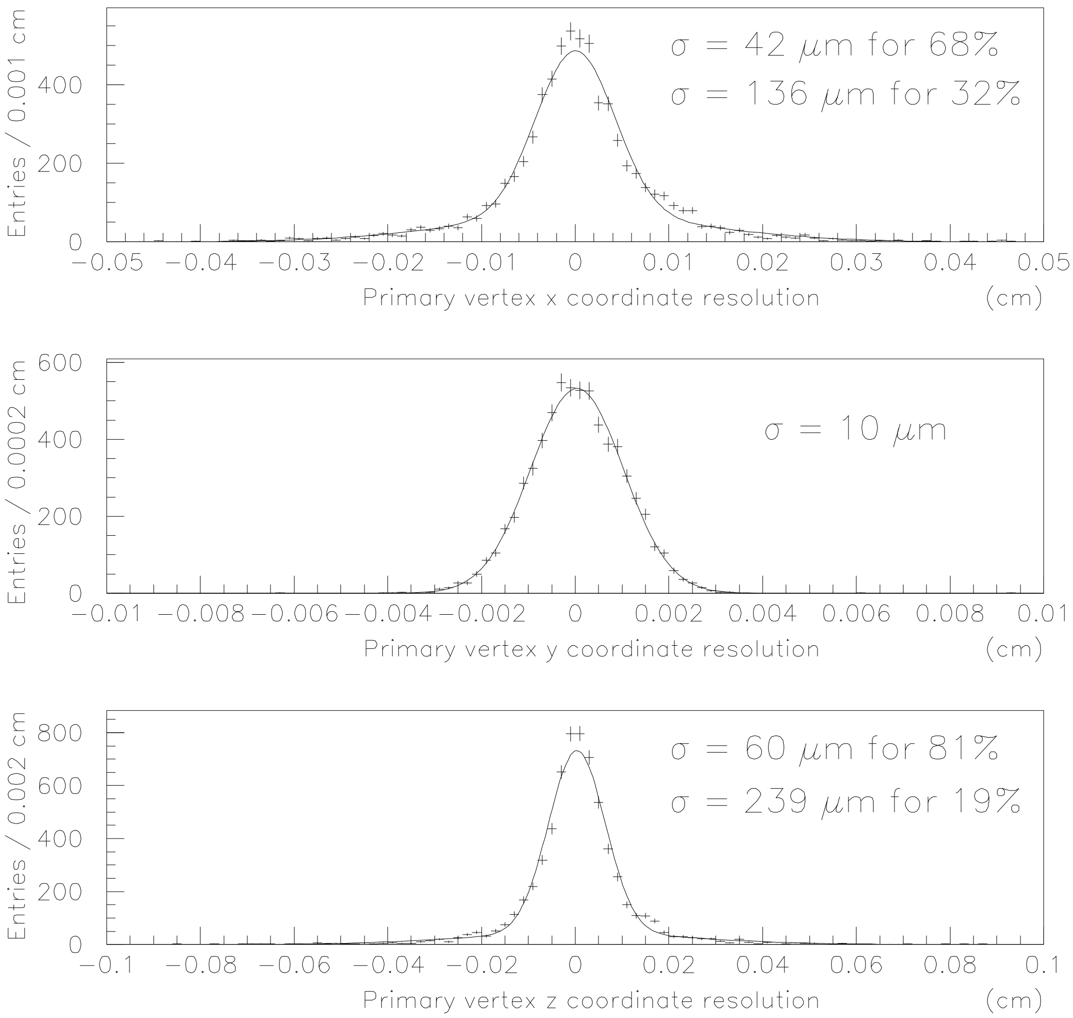


Figure 6.12: *Primary vertex coordinates resolution for the 1994-1995 data set. The solid line is a two (one) Gaussian fit.*

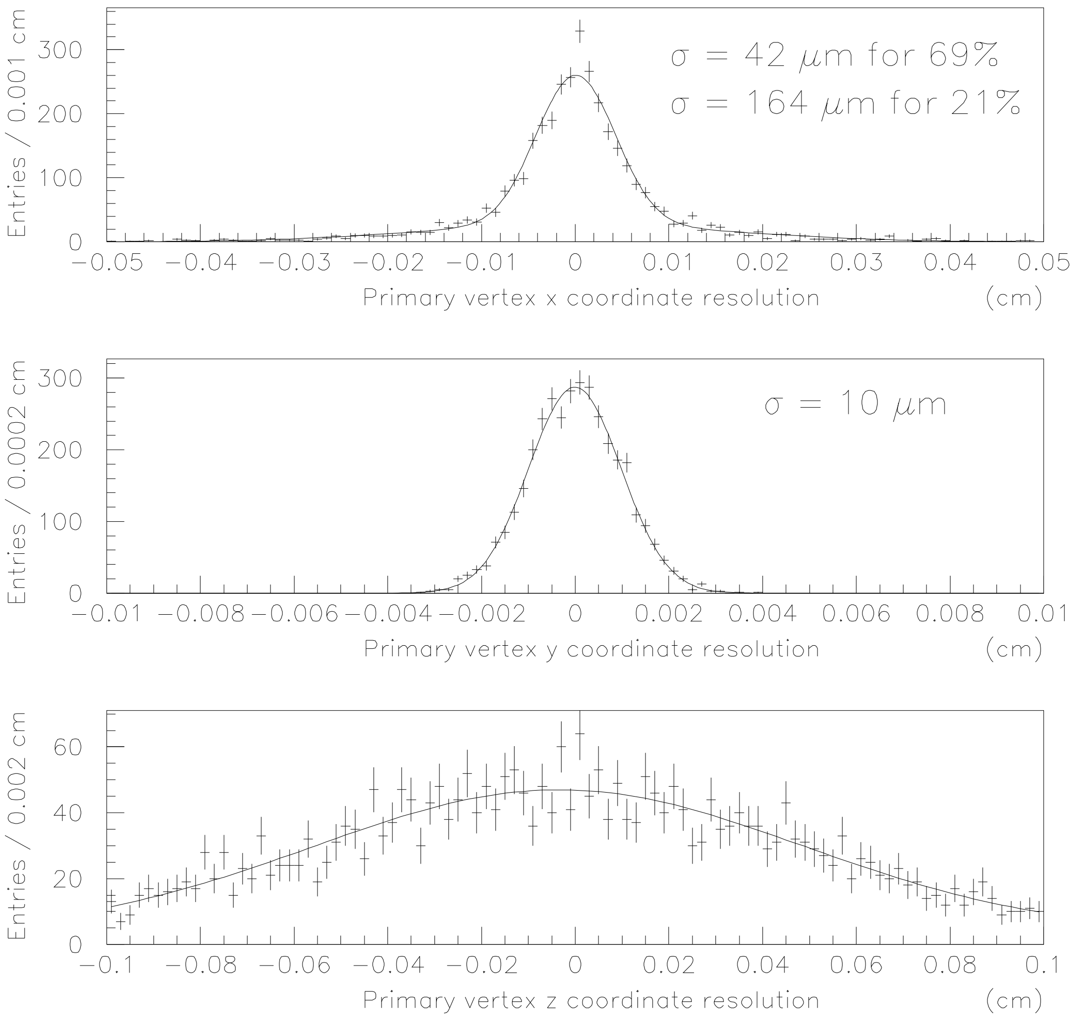


Figure 6.13: *Primary vertex coordinates resolution for the 1991-1993 data set. The solid line is a two (one) Gaussian fit.*

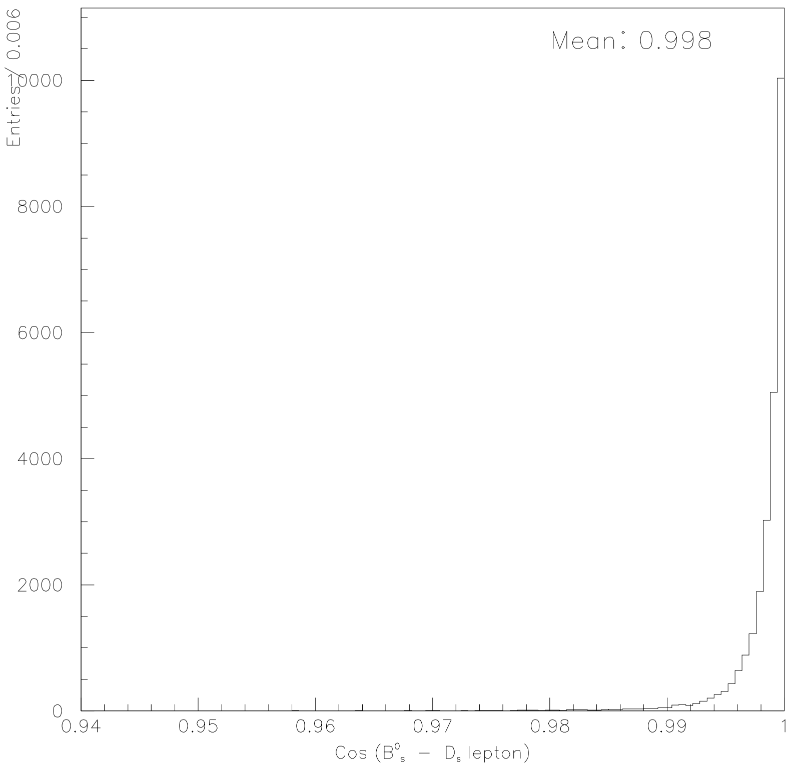


Figure 6.14: *Generated distribution of* $\cos\theta_{B^0_s - D^-_s l^+}$, *where* $\theta_{B^0_s - D^-_s l^+}$ *is the angle between the* B^0_s *and the* $D^-_s l^+$.

the $D^-_s l^+$ direction is a good estimate of the B^0_s direction. This can been seen in figure 6.14, which shows the angle between the B^0_s and the $D^-_s l^+$. It is peaked at small angles.

Table 6.5 gives the obtained decay length resolutions for the different modes for the 1994-1995 data set. For the modes with only charged tracks the core width is around 220 μm while for the modes with one particle missing in the vertex fit it is 20% worse.

The enhanced core width in the mode $D^-_s \longrightarrow \phi\pi^-\pi^+\pi^-$ (which subsequently is reflected onto the time-integrated proper time resolution) is due to the fact that the distribution contains a narrow peak around 0 which the binned fit is not able to pick up. One alternative would be to use a third Gaussian distribution in the fit. However, the proper time resolution is obtained from an unbinned fit and gives the same proper time resolution as for the other equivalent modes (see table 6.7).

Figure 6.16 shows the difference between the reconstructed and generated primary vertex projected onto the $D^-_s l^+$ momentum. A two-Gaussian fit gives a core resolution of

$$\sigma = 36.7\ \mu\text{m}$$

for 66.7% of the events, and $\sigma = 141.2\ \mu$m for the second Gaussian. This indicates that the main contribution to the B^0_s decay length resolution comes from the $D^-_s l^+$ vertex and not

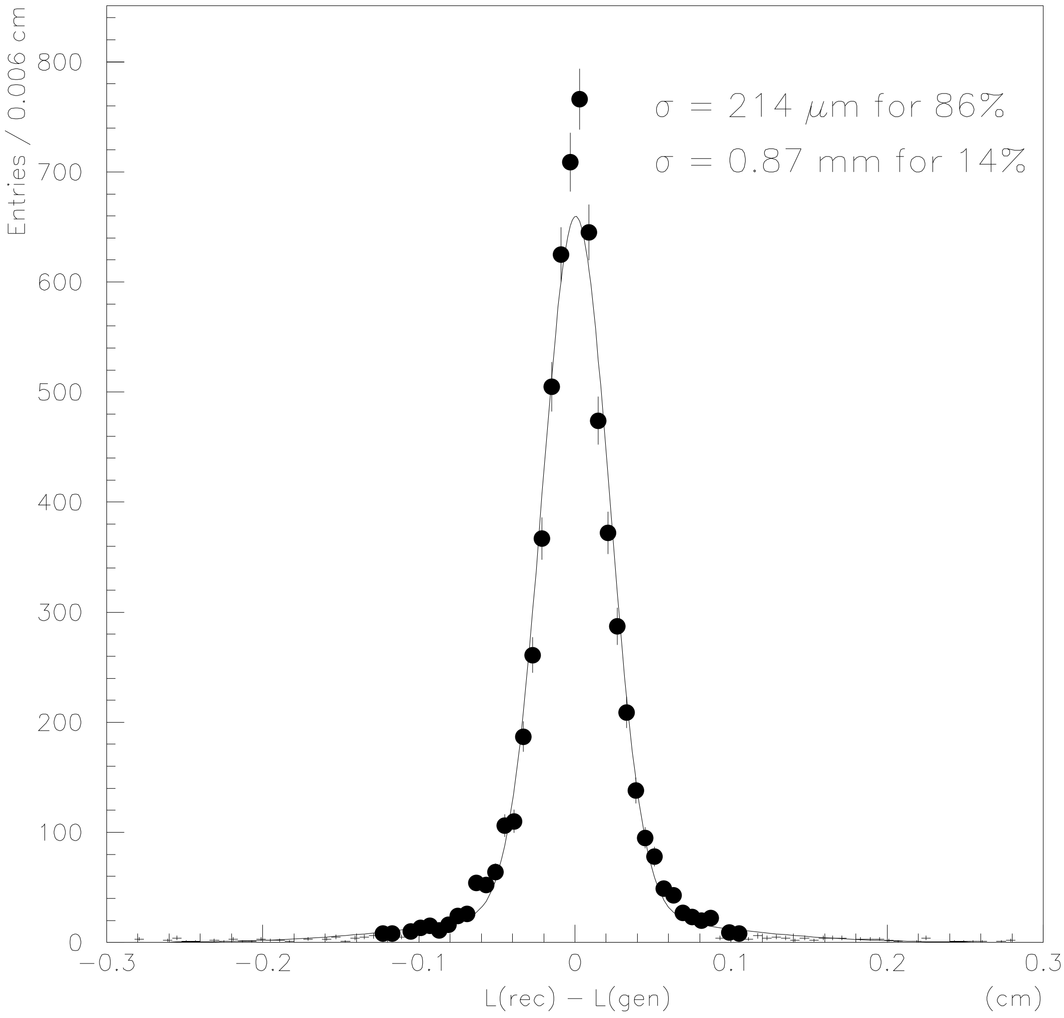


Figure 6.15: B_s^0 *decay length resolution. The solid line is the fit described in the text.*

D_s^- decay modes	Single Gaussian σ_L (μm)	First Gaussian σ_L (μm)	First Gaussian fraction (%)	Second Gaussian σ_L (mm)
$D_s^- \longrightarrow \phi\pi^-$	236 ± 3	214 ± 3	86	0.87 ± 0.03
$D_s^- \longrightarrow K^{*0}K^-$	224 ± 7	223 ± 8	90	1.2 ± 0.1
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	253 ± 13	230 ± 16	78	0.93 ± 0.09
$D_s^- \longrightarrow K^{*0}K^{*-}$	299 ± 12	248 ± 12	78	0.92 ± 0.06
$D_s^- \longrightarrow \phi\pi^-\pi^0$	286 ± 3	253 ± 4	84	0.95 ± 0.03
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	289 ± 5	253 ± 5	84	0.90 ± 0.04

Table 6.5: B_s^0 *decay length resolution with both a single and a two Gaussian fit for the 1994-1995 data set.*

D_s^- decay modes	Single Gaussian σ_L (μm)	First Gaussian σ_L (μm)	First Gaussian fraction (%)	Second Gaussian σ_L (mm)
$D_s^- \longrightarrow \phi\pi^-$	276 ± 6	226 ± 6	66	1.1 ± 0.04
$D_s^- \longrightarrow K^{*0}K^-$	279 ± 13	219 ± 13	73	0.91 ± 0.06
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	333 ± 18	283 ± 17	70	2.4 ± 2.0

Table 6.6: B_s^0 *decay length resolution with both a single and a two Gaussian fit for the 1991-1993 data set.*

from the reconstruction of the primary interaction point.

For the 1991-1993 data set a different approach had to be used. Because of a lack of z-information from the vertex detector, the same procedure was used to obtain the B_s^0 decay length $L_{R\Phi}$ in the $R\Phi$ plane only. The true decay length was then estimated from

$$L = L_{R\Phi}/\sin(\theta) \tag{6.8}$$

where θ is the polar angle of the track. Compared to using the (barely available) z-information from the vertex fit, this technique provides an improvement of the core resolution for the hadronic modes from 350 μm to 226 μm. Table 6.6 gives the B_s^0 decay length resolutions for the different modes for the 1991-1993 data set. While the core width is not significantly broadened compared to the 1994-1995 data set, the second Gaussian becomes more important here. For example, for the $D_s^- \longrightarrow \phi\pi^-$ mode the fraction increases from 14% to 34%.

6.3.2 Improvement of the Decay Length Resolution

One possible way to improve the decay length resolution follows directly from the topology of the decay chain shown in figure 5.1. When reconstructing the secondary vertex between the lepton and the D_s^- pseudotrack, any error on the vertex point is translated directly into an error on the B_s^0 decay length. The error induced on the decay length by the error on the vertex point is proportional to $\cos\alpha$, where α is the angle between the lepton and the D_s^-

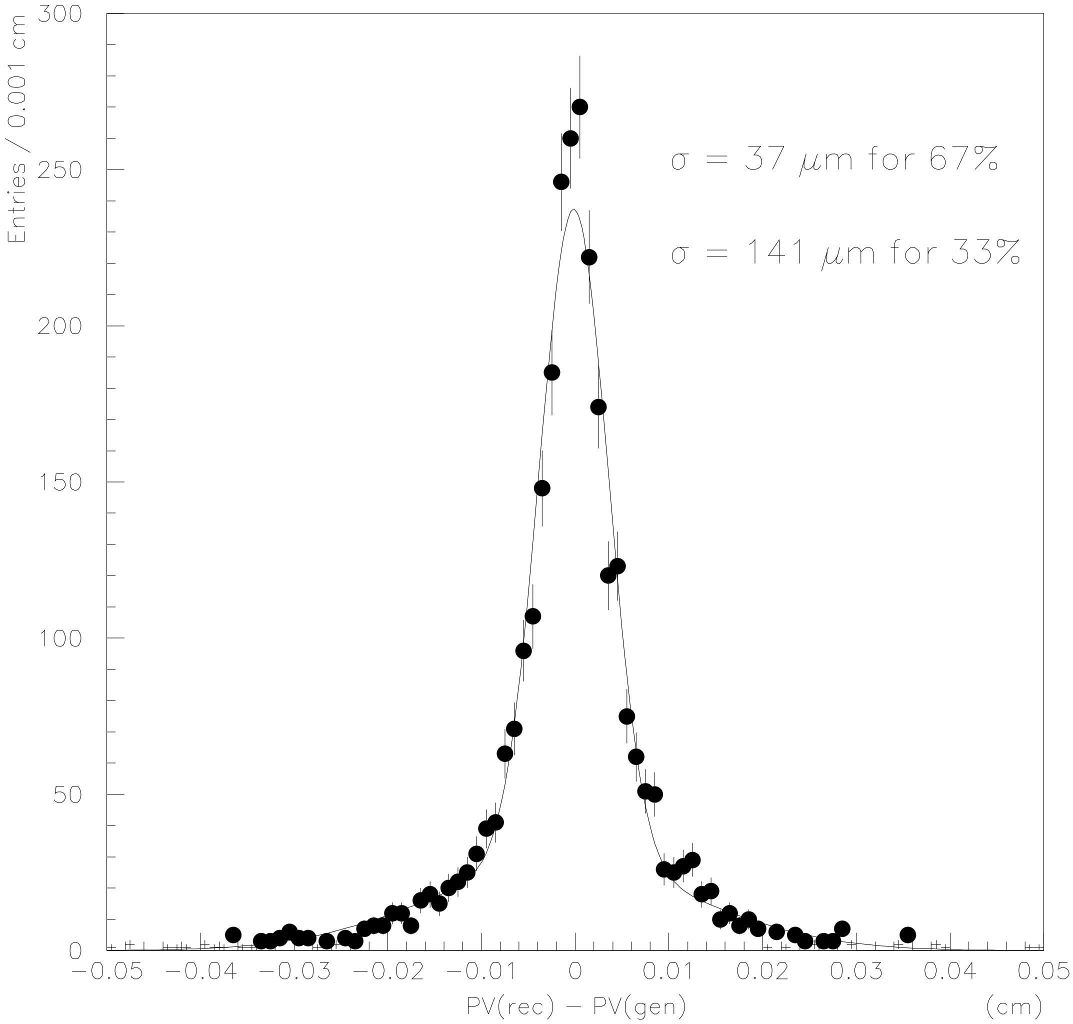


Figure 6.16: *Primary vertex resolution (projected onto the $D_s^- l^+$ direction). The solid line the fit described in the text.*

pseudotrack. Thus for large angle crossings the error on the decay length is smaller than for small angle crossings.

Figure 6.17 shows the correlation between the two. The best possible decay length resolution is around 125 μm. Unfortunately, as figure 6.18 shows, the mean angle between the two tracks is rather small. Thus requiring a large angle reduces the reconstruction efficiency, as can be seen in figure 6.19. Here, 100% is defined as the total after all the other selection criteria, but before any selection on the opening angle. This means that for a decay length resolution of 125 μm the efficiency in the $D_s^- \longrightarrow \phi\pi^-$ mode is approximately 6% of 12.0%, i.e. 0.7%.

Even if we were willing to pay this price, this will not improve the total proper time resolution by as much as one could expect. In figure 6.20 we see the generated momentum for all B_s^0 and for the B_s^0 satisfying $\cos\alpha < 0.95$. The mean momenta are 32 GeV/c and 26 GeV/c, respectively (the decrease is due to a lower lepton momentum). From figure 6.10 we know that the relative momentum resolution increases with decreasing momentum. In other words, what is potentially gained on the decay length resolution is lost on the relative momentum resolution. For this reason no selection on the angle between the lepton and the D_s^- was applied.

It should be noted, however, that these two factors enter the proper time resolution in two different places (see formula 6.2), and that it is possible to improve the time independent term in the proper time resolution this way.

6.4 Proper Time Resolution

6.4.1 Time Dependent Proper Time Resolution

One way to obtain the proper time resolution is to use the standard error propagation of formula 6.2, but as it was mentioned, it is only an approximate formula. For this reason another method was chosen. The difference between the reconstructed and generated proper time

$$t_{\text{rec}} - t_{\text{gen}} \tag{6.9}$$

was fitted using the MINUIT [69] package with two Gaussian distributions using an unbinned log[1] likelihood fit

$$G(t_m, t_{\text{gen}}) = \frac{f_1}{\sqrt{2\pi}\sigma_1(t_{\text{gen}})} e^{-\frac{(t_m - t_{\text{gen}})^2}{2\sigma_1^2(t_{\text{gen}})}} + \frac{1 - f_1}{\sqrt{2\pi}\sigma_2(t_{\text{gen}})} e^{-\frac{(t_m - t_{\text{gen}})^2}{2\sigma_2^2(t_{\text{gen}})}} \tag{6.10}$$

where f_1 is the fraction of the first Gaussian. Both Gaussians have time dependent standard deviations

$$\sigma_i^2(t_{\text{gen}}) = a_i^2 + (b_i \times t_{\text{gen}})^2 = a_i \otimes (b_i\, t_{\text{gen}}) \tag{6.11}$$

with a_i, b_i constants. These distributions are not strictly speaking Gaussians because of the time dependent resolutions. For simplicity, however, they are referred to as Gaussian in this thesis.

[1]'log' means here the natural logarithm.

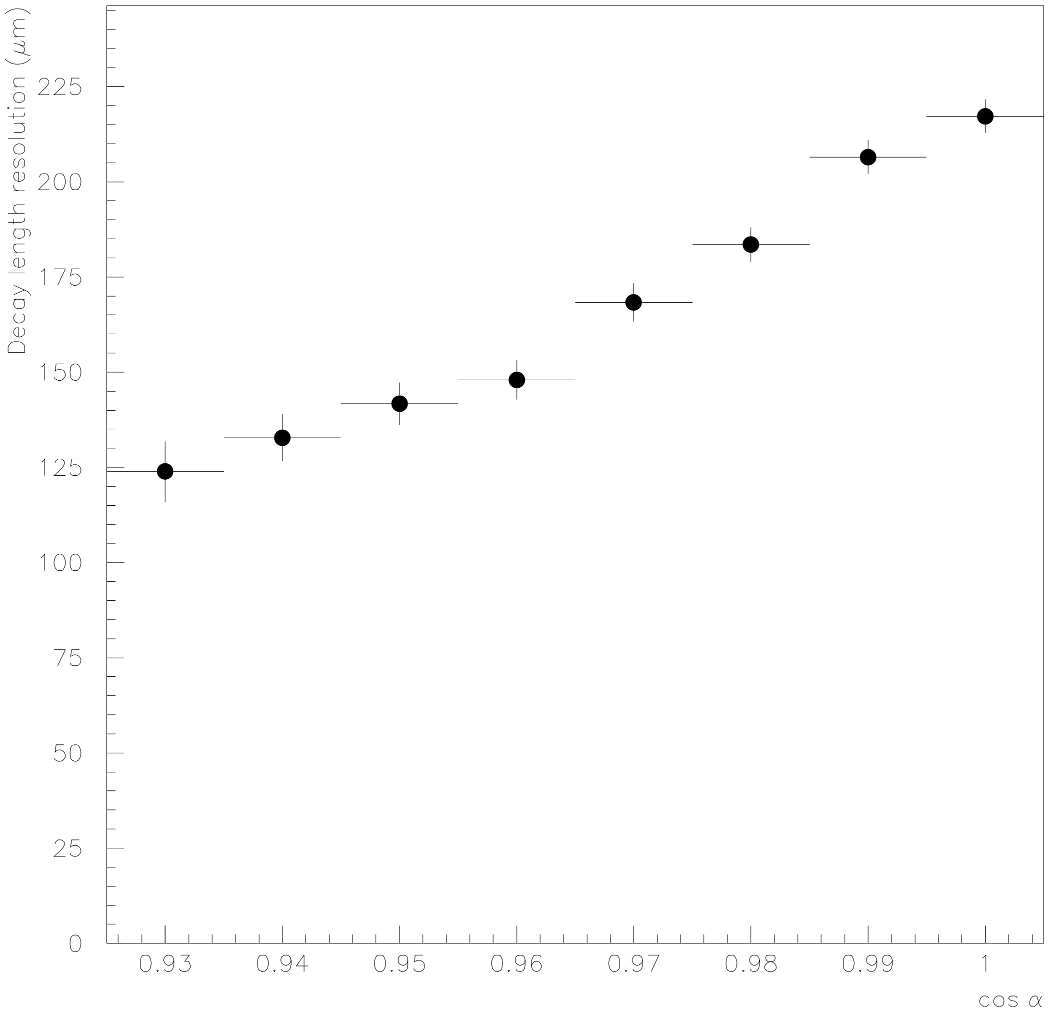


Figure 6.17: B_s^0 *decay length resolution versus the angle* α *between the lepton and the* D_s^-.

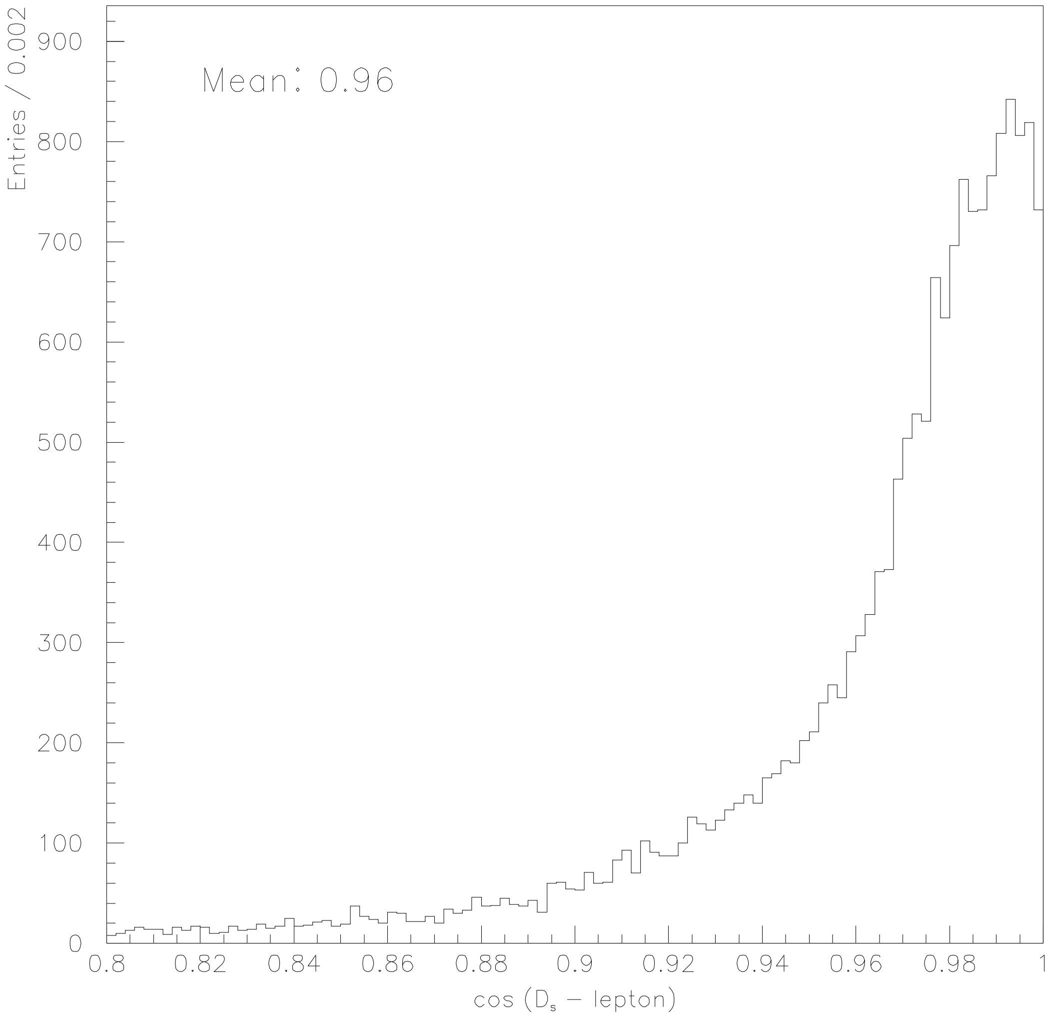


Figure 6.18: *Generated angle between the lepton and the D_s^-.*

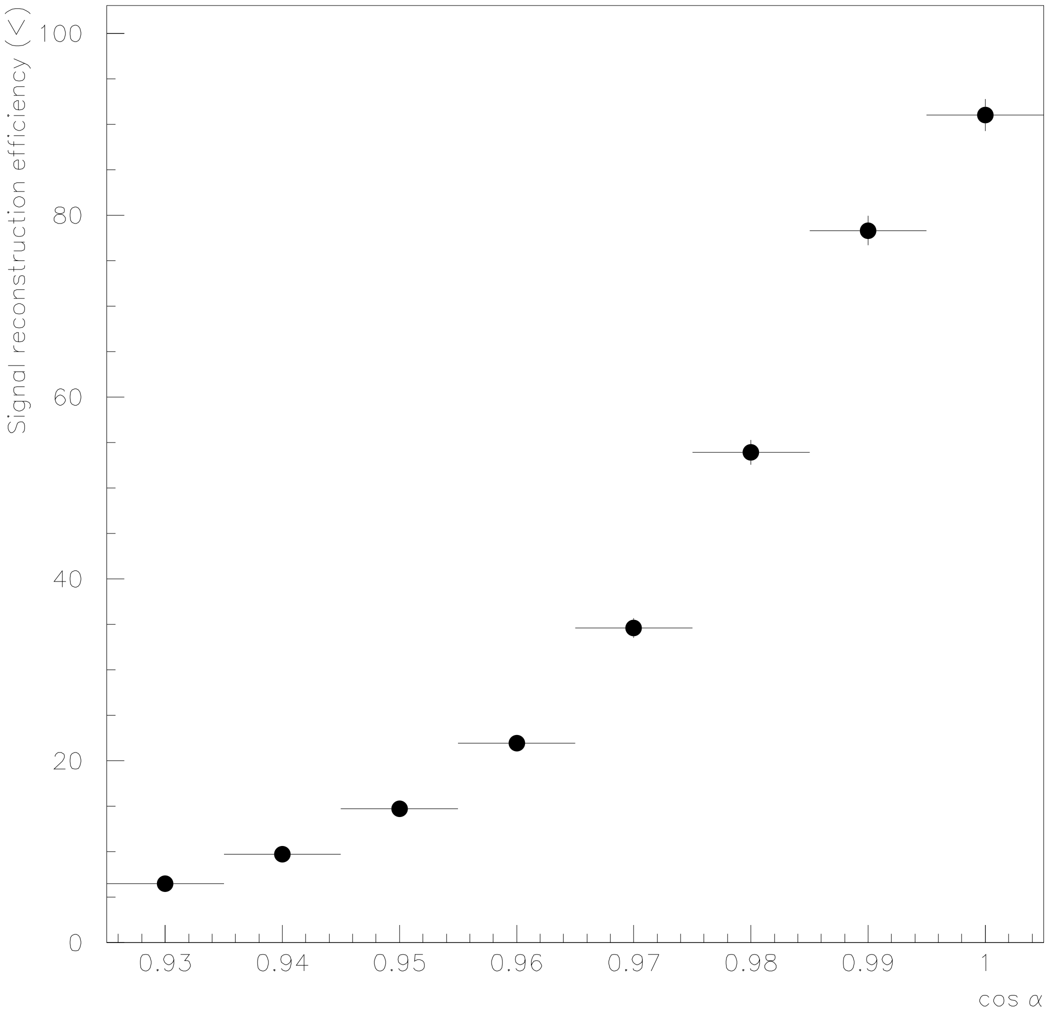


Figure 6.19: B_s^0 *reconstruction efficiency versus the angle* α *between the lepton and the* D_s^-*. Here, 100% is defined as the total after all the other selection criteria, but before any selection on the opening angle.*

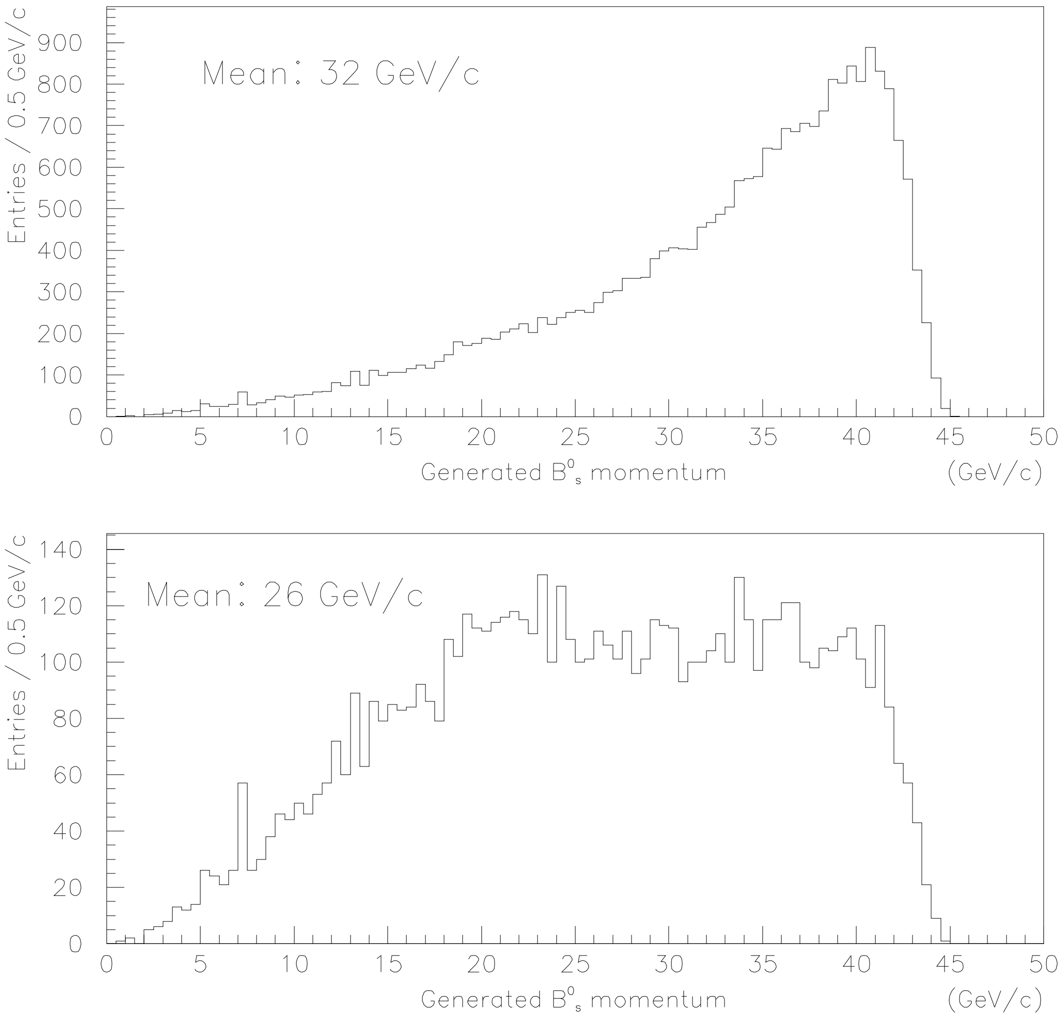


Figure 6.20: *Generated momentum for all B_s^0 (upper) and for the B_s^0 satisfying* cos $(\theta_{l+\,-\,D_s^-}) < 0.95$ (lower).

D_s^- decay mode	First Gaussian σ_t (ps)	First Gaussian fraction (%)	Second Gaussian σ_t (ps)
$D_s^- \longrightarrow \phi\pi^-$	0.10 ⊗ 0.07t	82	0.35 ⊗ 0.15t
$D_s^- \longrightarrow K^{*0}K^-$	0.11 ⊗ 0.07t	88	0.44 ⊗ 0.13t
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	0.11 ⊗ 0.09t	77	0.41 ⊗ 0.15t
$D_s^- \longrightarrow K^{*0}K^{*-}$	0.12 ⊗ 0.07t	74	0.40 ⊗ 0.15t
$D_s^- \longrightarrow \phi\pi^-\pi^0$	0.11 ⊗ 0.08t	82	0.48 ⊗ 0.24t
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	0.12 ⊗ 0.07t	79	0.37 ⊗ 0.16t

Table 6.7: B_s^0 *proper time resolution for the 1994-1995 data set.*

D_s^- decay mode	First Gaussian σ_t (ps)	First Gaussian fraction (%)	Second Gaussian σ_t (ps)
$D_s^- \longrightarrow \phi\pi^-$	0.11 ⊗ 0.06t	71	0.51 ⊗ 0.09t
$D_s^- \longrightarrow K^{*0}K^-$	0.11 ⊗ 0.07t	74	0.49 ⊗ 0.07t
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	0.15 ⊗ 0.08t	81	0.62

Table 6.8: B_s^0 *proper time resolution for the 1991-1993 data set.*

Tables 6.7 and 6.8 show the proper time resolutions obtained from the fit. Using equation (6.2), the coefficient b should be approximately equal to the relative momentum resolution, as observed. For a mean momentum of 32 GeV/c and a decay length resolution of 220 μm, the coefficient a should be around 0.1, as observed. Due to the increased z resolution for the 1991-1993 data set, the contribution of the second Gaussian is increased compared to the 1994-1995 data set. Because of the even worse resolution in the semileptonic modes, a constant resolution was used in the second Gaussian for these modes in this data set.

Figure 6.22 shows the difference $t_{\rm rec} - t_{\rm gen}$ in different bins of $t_{\rm gen}$ with the fit described above. This plot of the fit is just an approximation, however, since we try to plot a two variable function with $t_{\rm gen}$ and $t_{\rm rec} - t_{\rm gen}$ as variables. In the plot, $t_{\rm gen}$ (which is used in $\sigma(t_{\rm gen})$) is kept constant in each bin, while $t_{\rm rec} - t_{\rm gen}$ is used as the free variable. For the last bin this approximation is no longer really valid. The proper time resolution is well described by the fit.

In the previous figure we can see that the Root Mean Square RMS of the proper time resolution increases with increasing proper time. This can been seen directly in figure 6.21 which shows the proper time resolution versus the proper time. The proper time resolution increases with increasing proper time.

6.4.2 Time Integrated Proper Time Resolution

Fitting the difference between the reconstructed and generated proper time with two Gaussian distributions, both with time-independent resolutions, yields a time-integrated proper time resolution for the 1994-1995 data set with a core width of (see figure 6.23)

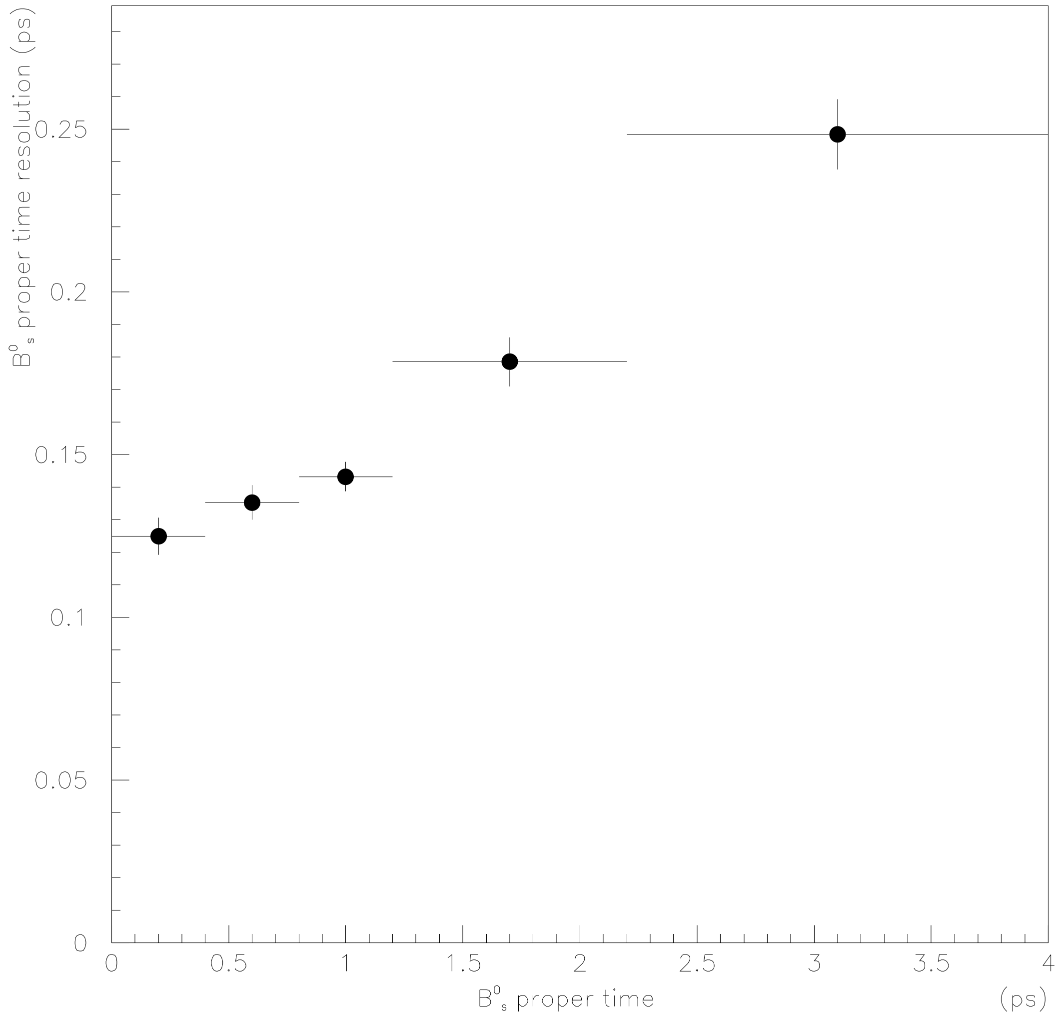


Figure 6.21: *B_s^0 proper time resolution versus the proper time.*

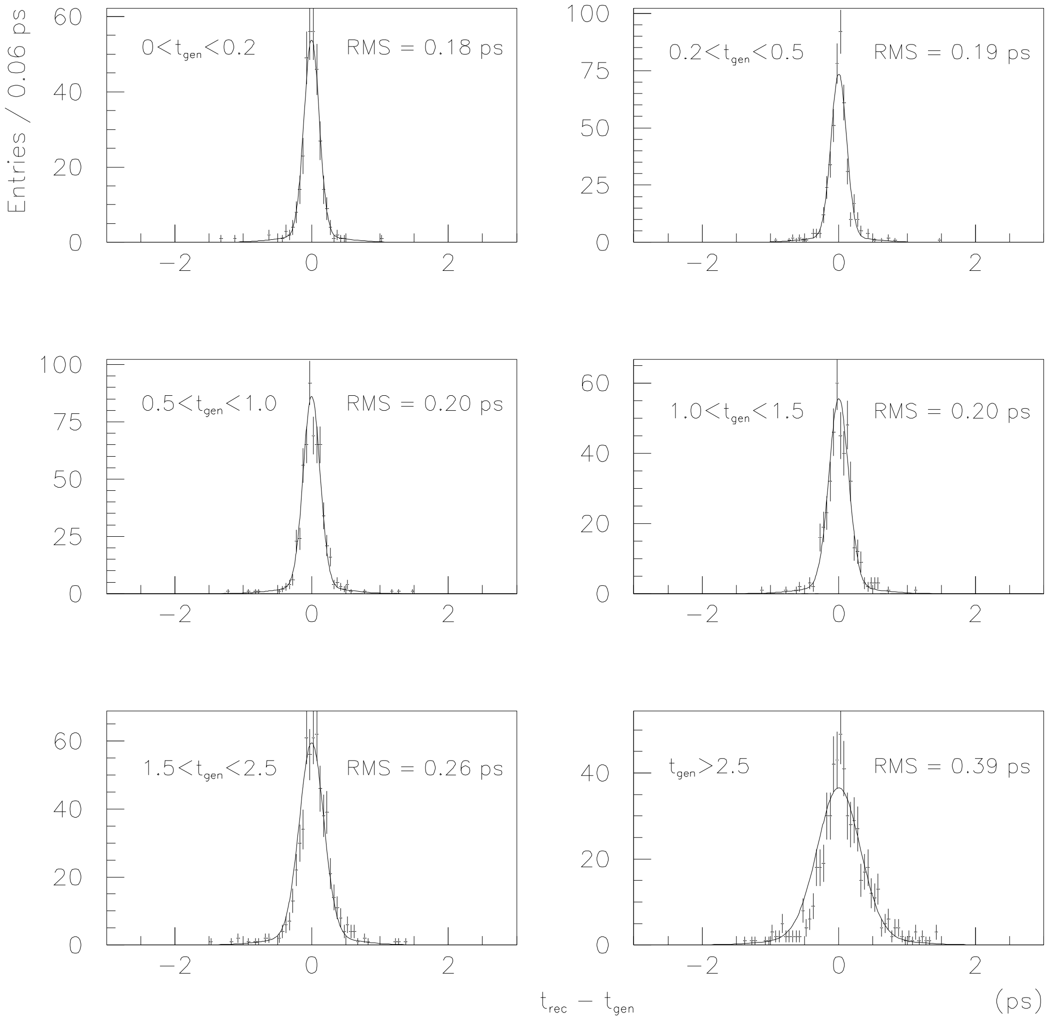


Figure 6.22: *B_s^0 proper time resolution in different bins of $t_{generated}$ (RMS = Root Mean Square). The solid lines are the fit described in the text.*

D_s^- decay mode	Single Gaussian σ_t (ps)	First Gaussian σ_t (ps)	First Gaussian fraction (%)	Second Gaussian σ_t (ps)
$D_s^- \longrightarrow \phi\pi^-$	0.169 ± 0.003	0.142 ± 0.003	78	0.46 ± 0.02
$D_s^- \longrightarrow K^{*0}K^-$	0.177 ± 0.006	0.115 ± 0.006	83	0.54 ± 0.05
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	0.21 ± 0.01	0.15 ± 0.01	68	0.49 ± 0.05
$D_s^- \longrightarrow K^{*0}K^{*-}$	0.205 ± 0.009	0.161 ± 0.009	72	0.54 ± 0.04
$D_s^- \longrightarrow \phi\pi^-\pi^0$	0.18 ± 0.01	0.14 ± 0.01	74	0.54 ± 0.07
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	0.198 ± 0.003	0.154 ± 0.003	74	0.50 ± 0.01

Table 6.9: B_s^0 *time integrated proper time resolution for the 1994-1995 data set.*

D_s^- decay mode	Single Gaussian σ_t (ps)	First Gaussian σ_t (ps)	First Gaussian fraction (%)	Second Gaussian σ_t (ps)
$D_s^- \longrightarrow \phi\pi^-$	0.205 ± 0.005	0.147 ± 0.005	68	0.56 ± 0.02
$D_s^- \longrightarrow K^{*0}K^-$	0.20 ± 0.01	0.15 ± 0.01	70	0.52 ± 0.04
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	0.23 ± 0.01	0.18 ± 0.02	72	0.63 ± 0.08

Table 6.10: B_s^0 *time-integrated proper time resolution for the 1991-1993 data set.*

Selection criteria	First Gaussian σ_t (ps)	First Gaussian σ_t (%)	Second Gaussian σ_t (ps)
Usual	0.142 ± 0.003	78	0.46 ± 0.02
cos $(\theta_{D_s^- - l^+}) < 0.95$	0.13 ± 0.01	68	0.43 ± 0.03

Table 6.11: B_s^0 *proper time resolution with and without a selection on the angle between the* D_s^- *and the lepton.*

$$\sigma_t = 0.14 \text{ ps}$$

For the 1991-1993 data set the corresponding proper time resolution is

$$\sigma_t = 0.15 \text{ ps}$$

but the fraction of the first Gaussian is 68% versus 78% for the 1994-1995 data set.

Tables 6.9 and 6.10 show the proper time resolutions for the individual modes for the two data sets.

In table 6.11 the proper time resolutions are listed for the whole sample and for the candidates satisfying the criteria cos $(\theta_{l^+ - D_s^-}) < 0.95$ (see also figure 6.24). The gain is smaller than expected from the improved decay length resolution, reflecting the higher relative momentum resolution.

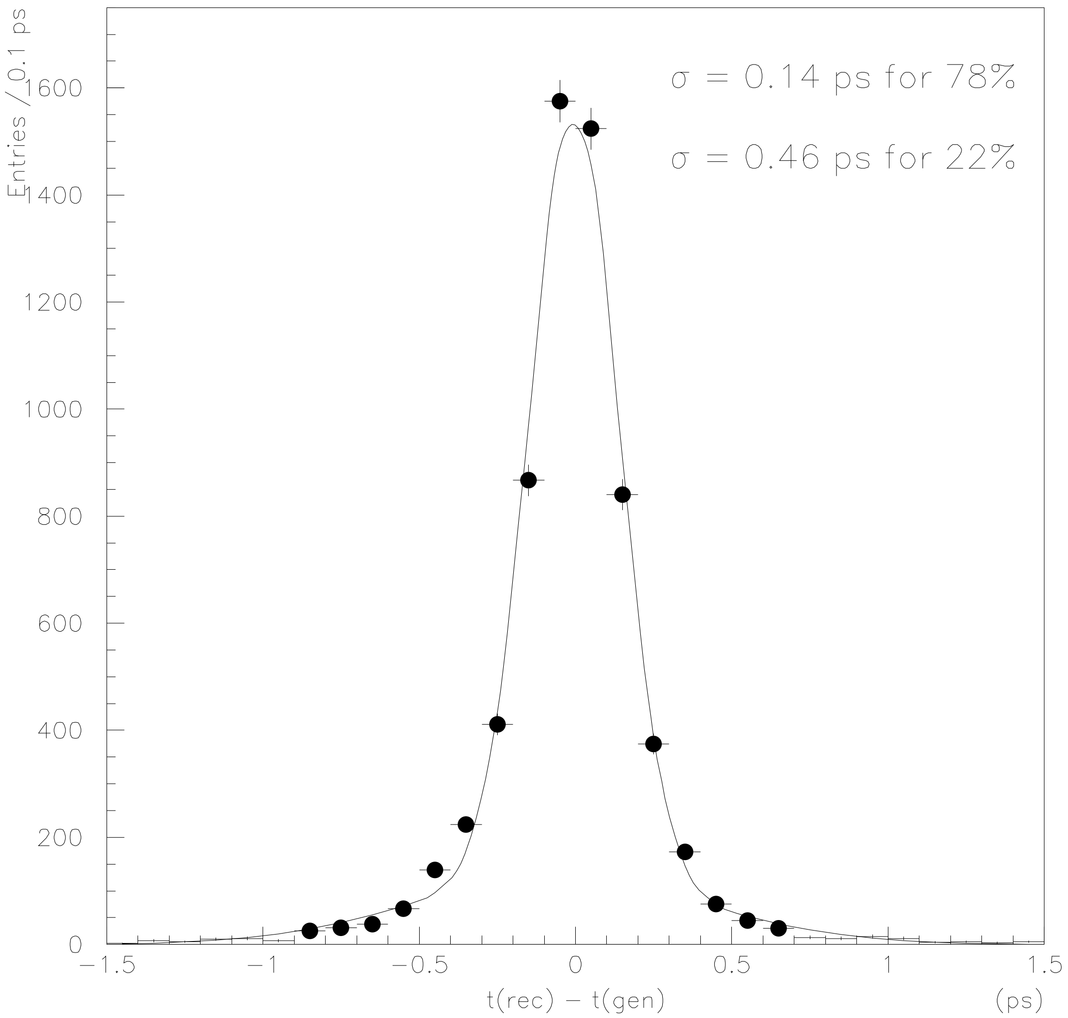


Figure 6.23: B_s^0 *proper time resolution. The solid line is the fit described in the text.*

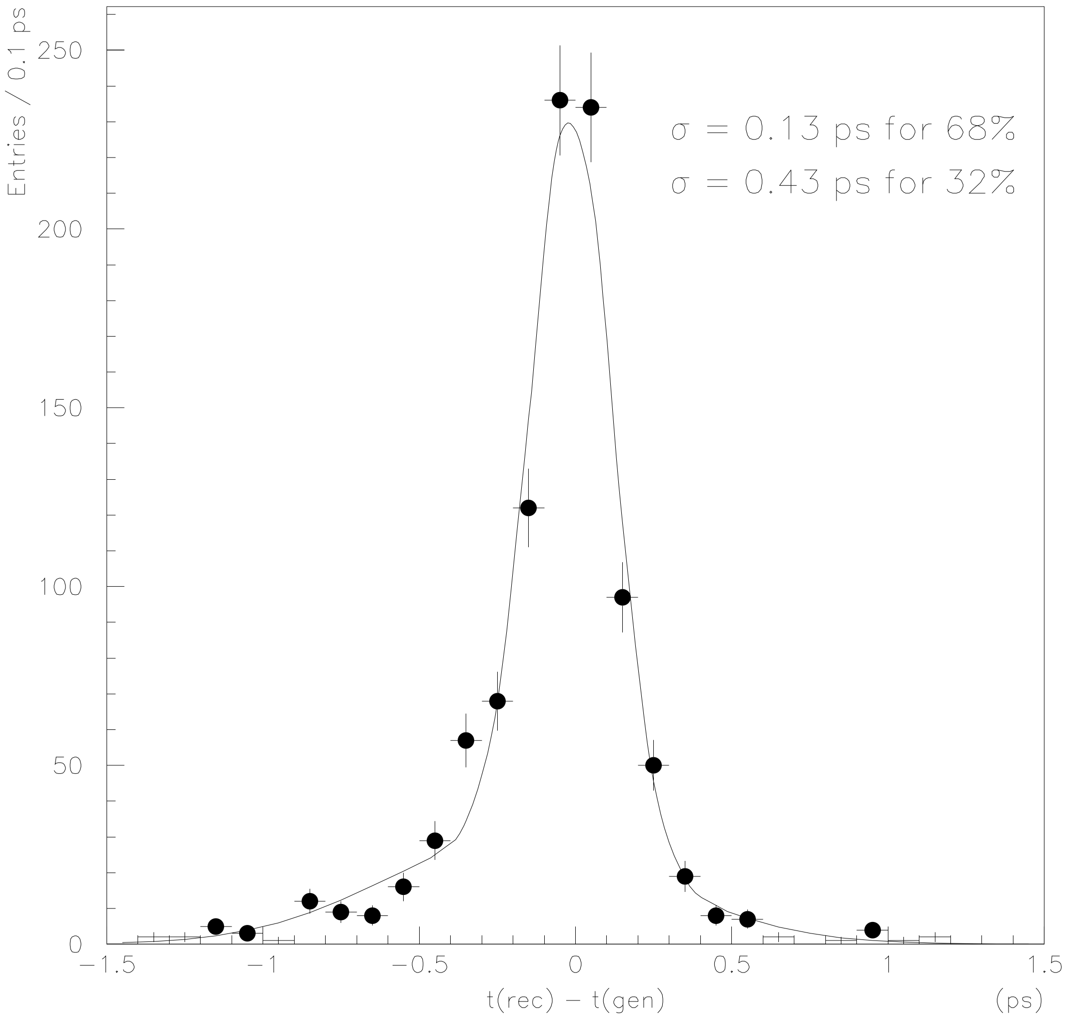


Figure 6.24: B_s^0 *proper time resolution for candidates satisfying the criteria cos* $(\theta_{l+ - D_s^-}) < 0.95$. *The solid line is the fit described in the text.*

Resolution type	First Gaussian mean	First Gaussian fraction (%)	Second Gaussian mean
Momentum	-0.004 $\pm$ 0.001	78	0.093 $\pm$ 0.006
Decay length	5 $\pm$ 3 μm	86	-31 $\pm$ 31 μm
Proper time	-0.004 $\pm$ 0.002 ps	78	-0.06 $\pm$ 0.01 ps

Table 6.12: *Reconstruction biases.*

6.4.3 Proper Time Bias

The reconstructed proper time is obtained from the reconstructed momentum and decay length. This means that any bias we have in the reconstruction of these quantities will affect the obtained proper time.

Table 6.12 shows the mean values for the two Gaussian distributions used to fit the relative momentum resolution, the decay length resolution and the proper time resolution (see also figures 6.6, 6.15 and 6.23). As figure 6.6 shows, there is a tendency to overestimate the momentum, because of the additional photon (there is no bias in the reconstructed neutrino energy (see figure 6.4)). This implies that the proper time is underestimated, since $t \propto \frac{1}{p}$. The bias is small, however, only -0.06 ps for 22% of the events. The bias in the decay length reconstruction is negligible. This confirms the conclusion from section 5.7 that there is no need for an acceptance function.

Demanding candidates with $\cos(\theta_{l^+ - D_s^-}) < 0.95$ introduces an increased bias (see figure 6.24). Since these candidates have lower momentum, the effect of including a too energetic photon is more important.

Chapter 7

B_s^0 Lifetime

7.1 General Remarks

The B_s^0 lifetime is obtained from an unbinned maximum log likelihood fit. The method is the following: the true exponential proper time distribution, with the B_s^0 lifetime as a free parameter, is convoluted with a resolution function. The contributions from the combinatorial background, the *cascade* decays and the kinematic *reflection* background are taken into account. The obtained proper time distribution is compared to the proper time distribution obtained from data. The fit itself is performed with the MINUIT package [69]. All relevant parameters are either fixed to values obtained from Monte Carlo simulations or directly from data. Thus, the B_s^0 lifetime is the only free parameter in the fit.

As already mentioned, only B_s^0 candidates with a $D_s(\phi)$ mass within $\pm 2\sigma(\pm 1.5\Gamma)$ of the fitted mean $D_s(\phi)$ mass were used. In addition, the measured proper time of the candidates had to be in the time interval

$$-3.0\,ps\ < t_i < 10.0\,ps$$

A total of 392 candidates were selected. In six events (five in the 1994-1995 data set, one in the 1991-1993 data set) a candidate was found in the same hemisphere in the same event, but in two different modes. The overlaps were between the modes $\phi\pi$ and $\phi l\nu_l$ (four events), $\phi\pi$ and $K^{*0}K$ (one event) and $\phi\pi\pi\pi$ and $\phi l\nu_l$ (one event). Taking into account the narrow ϕ mass resonance compared to K^{*0}, and the additional constraint of the D_s mass, the candidates in the modes $\phi\pi$ and $\phi\pi\pi\pi$ were kept.

Of the 392 candidates, 197 B_s^0 signal candidates were obtained. Table 7.1 gives the sample composition mode by mode.

7.2 The Likelihood Function

The likelihood function for the B_s^0 lifetime is

$$\mathcal{L} = \prod_{i=1}^{N(\text{right}-\text{sign})} \mathcal{L}_{\text{i}} \tag{7.1}$$

where

D_s decay mode	Number of candidates 1994-1995	Number of B_s^0 signal events 1994-1995	Number of candidates 1991-1993	Number of B_s^0 signal events 1991-1993
$D_s^- \longrightarrow \phi\pi^-$	55	33	46	25
$D_s^- \longrightarrow K^{*0}K^-$	56	26	50	20
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	20	10		
$D_s^- \longrightarrow K^{*0}K^{*-}$	26	17		
$D_s^- \longrightarrow \phi\pi^-\pi^0$	26	14		
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	69	38	44	14
All modes:	252	138	140	59

Table 7.1: *Sample composition: Total number of events and the number of B_s^0 signal candidates in the sample for the individual modes and the two data sets.*

$$\mathcal{L}_{\mathrm{i}} = (1 - f_{\mathrm{comb}}) \times (f_{B_s^0} P_{B_s^0}(t_i) + f_{DD_s} P_{DD_s}(t_i) + f_{reflec} P_{reflec}(t_i)) + f_{\mathrm{comb}} P_{\mathrm{comb}}(t_i) \quad (7.2)$$

Here (the values of these parameters and additional information can be found in/from section 5.8):

- f_{comb} is the fraction of combinatorial background events. To improve the sensitivity we use, instead of the average fractions found from the mass plots, the combinatorial background fractions on an *event-by-event* basis. These were determined by the measured value of the $D_s(\phi)$ mass of the candidate. It was assumed that the further the candidate's mass was from the nominal one, the more likely it was that it was a combinatorial background event. The combinatorial background fraction was then given by

$$f_{\mathrm{comb}} = \frac{B(m_i)}{S(m_i) + B(m_i)} \quad (7.3)$$

  where $B(m_i)$ and $S(m_i)$ are the contributions from the background and signal fit, respectively, and m_i is the $D_s(\phi)$ mass of the candidate. This procedure reduces the statistical uncertainty on the fitted lifetime with 7%, and does not change the fitted value significantly.

- $f_{B_s^0}$ is the fraction of B_s^0 signal events in the $D_s(\phi)$ peak.
- f_{DD_s} is the fraction of *cascade* events in the $D_s(\phi)$ peak.
- f_{reflec} is the fraction of kinematic *reflection* background in the $D_s(\phi)$ peak.
- t_i is the measured proper time for candidate i.

The signal and background fractions f_i are found separately for each decay mode and the two data sets.

The probability densities are of the form

$$
\begin{aligned}
P &= G(t_i,t) \otimes exp(t,\tau_{B_s^0}) \\
&= \frac{1}{C_{\text{Normalization}}} \int_{t=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_t}\, e^{-\frac{(t_i-t)^2}{2\sigma_t^2}}\, \frac{1}{\tau}\, e^{\frac{-t}{\tau}}\, dt
\end{aligned}
\tag{7.4}
$$

where the parameter $C_{\text{Normalization}}$, given by

$$
C_{Normalization} = \int_{t_i=-3}^{t_i=10} \int_{t=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_t}\, e^{-\frac{(t_i-t)^2}{2\sigma_t^2}}\, \frac{1}{\tau} e^{\frac{-t}{\tau}}\, dt_i\, dt \tag{7.5}
$$

is needed to normalize the likelihood function properly.

The exact forms of the different probability densities are the following:

- It was shown in the previous chapter that the proper time resolution for the B_s^0 signal is well described by two Gaussian distributions with time dependent resolutions (separately for each decay mode and the two data sets). The signal probability is thus the convolution of an exponential of slope corresponding to the B_s^0 lifetime with the two Gaussian distributions.

  In section 5.7 it was verified that the selection criteria do not introduce any bias with respect to the B_s^0 proper time. Therefore, there is no need to introduce an *acceptance* function in the probability density function.

- The proper time distribution of the combinatorial background was parameterized for each decay mode and each data set separately (see also section 7.4). A Gaussian distribution G_{t_i} with a time independent resolution was used to describe the non-flying component. An exponential was used to describe any additional negative tails (poorly reconstructed events, $t_i < 0$). A second exponential was used to describe the flying background ($t_i > 0$) (see tables 7.2 and 7.3).

  $$
  P_{\text{comb}}(t_i) = f_1 G_{t_i} \;+\; f_2\, exp(+t_i,\tau^+) \;+\; (1-f_1-f_2)\, exp(-t_i,\tau^-) \tag{7.6}
  $$

  To describe the background under the $D_s(\phi)$ peak, data events from the $D_s(\phi)$ upper side-band, defined to be between 2.05(1.035) GeV/c^2 and 2.2(1.06) GeV/c^2, and the like-sign $D_s^{\pm} l^{\pm}$ events were used. Due to the possibility of a partial reconstruction of the D_s^- (for example $D_s^- \longrightarrow \phi\pi^-\pi^0$ with a missing π^0 reconstructed as $D_s^- \longrightarrow \phi\pi^-$) the lower side-band was not used.

- The kinematic *reflection* decays was parameterized as the signal, i.e. with the same resolution function, but with the B_d^0 lifetime.

- The *cascade* "DD_s" background proper time distribution was parameterized separately depending on whether it originated from a $\overline{B_d^0}$ or a B^-. The probability density was

D_s^- decay mode	Fraction Gaussian f_1	Fraction positive exponential f_2	Lifetime flying background τ^-
$D_s^- \longrightarrow \phi\pi^-$	0.30 ± 0.09	0.13 ± 0.04	0.99 ± 0.27
$D_s^- \longrightarrow K^{*0}K^-$	0.26 ± 0.08	0.14 ± 0.07	1.68 ± 0.25
$D_s^- \longrightarrow \phi l^- \bar{\nu}_l$	0.11 ± 0.06	0.22 ± 0.07	1.07 ± 0.19

Table 7.2: *Combinatorial background proper time parameterization for the 1991-1993 data set: Fractions and lifetime of the flying background.*

D_s^- decay mode	Fraction Gaussian f_1	Fraction positive exponential f_2	Lifetime flying background τ^-
$D_s^- \longrightarrow \phi\pi^-$	0.11 ± 0.04	0.20 ± 0.04	0.87 ± 0.12
$D_s^- \longrightarrow K^{*0}K^-$	0.20 ± 0.11	0.19 ± 0.04	0.68 ± 0.18
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	0.15 ± 0.07	0.19 ± 0.04	1.13 ± 0.20
$D_s^- \longrightarrow K^{*0}K^{*-}$	0.30 ± 0.32	0.11 ± 0.06	1.11 ± 0.22
$D_s^- \longrightarrow \phi\pi^-\pi^0$	0.11 ± 0.08	0.15 ± 0.06	1.37 ± 0.24
$D_s^- \longrightarrow \phi l^- \bar{\nu}_l$	0.37 ± 0.11	0.15 ± 0.08	1.38 ± 0.27

Table 7.3: *Combinatorial background proper time parameterization for the 1994-1995 data sets: Fractions and lifetime of the flying background.*

given by the convolution of an exponential and a Gaussian distribution with a constant resolution. The contributions were assumed to be mode independent for the hadronic modes. They were also assumed to be the same for the two data sets (see section 5.8.3).

For these decays, the momentum of the candidate is underestimated since the other tracks from the charmed meson which decays semileptonically are missing. This implies that an *effective* lifetime τ^{eff}, larger than the nominal B meson lifetime, is needed to describe the decay.

All parameters were obtained from Monte Carlo simulation studies. Tables 7.4 and 7.5 give the parameters for both B_d^0 and B^-. The parameters were assumed to the the same for the two data sets. For the semileptonic modes the B^- contribution was assumed to be the same as the one from the B_d^0. As expected, the fitted effective lifetimes are all higher than the generated ones.

7.3 Results

The fitted B_s^0 lifetime from the unbinned maximum likelihood for the entire data sample is

$$\tau(B_s^0) = 1.48\,^{+0.16}_{-0.14}\ \text{(stat) ps}$$

B meson type	τ^{gen} (ps)	τ^{eff} (ps)	σ (ps)
B_d^0	1.60	1.90 ± 0.06	0.35 ± 0.03
B^-	1.60	1.97 ± 0.09	0.19 ± 0.03

Table 7.4: *Effective lifetimes and fitted standard deviations for the resolution function for B_d^0 and B^-* cascade *decays - hadronic modes.*

B meson type	τ^{gen} (ps)	τ^{eff} (ps)	σ (ps)
B_d^0	1.60	1.97 ± 0.07	0.51 ± 0.03

Table 7.5: *Effective lifetimes and fitted standard deviations for the resolution function for B_d^0* cascade *decays - semileptonic modes.*

D_s^- decay mode	Fitted lifetime (ps)
$D_s^- \longrightarrow \phi\pi^-$	$1.35^{+0.26}_{-0.22}$
$D_s^- \longrightarrow K^{*0}K^-$	$1.51^{+0.39}_{-0.32}$
$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$	$1.91^{+1.33}_{-0.73}$
$D_s^- \longrightarrow K^{*0}K^{*-}$	$1.60^{+0.73}_{-0.52}$
$D_s^- \longrightarrow \phi\pi^-\pi^0$	$1.56^{+0.62}_{-0.41}$
$D_s^- \longrightarrow \phi l^-\bar{\nu}_l$	$1.49^{+0.31}_{-0.26}$
All modes 1991-1993	$1.53^{+0.32}_{-0.28}$
All modes 1994-1995	$1.46^{+0.18}_{-0.16}$
All modes 1991-1995	$1.48^{+0.16}_{-0.14}$

Table 7.6: *Separately fitted B_s^0 lifetime for each decay mode for the entire data sample.*

where the uncertainty is statistical only.

The decay length for the whole sample is shown in figure 7.1, while the proper time distribution (with the fit) for the data events in the signal region is shown in figure 7.2 for the whole sample and figure 7.3 for each decay mode separately.

The behaviour of the likelihood function for the whole sample around the minimum is shown in figure 7.4. It has a clear and well-behaved minimum.

Table 7.6 and figure 7.5 give the fitted lifetimes for each decay mode separately. They are all compatible within one standard deviation.

To check the reliability of the fitting method, a sample of semileptonic B_s^0-mesons was generated with a lifetime of 1.60 ps. The sample was passed through the same selection criteria as the data. The fitted lifetime was $\tau(B_s^0) = 1.576\,^{+0.024}_{-0.024}$ ps, i.e. it is compatible with the generated lifetime within the statistical uncertainty.

Source	Contribution (ps)
Fitting method	± 0.024
Signal proper time resolution function	± 0.002
Background proper time parameterization	± 0.04
B_s^0 purity and combinatorial background fractions	± 0.04
Total	± 0.06

Table 7.7: *Systematic uncertainties on the B_s^0 lifetime.*

7.4 Systematic Uncertainties

Systematic uncertainties arise from uncertainties on the different parameters used in the lifetime fit. They come from different sources:

- Fitting method:
  The fitting method itself was checked by using a sample of B_s^0 signal events with a generated lifetime of 1.60 ps. The fitted lifetime was: (1.576 ± 0.024) ps, i.e. the two values agree within the uncertainty. The statistical uncertainty was interpreted as a potential systematic uncertainty due to the fitting method.

- Signal proper time resolution function:
  The systematic uncertainty resulting from the proper time resolution functions was obtained by varying the fraction and the two parameters describing the time dependence of the narrow Gaussian within their statistical uncertainties. This was done for each mode and each data set.

- Background proper time parameterization:
  The background has been parameterized separately for each mode and for each data set, using real data from the side-band and like-sign combinations. The lifetime of the flying background was varied according to its uncertainty for each mode and each data set. In addition, the background was parameterized with one common parameterization for all the hadronic modes (but differently for the two data sets). This parameterization was then evaluated on both side-band and like-sign events, and the two of them separately.

- B_s^0 purity and combinatorial background fractions:
  The values of f_{comb}, f_{DD_s} and f_{refl}, with the corresponding (effective) lifetimes and standard deviations, were varied according to their uncertainties. In addition, the B_d^0 lifetime was varied within its uncertainty.

The main contributions come from the background proper time parameterization, the B_s^0 purity (including *cascade* decays and kinematic *reflections*) and combinatorial background fractions. The estimate of the first is rather conservative. As can be seen in figures 7.6 and 7.7, there are similarities between the hadronic modes within the same data set. One could, therefore, use a common background parameterization for these modes. The advantage is the increased statistics one obtains. By using individual parameterizations, as is done in this analysis, one should in principle get a more precise description of the background for each mode. Unfortunately, one suffers from reduced statistics in some modes, and, therefore, has large statistical uncertainties. These are reflected into the systematic uncertainties and give a much larger contribution than in the case with a common parameterization. One can argue for both methods, but when the choice was between the risk of underestimating the systematic uncertainties and being too conservative, the latter was preferred. It should be added that the fitted lifetime does not change significantly when one changes between the two methods. One should also bear in mind that the measurement of the B_s^0 lifetime is completely dominated by the statistical uncertainty.

The different contributions were assumed to be independent and added in quadrature to give the total systematic uncertainty. The final result for the lifetime then becomes

$$\tau(B_s^0) = 1.48\,^{+0.16}_{-0.14}\ (\mathrm{stat})\ \mathrm{ps}\ ^{+0.06}_{-0.06}\ (\mathrm{syst})\ \mathrm{ps}$$

The world average [1] B_s^0 lifetime is $\tau(B_s^0) = (1.54 \pm 0.07)$ ps. Within the uncertainties, the two values are compatible.

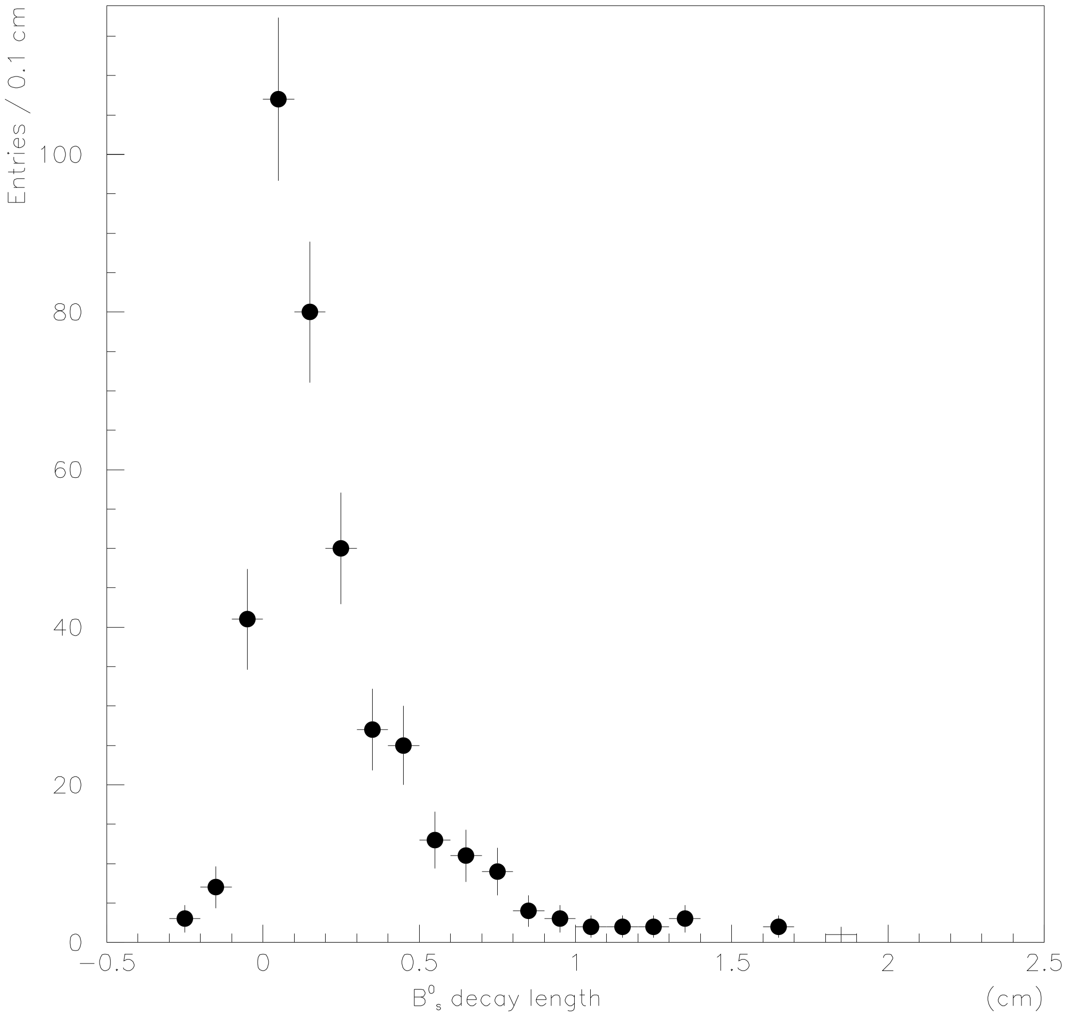


Figure 7.1: B_s^0 *decay length for the entire data sample.*

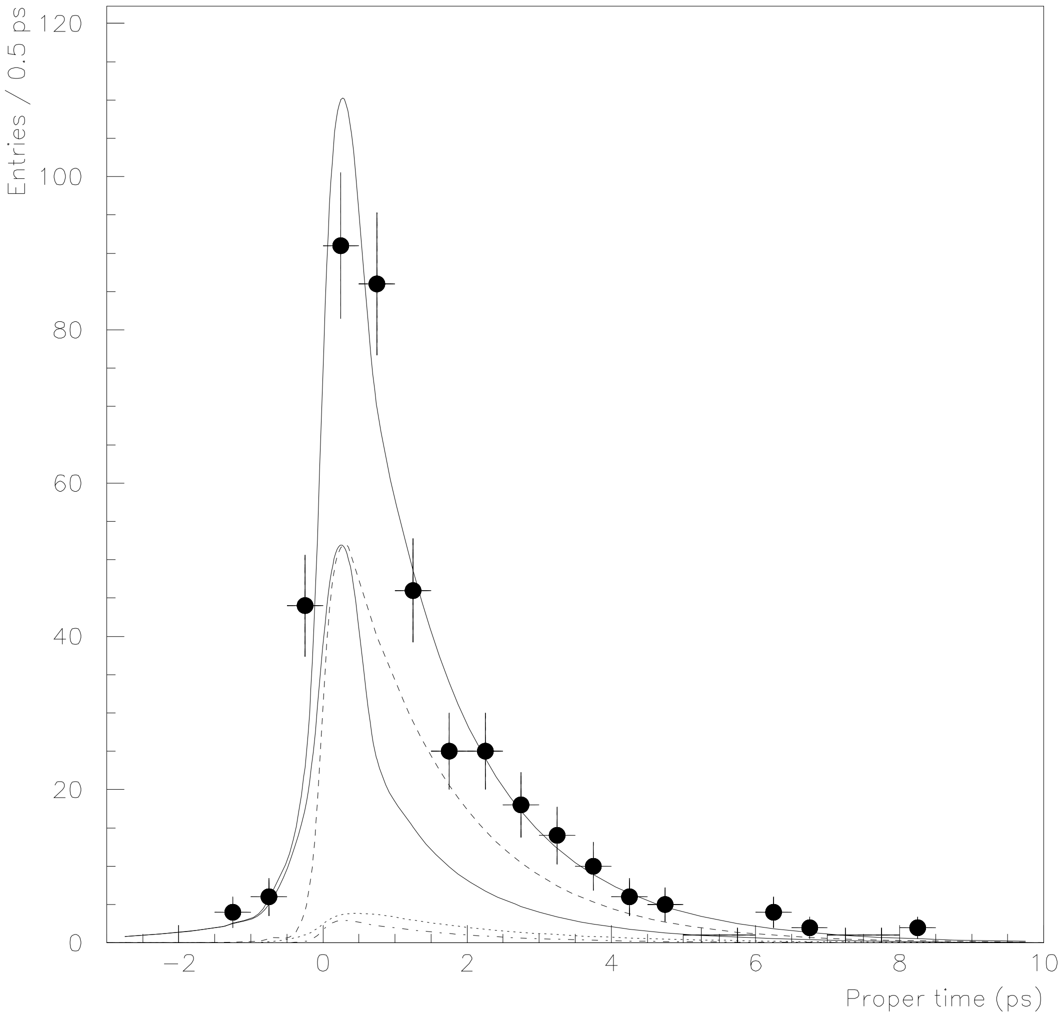


Figure 7.2: B_s^0 *proper time distribution for the entire data sample. The solid line through the data points is the likelihood fit described in the text. Also shown are the contributions from the* B_s^0 *signal (- - -), combinatorial background (continuous line), the* cascade *background (....) and the kinematic* reflection *background (_._._.).*

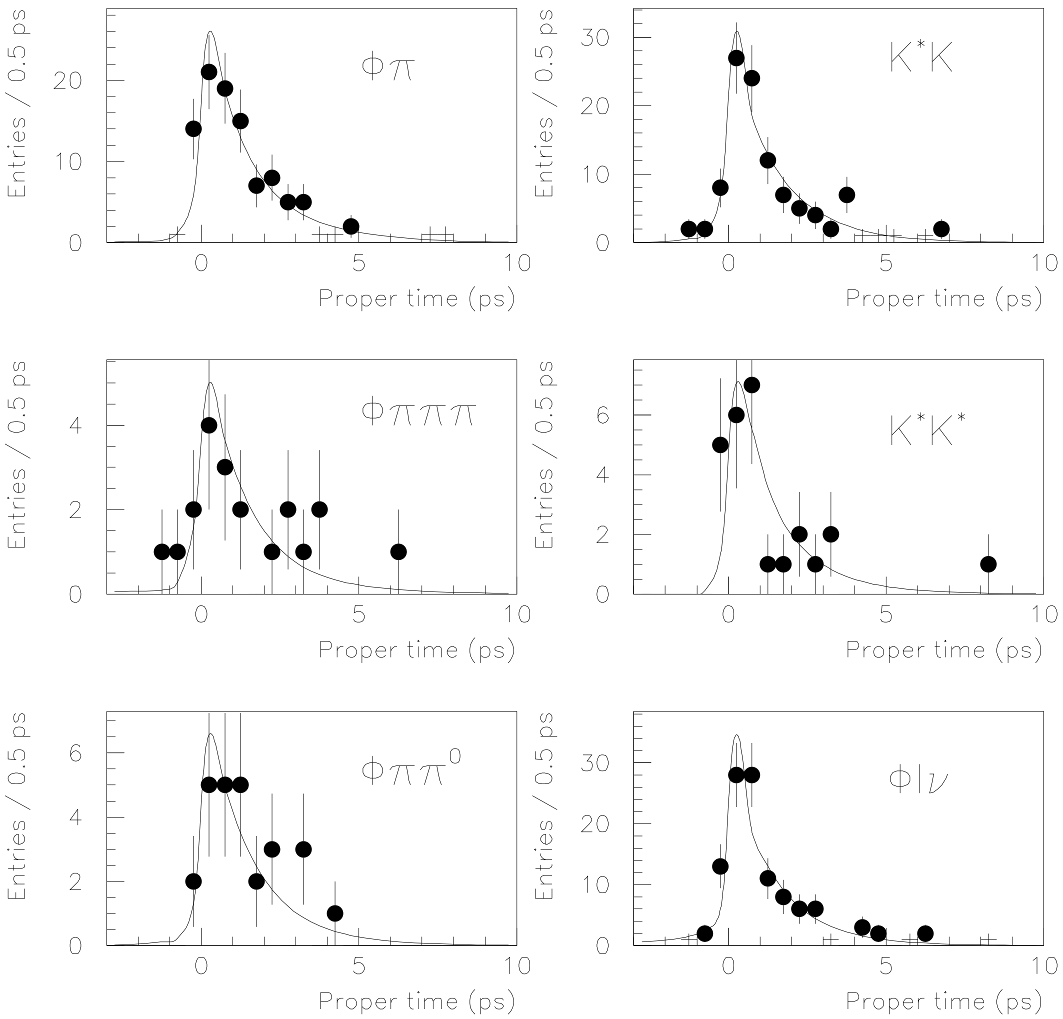


Figure 7.3: B_s^0 *proper time distribution for each individual decay mode for the entire data sample. The solid lines are the fit to the common lifetime described in the text.*

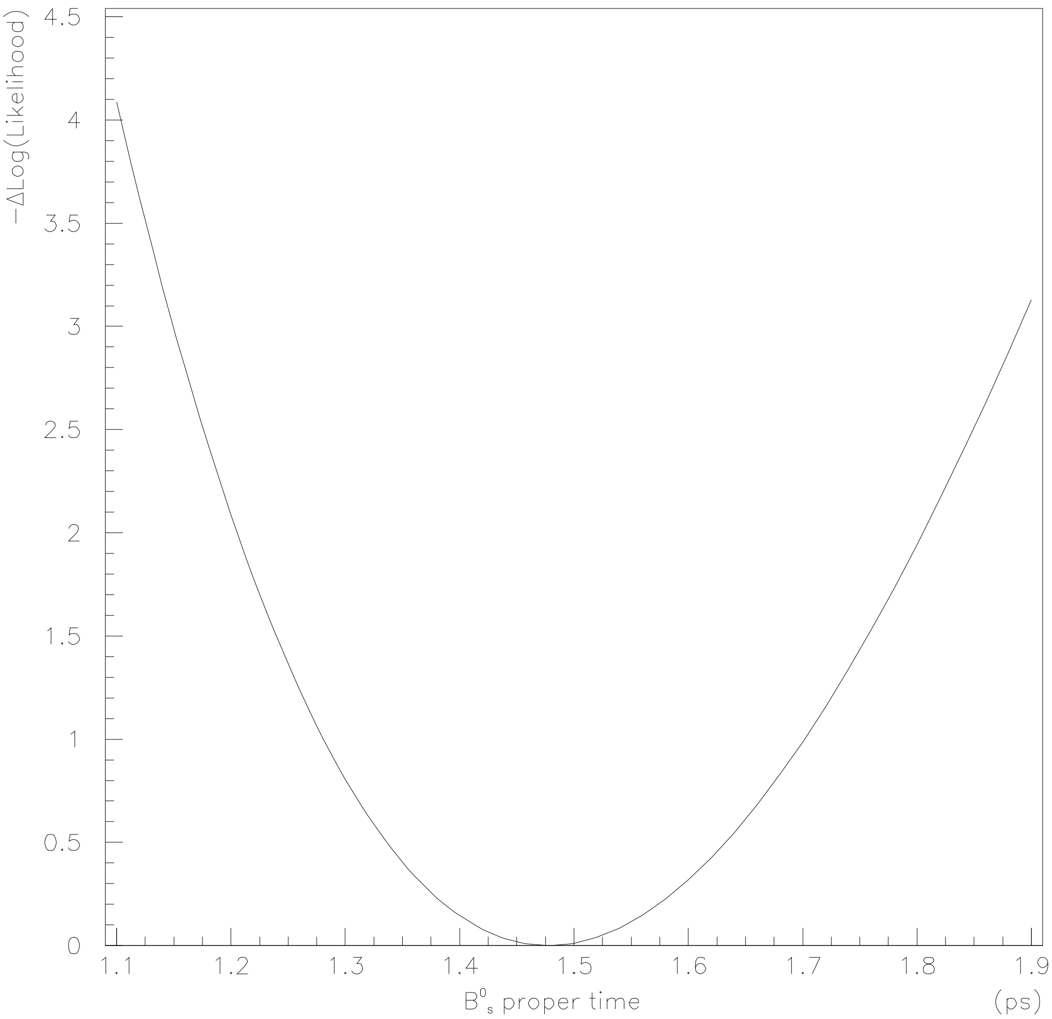


Figure 7.4: *Negative log-likelihood difference with respect to the minimum versus the B_s^0 lifetime.*

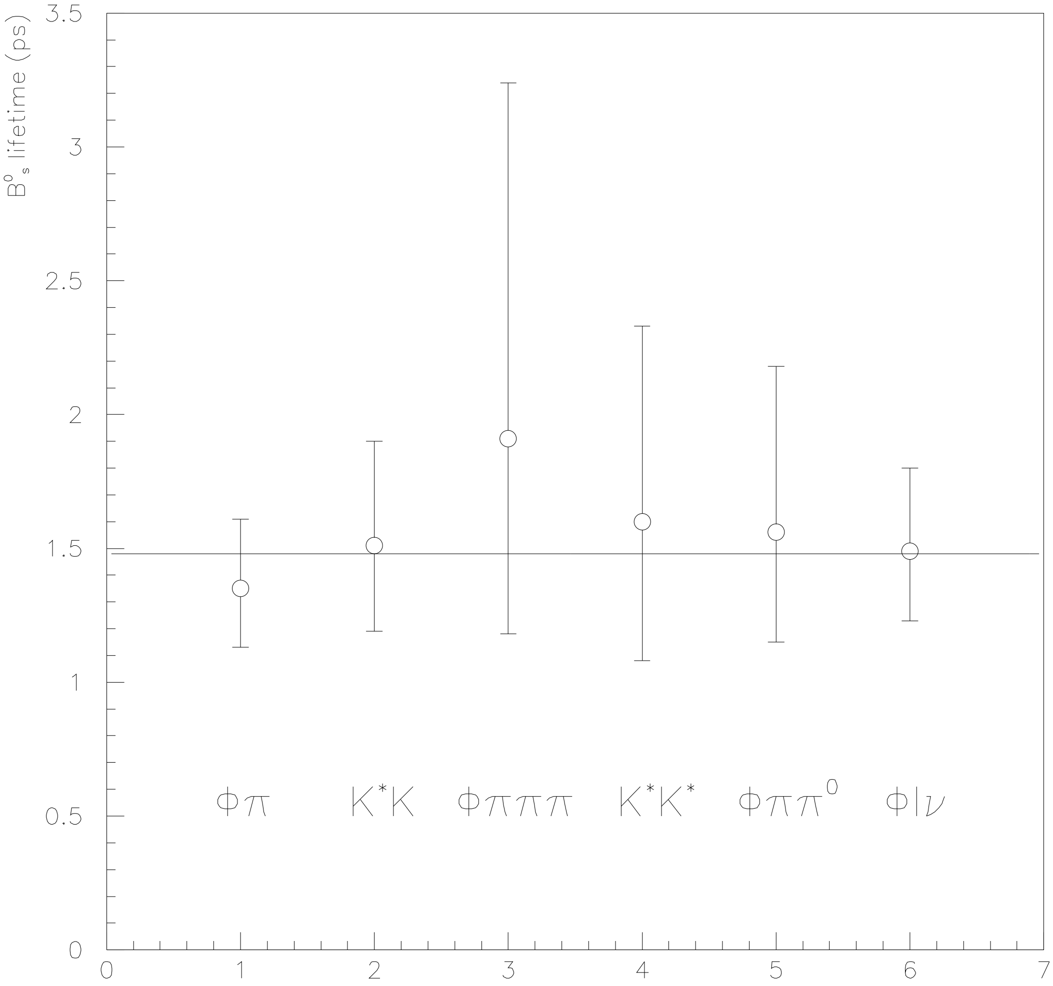


Figure 7.5: *Fitted B_s^0 lifetime for each decay mode for the entire data set.*

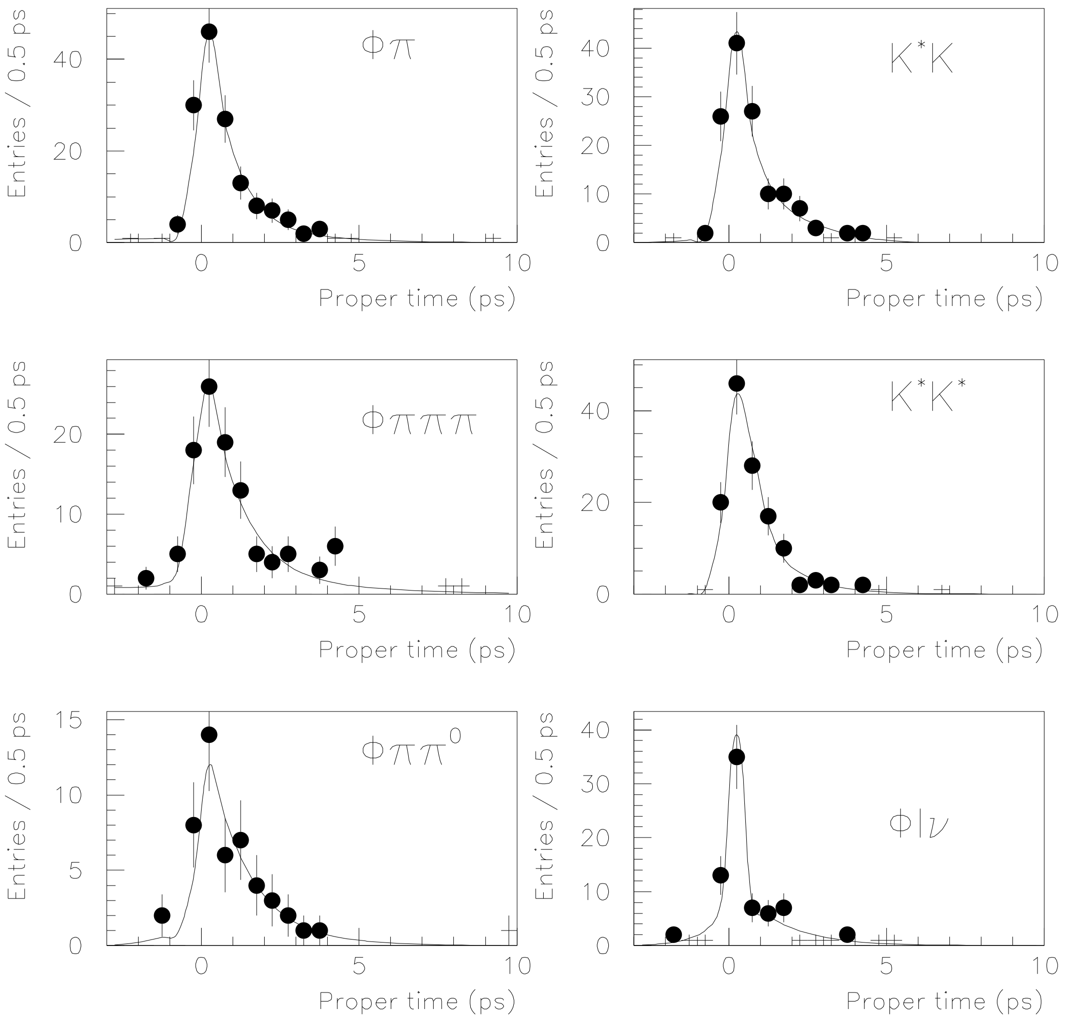


Figure 7.6: *Combinatorial background proper time distributions (side-band and like-sign) for each decay mode for the 1994-1995 data set. The solid lines are the fits described in the text.*

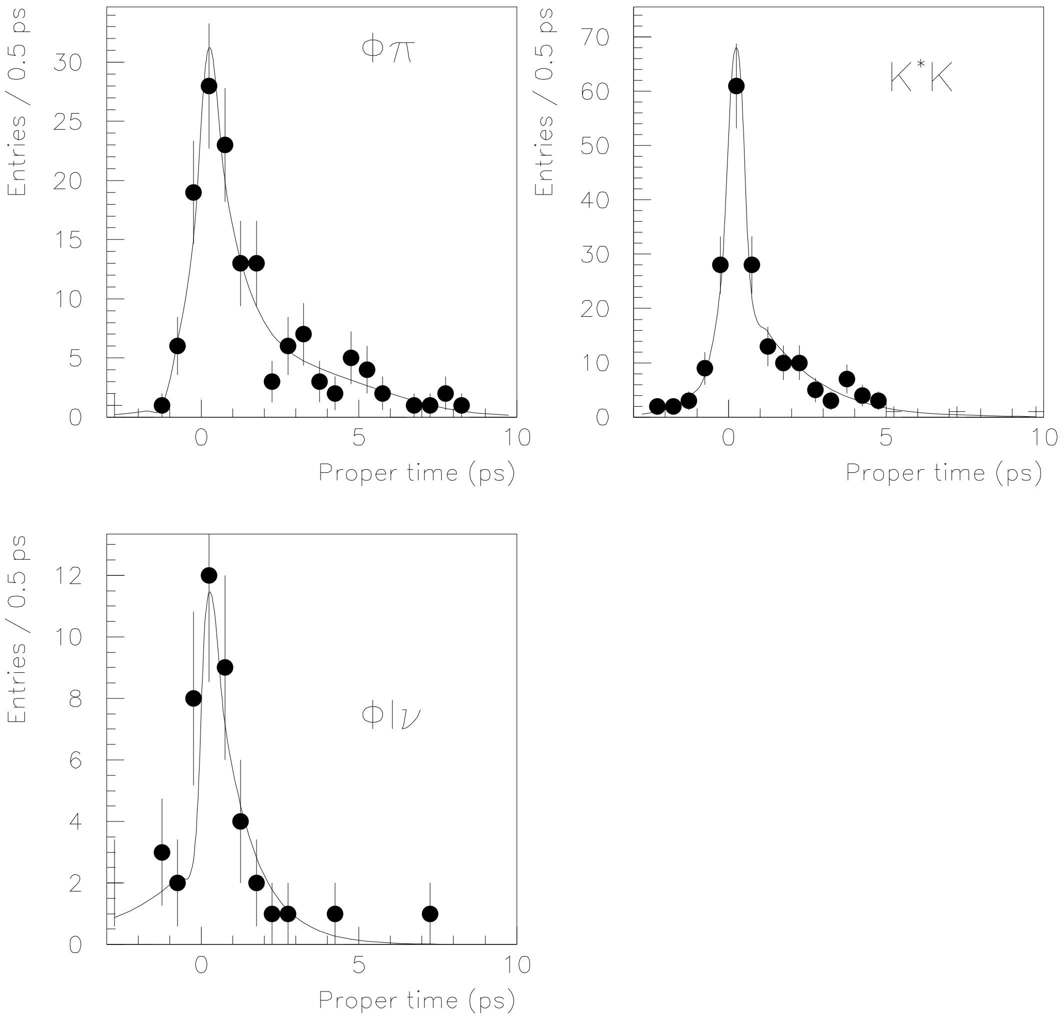


Figure 7.7: *Combinatorial background proper time distributions (side-band and like-sign) for each decay mode for the 1991-1993 data set. The solid lines are the fits described in the text.*

Chapter 8

B_s^0 Lifetime Difference

As was mentioned in chapter 2, there are a *light* (L) and a *heavy* (H) B_s^0 state which have different lifetimes: $\tau_L < \tau_H$. The lifetime that was fitted in the previous section is the mean of these two since $B_s^0 \longrightarrow D_s^{(*)-} l^+ \nu_l$ contains equal mixtures of the two *CP* eigenstates. One natural extension of the previous fit is, therefore, to fit for this lifetime difference using the same sample. If the mean lifetime is fixed, we have:

$$\tau_L = \tau_{B_s^0} - \frac{1}{2}\Delta\tau \tag{8.1}$$

$$\tau_H = \tau_{B_s^0} + \frac{1}{2}\Delta\tau \tag{8.2}$$

i.e. instead of fitting for two separate lifetimes we fit for the lifetime difference $\Delta\tau$ while keeping $\tau_{B_s^0}$ fixed at the world average value. Thus, the fit is a one parameter fit. Using the same fitting procedure as for the B_s^0 lifetime, we just parameterize $P_{B_s^0}$ as (see appendix B)

$$\frac{1}{2\tau_{B_s^0}}\left(e^{-\frac{t}{\tau_L}} + e^{-\frac{t}{\tau_H}}\right) \tag{8.3}$$

where $\tau_{B_s^0} = \frac{\tau_L+\tau_H}{2}$ is the mean lifetime. This can be written in terms of the decay widths for the two states (L) and (H):

$$\frac{\Gamma_L\Gamma_H}{\Gamma_L+\Gamma_H}\left(e^{-\Gamma_L} + e^{-\Gamma_H}\right) \tag{8.4}$$

which can be written in terms of the mean decay width and decay width difference:

$$\left(1 - \left(\frac{\Delta\Gamma}{2\Gamma}\right)^2\right)\frac{\Gamma}{2}e^{-\Gamma t}\left(e^{\frac{\Delta\Gamma t}{2}} + e^{\frac{-\Delta\Gamma t}{2}}\right) \tag{8.5}$$

Here:

- Γ is the total B_s^0 decay width, i.e. $\Gamma = \frac{1}{\tau_{B_s^0}}$
- Γ_L and Γ_H are the decay widths of the light and heavy states, respectively.
- $\Delta\Gamma$ is the decay width difference: $\Delta\Gamma = \Gamma_L - \Gamma_H$

To be precise, the fit parameter is actually the relative decay width difference $\frac{\Delta\Gamma}{\Gamma}$. As seen in figure 8.1, the minimum of the Likelihood function is at 0. Thus, the fit is not sensitive enough to separate the two lifetimes. The fitted value is

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} = 0.0^{+0.43}$$

The *probability density function* is then constructed to be

$$\mathcal{P} = e^{-\Delta\mathcal{L}} \tag{8.6}$$

where $\Delta\mathcal{L}$ is the likelihood difference with respect to the minimum. Integrating the probability density, an upper limit at 95% C.L. is obtained for the relative decay width difference (see figure 8.2):

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} \leq 56\% \quad @\ 95\%\ \text{C.L.}$$

This limit takes into account statistical uncertainties only. Since this limit hardly constrains $\Delta\Gamma_{B_s^0}$ at all, systematic uncertainties are not considered.

If, for the B_s^0 lifetime, the fitted value of 1.48 ps is used, the limit becomes

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} \leq 61\% \quad @\ 95\%\ \text{C.L.}$$

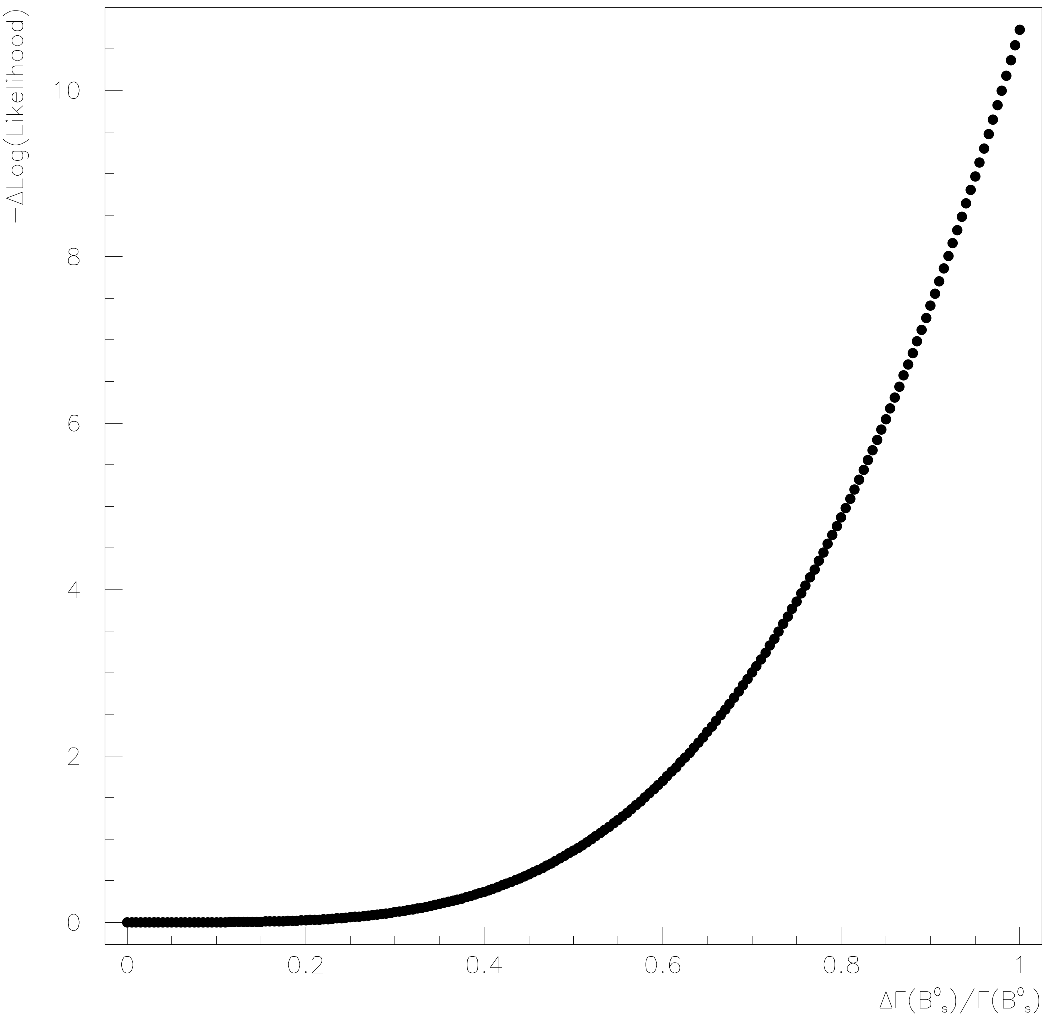


Figure 8.1: *Negative log-likelihood difference with respect to the minimum versus* $\Delta\Gamma(B_s^0)/\Gamma(B_s^0)$.

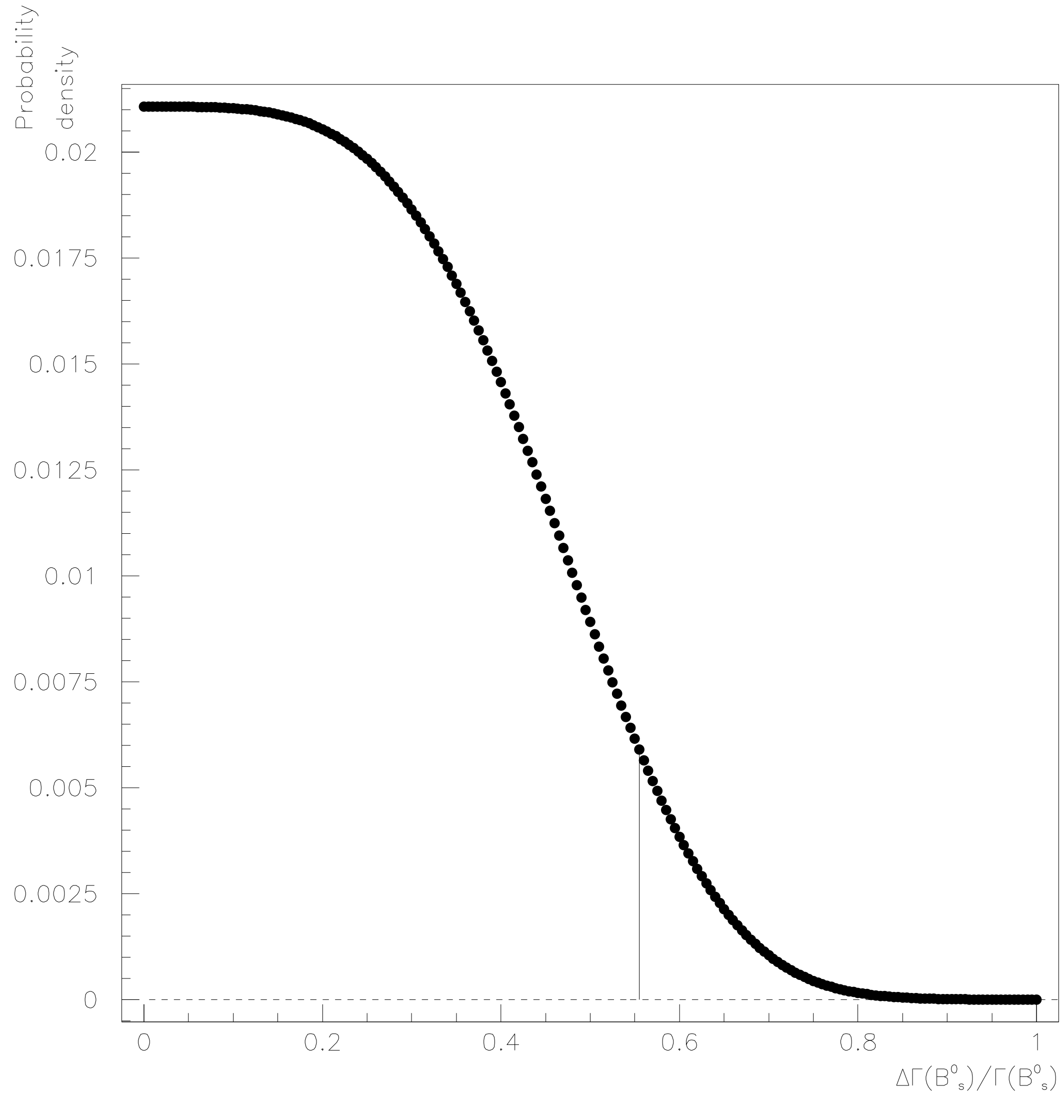


Figure 8.2: *Probability density function for $\Delta\Gamma(B_s^0)/\Gamma(B_s^0)$. The solid line shows the limit at 95% C.L.*

Chapter 9

$B_s^0 - \overline{B_s^0}$ Oscillations

9.1 General Remarks

The procedure to find the B_s^0 oscillation frequency is somewhat different than that of fitting the B_s^0 lifetime. We have to correctly decide if we have a B_s^0 or a $\overline{B_s^0}$ at both decay and production times and thus establish if the B_s^0 oscillated. We then have to fit the corresponding proper time distribution in the likelihood fit. However, as we shall see, there are several complications and another approach will be introduced: the *Amplitude method* [55].

9.2 $b(\bar{b})$ Tagging at Production Time

9.2.1 Tagging Variables Description

While the identity of the B_s^0 (and thus $b(\bar{b})$) at decay time is given by the lepton charge, it is more complicated to correctly identify the identity of the b-quark at production time. DELPHI has chosen to use an algorithm[1] based on probability density functions containing a combination of several variables that are sensitive to the initial quark state. For each variable V_i, one constructs the probability density functions $f_b(V_i)(f_{\bar{b}}(V_i))$ for $b(\bar{b})$ quarks and finds the ratio

$$R_i = f_{\bar{b}}(V_i)/f_b(V_i) \tag{9.1}$$

The combined tagging variable is then defined by

$$x_{\text{tag}} = \frac{1-R}{1+R} \tag{9.2}$$

where $R = \prod_i R_i$. This means that

$$-1 \leq x_{\text{tag}} \leq 1 \tag{9.3}$$

where a high value corresponds to a high probability that it was a b quark at production time. For independent variables V_i, this gives the optimal separation between b and $\bar{b}$ initial states.

[1] The tagging algorithm was made by G. Borisov, DELPHI [70]. This section is only meant to give a short description of this alghoritm with emphasis on the variables that are used. See reference [43] for more details.

Some variables are defined separately for both hemispheres while others are only defined in one of them. In the opposite hemisphere the following variables were used:

- Mean jet charge:
$$Q_{\text{jet}} = \frac{\sum_{i=1}^{n} q_i(\vec{p}_i \cdot \vec{e}_s)^\kappa}{\sum_{i=1}^{n}(\vec{p}_i \cdot \vec{e}_s)^\kappa} \tag{9.4}$$
where q_i and $\vec{p}_i$ are the charge and momentum of particle i, $\vec{e}_s$ is the unit vector along the thrust axis, $\kappa = 0.6$ [71] and the sum is over all charged tracks in the hemisphere.

- The weighted sum of the kaon candidates:
$$Q_K = \sum_{i=1}^{n_K} q_i(\vec{p}_i \cdot \vec{e}_s)^\kappa \tag{9.5}$$
where the sum is over all tracks identified as kaons.

- Secondary tracks: the sum of the charges of all tracks with large impact parameters with respect to the primary vertex.

- Primary tracks: the sum of the charges of all tracks compatible with the primary vertex.

- P_t leptons: charge-weighted transverse momentum of the lepton candidate with the highest momentum.

For the variables used in the same hemisphere as the B_s^0 candidate, all the tracks used in the B_s^0 candidate were removed before the variables were computed. They are, therefore, not sensitive to whether the B_s^0 candidate oscillated or not. The variables are:

- Mean jet charge: here the vector $\vec{e}_s$ is the unit vector along the B_s^0 candidate direction.

- Fragmentation kaon: the charge-weighted rapidity with respect to the direction of the thrust axis of the identified kaon with the highest momentum being compatible with originating from the primary vertex.

- Λ^0: the (proton) charge-weighted momentum of any reconstructed Λ^0.

Finally the polar angle of the thrust axis was used. The distribution of the tagging variable x_{tag} for b and $\bar{b}$ quark initial state is shown in figure 9.1.

In the 1995-1995 data set, 47% of the candidates were tagged as oscillated. The corresponding fraction for the 1991-1993 data set was 41%.

9.2.2 Tagging Purity

The average tagging purity was found from Monte Carlo simulation studies using B_s^0 signal events. For 100% efficiency it was found to be 74.6% $\pm$ 1.5%.

The improvement from using a simple jet charge can be seen in figure 9.2. The separation between b and $\bar{b}$ quarks using the jet charge is much smaller than that of the combined tagging variable. The corresponding purity for 100% efficiency is 65.6% $\pm$ 1.4%.

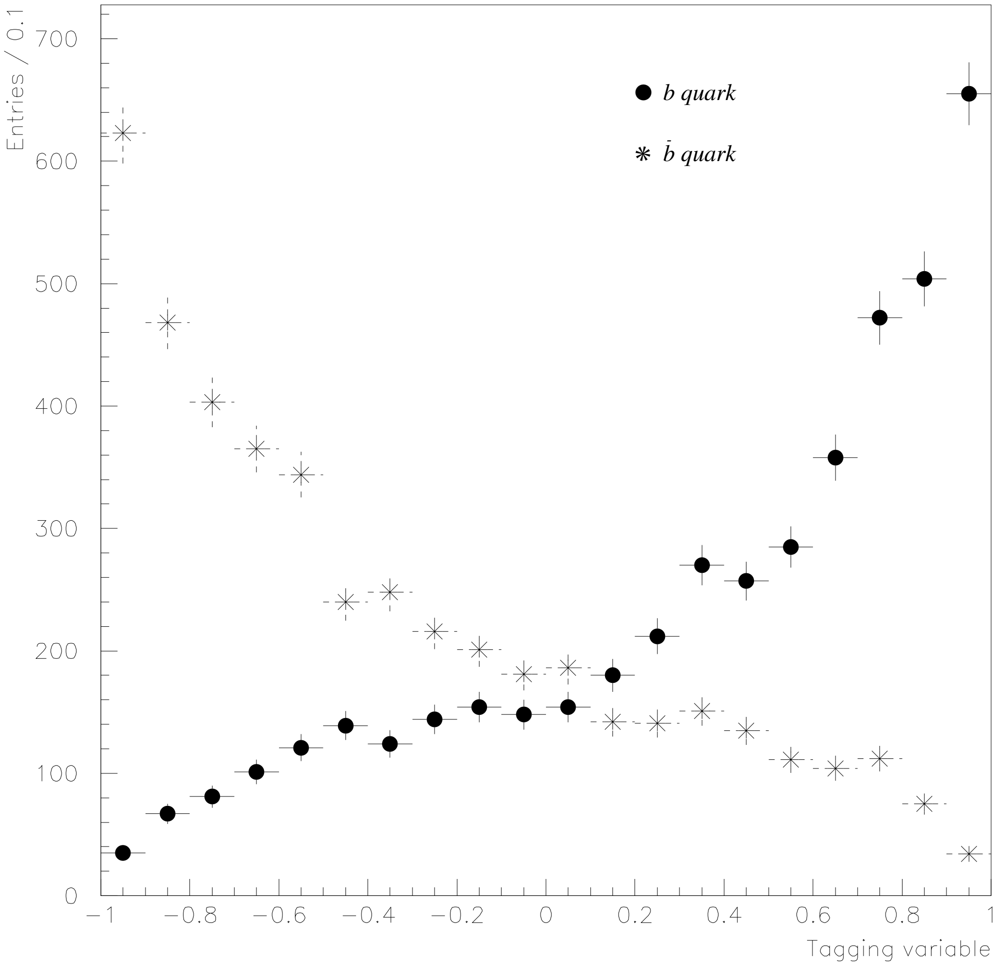


Figure 9.1: *Distribution of the initial quark state identity tagging variable for $b/\bar{b}$ initial states obtained from Monte Carlo simulations of* $B_s^0 \longrightarrow D_s^{(*)-} l^+ \nu_l$.

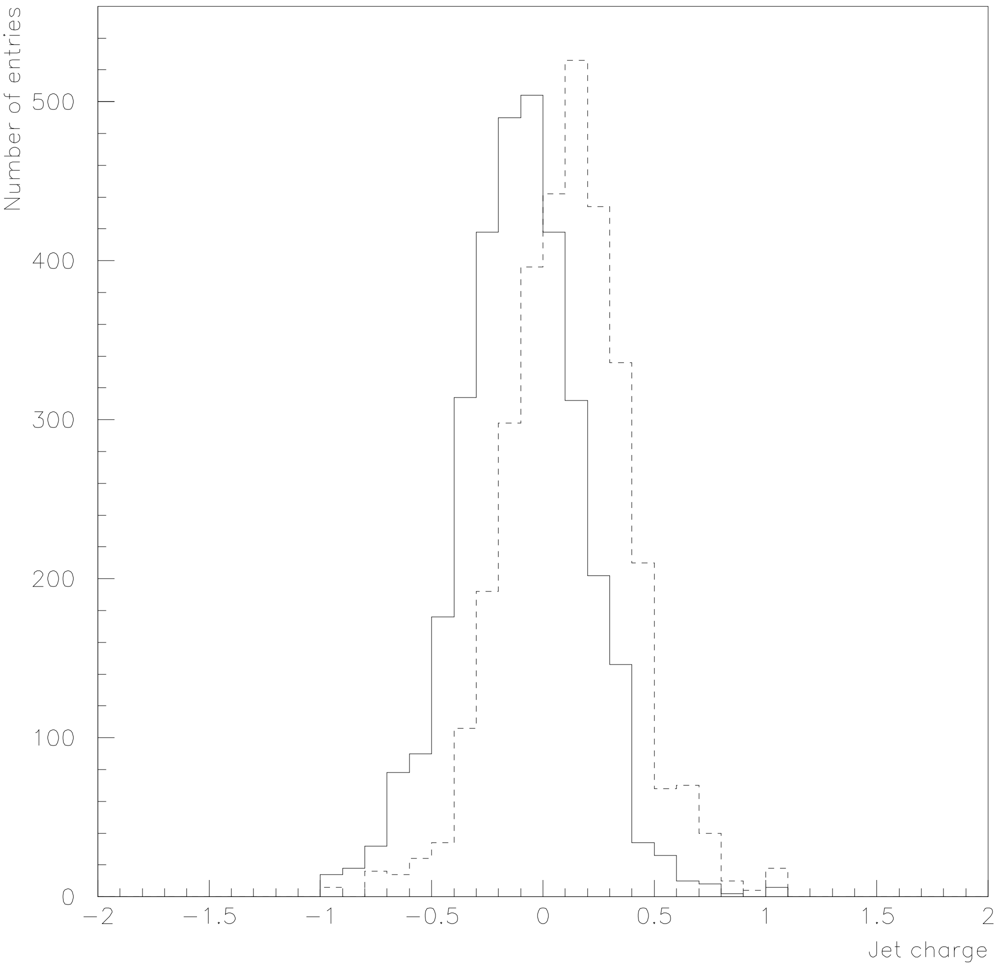


Figure 9.2: *Distribution of the jet charge for b (solid) and $\bar{b}$ (dashed) initial states obtained from Monte Carlo simulations of* $B_s^0 \longrightarrow D_s^{(*)-} l^+ \nu_l$.

9.3 The Likelihood Function

The likelihood function to be maximized for the B_s^0 oscillations is

$$\mathcal{L} = \prod_{i=1}^{N(\mathrm{mix})} P_{\mathrm{i}}^{\mathrm{mix}} \prod_{i=1}^{N(\mathrm{unmix})} P_{\mathrm{i}}^{\mathrm{unmix}} \tag{9.6}$$

where

$$P_{\mathrm{i}}^{\mathrm{mix}} = (1 - f_{\mathrm{comb}}) \times [f_{B_s^0} P_{B_s^0}^{\mathrm{mix}}(t_i) + f_{DD_s} P_{DD_s}^{\mathrm{mix}}(t_i) + f_{reflec} P_{reflec}^{\mathrm{mix}}(t_i)] + f_{\mathrm{comb}} P_{\mathrm{comb}}^{\mathrm{mix}}(t_i) \tag{9.7}$$

In the following, only the probability densities for the mixed events will be specified. To obtain those for the unmixed events one only has to change ϵ_b into $(1 - \epsilon_b)$. Here ϵ_b is the probability to correctly identify the $b(\bar{b})$ quark at production time with the tagging variable described in the previous section.

The different terms contributing to the mixed probability density are (see also section 7.2 for more details about the different terms):

- B_s^0 mixing probability density:

$$P_{B_s^0}^{\mathrm{mix}}(t_i) = [\ \epsilon_b \mathcal{P}_{B_s^0}^{\mathrm{mix}}(t) + (1 - \epsilon_b) \mathcal{P}_{B_s^0}^{\mathrm{unmix}}(t)\ ] \otimes G_{B_s^0}(t_i, t) \tag{9.8}$$

  where:

  - $\mathcal{P}_{B_s^0}^{\mathrm{mix}}(t)$ is the probability that starting with a B_s^0 meson at time t = 0, we observe a $\overline{B_s^0}$ decaying at time t:

$$\mathcal{P}_{B_s^0}^{\mathrm{mix}}(t) = \mathcal{P}\,(B_s^0 \longrightarrow \overline{B_s^0}) = \frac{1}{2\tau_{B_s^0}} e^{-\frac{t}{\tau_{B_s^0}}}\ (1 - \cos{(\Delta m_{B_s^0}\ t)}) \tag{9.9}$$

  - $G_{B_s^0}(t_i, t)$ is the signal resolution function.

- *Cascade* 'DD_s' decay mixing probability density:

  These events have three different contributions depending on their origin. Events originating from a B_d^0 have an oscillating contribution of the same type as the B_s^0 with $\Delta m_{B_d^0}$ and the effective lifetime $\tau_{B_d^0}^{eff}$ as fixed values. It contributes to the unmixed events, however, because of the opposite lepton charge (see section 5.8.3).

  Events coming from a B_s^0 (i.e. '$D_s D_s$' events) contribute equally to both the mixed and the unmixed events, since either of the D_s can decay semileptonically thus cancelling out any information about the identity of the B_s^0 meson at decay time. The proper time distribution is, therefore, an exponential with the effective B meson lifetime (taken to be the effective B_d^0 meson lifetime).

  Events from charged B mesons contribute to the mixed events because of the opposite lepton charge. These events do not oscillate of course, so their proper time distribution is a simple exponential with the effective B^+ meson lifetime.

  We have

$$f_{DDs}\, P^{\rm mix}_{DD_s}(t_m) = \left[\ f^{B^0_d}_{DDs}\, (\epsilon_b\, \mathcal{P}^{\rm unmix}_{B^0_d}(t) + (1-\epsilon_b)\, \mathcal{P}^{\rm mix}_{B^0_d}(t)) + \frac{f^{B^0_s}_{DsDs}}{2}\, e^{\frac{-t}{\tau_{eff}}} + f^{B^+}_{DDs}\, \epsilon_b\ e^{\frac{-t}{\tau_{eff}}}\ \right] \otimes G_{DD_s}(t_i, t) \quad (9.10)$$

where the effective lifetimes are defined in section 5.8.3, $G_{DD_s}(t_i, t)$ is the *cascade* resolution functions (one for each contribution), $f^{B^0_d}_{DDs}$ and $f^{B^+}_{DDs}$ are the fractions of *cascade* decay (see section 5.8) and $f^{B^0_s}_{DsDs} = 1.7\% \pm 0.8\%$. This last contribution was neglected in the lifetime fit.

- Mass *reflection* mixing probability density: these events come from B^0_d mesons so they have an oscillating contribution with $\Delta m_{B^0_d}$ and $\tau_{B^0_d}$ fixed to the world average values. We have

$$P^{\rm mix}_{reflec}(t_i) = [\ \epsilon_b \mathcal{P}^{\rm mix}_{B^0_d}(t) + (1-\epsilon_b)\mathcal{P}^{\rm unmix}_{B^0_d}(t)\ ] \otimes G_{B^0_s}(t_i, t) \quad (9.11)$$

- Combinatorial background mixing probability: the probability of classifying a combinatorial background event as mixed ($\epsilon^{\rm mix}_{\rm comb}$) was taken directly from the data (the $D_s(\phi)$ mass upper side-band). It was done separately for each data set. Due to the reduced statistics, this could not be done separately for each mode. 47.1% ± 4.7% and 40.4% ± 4.7% of the combinatorial background were found to be classified as mixed for the 1991-1993 and 1994-1995 data sets, respectively. The corresponding proper time distributions were the same as the ones used in the B^0_s lifetime fit.

$$P^{\rm mix}_{\rm comb}(t_i) = \epsilon^{\rm mix}_{\rm comb}\ \mathcal{P}_{\rm comb}(t) \quad (9.12)$$

Figure 9.3 shows a scan of the obtained negative log likelihood function versus the oscillation frequency $\Delta m_{B^0_s}$. As can be seen, the minimum is not significant enough to claim a measurement. In the following section we will extract a lower limit on $\Delta m_{B^0_s}$.

9.4 The Likelihood Function in the Case of No Sensitivity

In the case of the B^0_s lifetime, the likelihood function had a significant minimum around the fitted value and a measurement of the lifetime could be claimed. The lifetime at the minimum was an estimator of the true lifetime. In the case of the width difference, however, the fit was not sensitive to the true value of the lifetime difference. The likelihood had a minimum at zero width difference and only an upper limit on $\frac{\Delta\Gamma(B^0_s)}{\Gamma(B^0_s)}$ could be found. From figure 9.3 it is clear that a measurement of the oscillation frequency cannot be claimed and a limit must be determined. This case is, however, different from the previous case, and a new method is needed.

As can be seen from the definition of the likelihood, it cannot be absolutely normalized. It has to be referenced to a certain value and we have to find likelihood differences. Two possibilities exist: the likelihood can be referenced to its value at the frequency which minimizes the likelihood or to its value for an infinite oscillation frequency. The following sections, based on reference [55], will give more details about these two approaches.

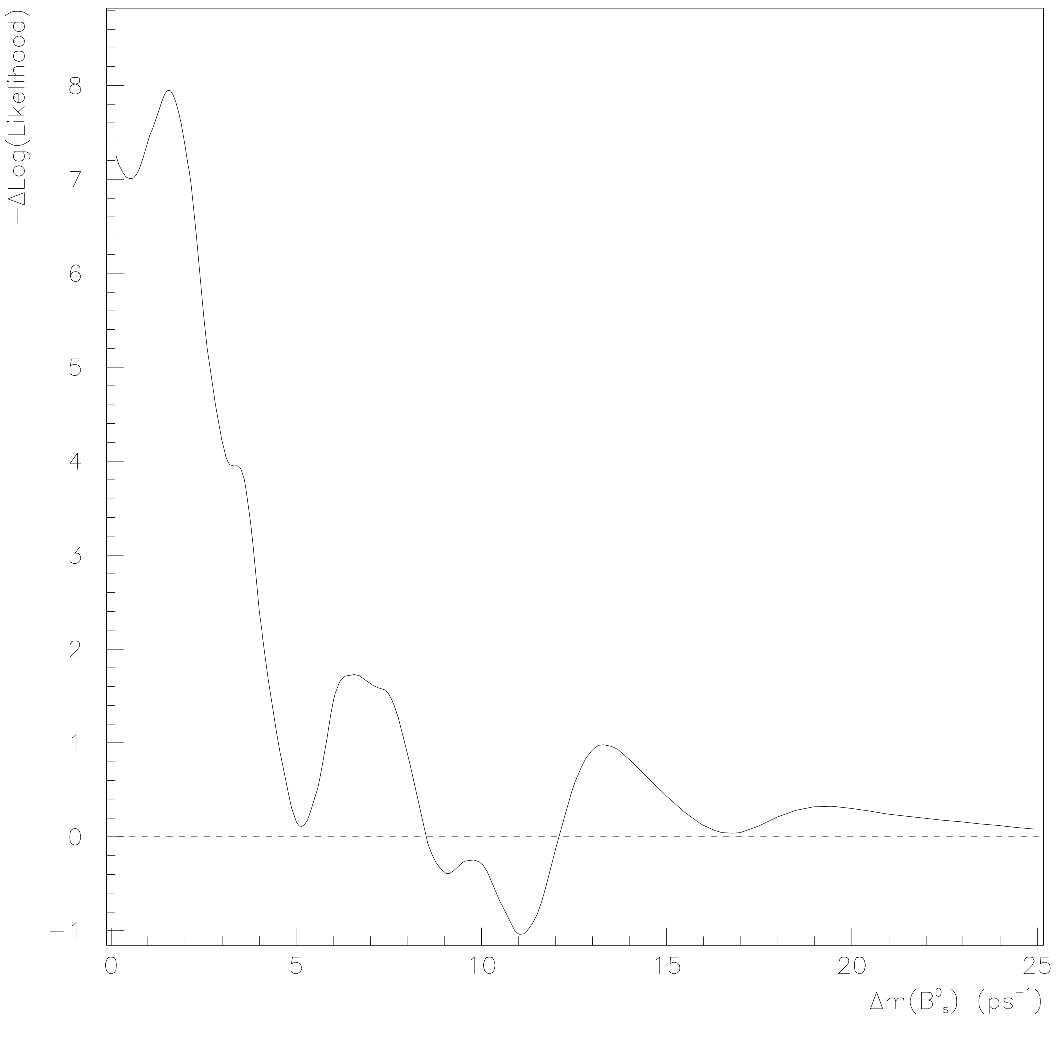


Figure 9.3: *Observed $\Delta m_{B_s^0}$ likelihood function referenced to an infinite oscillation frequency for the entire data sample.*

9.4.1 Likelihood Referenced to the Minimum

The classical use of the likelihood fit is to search for the frequency ν_{min} that minimizes the likelihood. One then looks at the difference with respect to this minimum:

$$\Delta\mathcal{L}^{\min}(\nu) = -\mathrm{Log}(L(\nu)) - (-\mathrm{Log}(L(\nu_{\min}))) \tag{9.13}$$

The minimum frequency $\nu_{\min}$ is in principle a good and unbiased estimator of the true frequency. There are then well established procedures so that one can exclude all frequencies ν satisfying

$$\Delta\mathcal{L}^{\min}(\nu) > \Delta\mathcal{L}_{\mathrm{cut}} \tag{9.14}$$

with a certain confidence level. For example, $\Delta\mathcal{L}_{\mathrm{cut}}$ is 1.34 for a single sided 95% C.L. Unfortunately, these well established values cannot be used here. Even if the real oscillation frequency lies beyond the sensitivity of the fit, one may still find a minimum in the likelihood at some value lower than the true value. In the absence of a signal, one may get a minimum in a region where statistical fluctuations due to backgrounds have the same strength as the expected signal. In this case, the minimum frequency is obviously no longer a good (unbiased) estimator of the true frequency, the statistical distribution is not necessarily Gaussian and the well established $\Delta\mathcal{L}_{\mathrm{cut}}$ values are no longer valid.

This can be shown by using a *toy Monte Carlo*, i.e. a Monte Carlo simulation where we directly simulate the different distributions (proper time, tagging variable, etc) for the different contributions in the sample. This enables us to make a large number of experiments and study their statistical behaviour. In the first case, 25 experiments of a sample with 400 events having the characteristics of the 1994-1995 $D_s^- \longrightarrow \phi\pi^-$ mode and $\Delta m_{B_s^0} = 3$ ps^{-1} were simulated. The values of $\Delta m_{B_s^0}$ and ΔLog (Likelihood) at the minimum of the likelihood function are plotted in figure 9.4. We see that in all experiments the minimum is at the generated frequency, and the minima are significant enough to claim a measurement.

In figure 9.5 we see the same plots for 100 experiments of a sample of 110 events with $D_s^- \longrightarrow \phi\pi^-$ characteristics and $\Delta m_{B_s^0} = 6$ ps^{-1}. In this case, we are at the limit of the sensitivity and we see that the minimum value of $\Delta m_{B_s^0}$ is distributed around the generated value.

In figure 9.6 the same variables are plotted for 200 experiments of 55 events with $D_s^- \longrightarrow \phi\pi^-$ characteristics and $\Delta m_{B_s^0} = 20$ ps^{-1}. We are now no longer sensitive to the true value and the minimum of $\Delta m_{B_s^0}$ can no longer be used as an estimator.

This difficulty can be resolved using the toy Monte Carlo to calibrate the likelihood. One generates a large number of experiments with exactly the same characteristics as the events in the real sample. For each experiment one finds $\nu_{\min}$ and $\Delta\mathcal{L}^{\min}(\nu)$. The value $\Delta\mathcal{L}_{\mathrm{cut}}$ can then be found so that we really have $\Delta\mathcal{L}^{\min}(\nu) < \Delta\mathcal{L}_{\mathrm{cut}}$ for 95% of the generated experiments. What is found [55] is that the actual *cut value* $\Delta\mathcal{L}_{\mathrm{cut}}$ is always larger than the theoretical value expected from Gaussian statistics. In addition, this cut value actually depends on the frequency which means one must make a *cut curve* $\Delta\mathcal{L}_{\mathrm{cut}}(\Delta m_{B_s^0})$. The limit will then be the intersection between the curve and the likelihood difference obtained from the data.

There are several disadvantages with this approach. In particular, it involves CPU-intensive calibration studies. It is also very difficult to add several different analyses (eventually from different experiments) together because of the need to calibrate the likelihood.

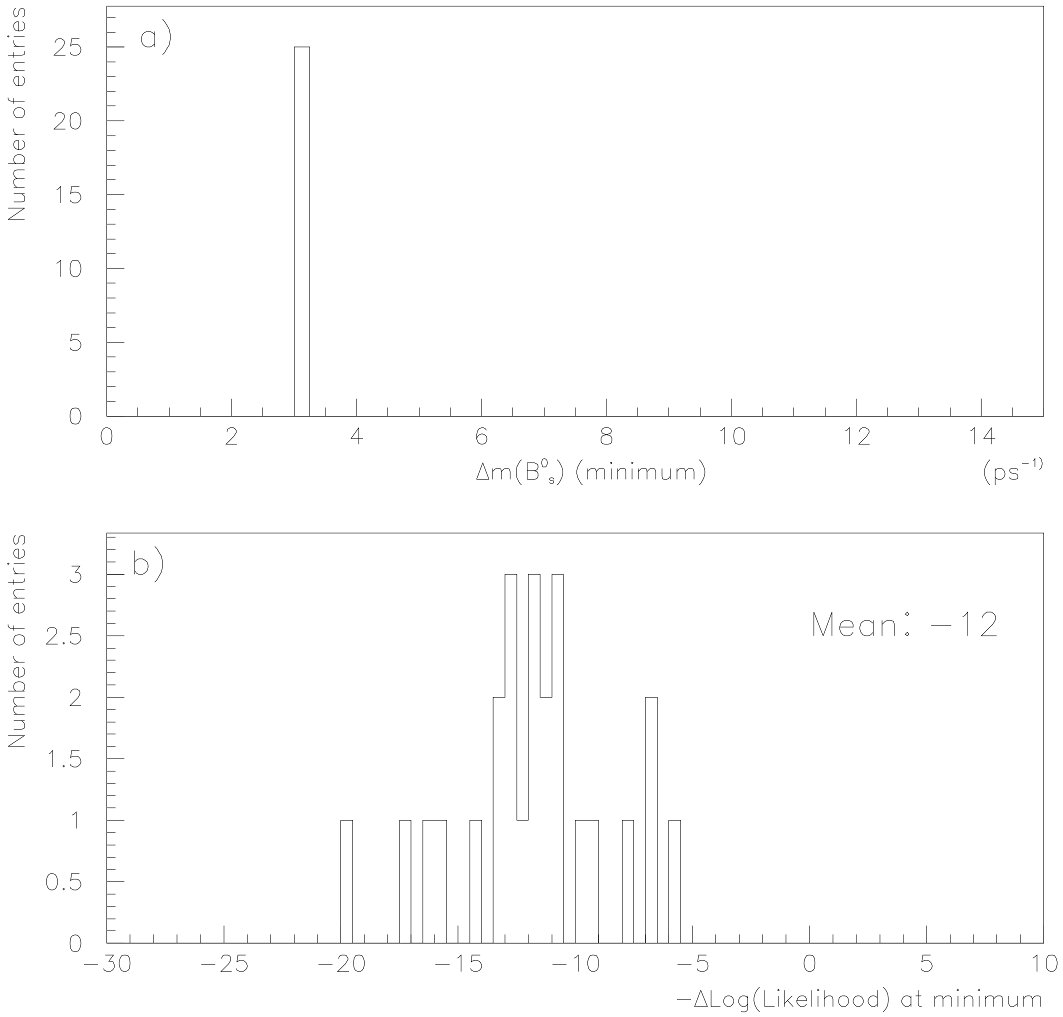


Figure 9.4: *Toy Monte Carlo simulations - Sensitive to the true value of* $\Delta m_{B_s^0}$*: 25 experiments of a sample with 400 events with* $D_s^- \longrightarrow \phi\pi^-$ *characteristics and* $\Delta m_{B_s^0} = 3$ ps^{-1} *were simulated. Figure a) shows the value of* $\Delta m_{B_s^0}$ *at the minimum of the likelihood function. Figure b) shows the value of* $-\Delta$*Log(Likelihood) at the minimum.*

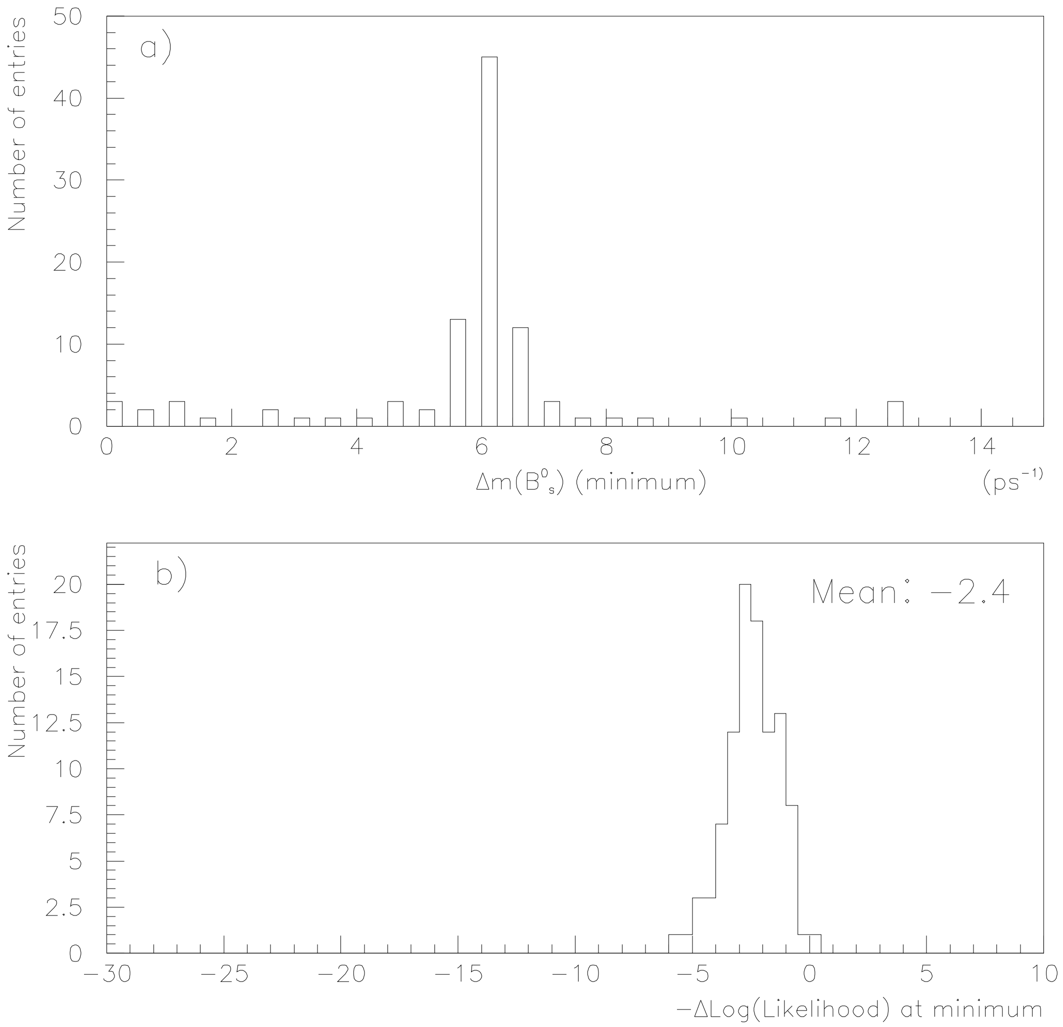


Figure 9.5: *Toy Monte Carlo simulations - At the limit of sensitivity to the true value of $\Delta m_{B^0_s}$: 100 experiments of a sample with 110 events with $D^-_s \longrightarrow \phi\pi^-$ characteristics and $\Delta m_{B^0_s} = 6$ ps^{-1} were simulated. Figure a) shows the value of $\Delta m_{B^0_s}$ at the minimum of the likelihood function. Figure b) shows the value of $-\Delta$Log(Likelihood) at the minimum.*

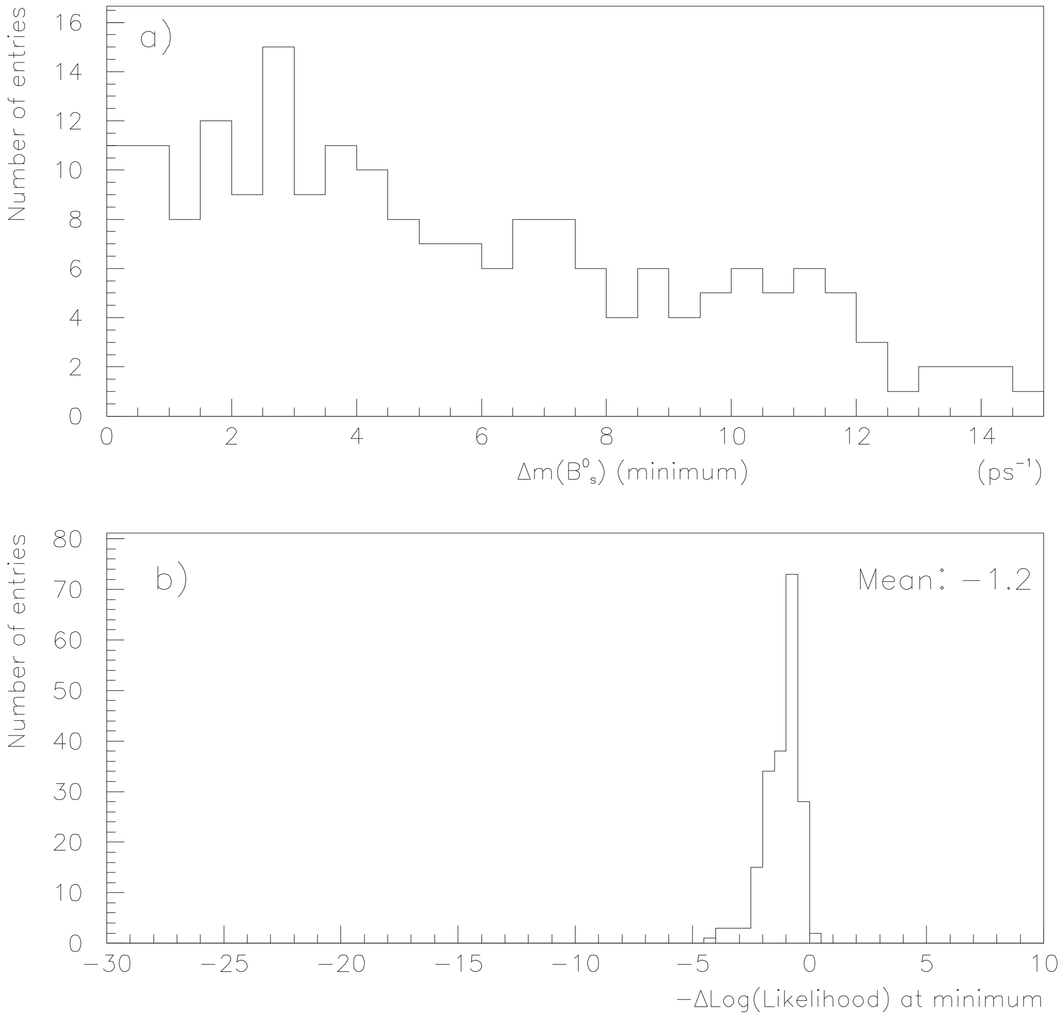


Figure 9.6: *Toy Monte Carlo simulations - No sensitivity to the true value of $\Delta m_{B_s^0}$: 200 experiments of a sample with 55 events with $D_s^- \longrightarrow \phi\pi^-$ characteristics and $\Delta m_{B_s^0}$ = 20 ps^{-1} were simulated. Figure a) shows the value of $\Delta m_{B_s^0}$ at the minimum of the likelihood function. Figure b) shows the value of $-\Delta$Log(Likelihood) at the minimum.*

Finally it can be shown that the sensitivity of the fit is decreased in this approach because the limit obtained is actually lower than the one determined with other methods [55]. This method is, therefore, not well suited for this analysis.

9.4.2 Likelihood Referenced to Infinity

In this case the reference point of the likelihood is the value of the likelihood at infinity, i.e. an infinite oscillation frequency:

$$\Delta\mathcal{L}^{\infty}(\nu) = -\mathrm{Log}(L(\nu)) - (-\mathrm{Log}(L(\infty))) \tag{9.15}$$

The central limit theorem of likelihood theory states that for a large number of events, the likelihood difference between two hypotheses (with oscillation frequency ν and with oscillation frequency $\nu = \infty$) is χ^2-distributed and the statistical fluctuations of $\Delta\mathcal{L}^{\infty}(\nu)$ are Gaussian. One can then exclude any frequency ν such that

$$\Delta\mathcal{L}^{\infty}(\nu) > \Delta\mathcal{L}^{\infty}_{\mathrm{Average}}(\nu) \; + \; 1.645 \cdot \sigma(\Delta\mathcal{L}^{\infty}(\nu)) \tag{9.16}$$

where:

- $\Delta\mathcal{L}^{\infty}(\nu)$ is the value of the likelihood difference obtained for the data sample at frequency ν.
- $\Delta\mathcal{L}^{\infty}_{\mathrm{Average}}(\nu)$ is the average value of the likelihood difference as expected for real oscillation with frequency ν.
- $\sigma(\Delta\mathcal{L}^{\infty}(\nu)$ is the variance of $\Delta\mathcal{L}^{\infty}_{\mathrm{Average}}(\nu)$

One, therefore, has to generate toy Monte Carlo experiments with different oscillation frequencies and for each frequency one needs to find the average $\Delta\mathcal{L}^{\infty}_{Average}(\nu)$ over all the experiments and the variance of this distribution. This defines a curve according to equation (9.16) as a function of the frequency and the intersection between this curve and the likelihood difference obtained for the the data sample defines the upper limit.

This method also needs lengthy, CPU-intensive toy Monte Carlo studies. This approach is interesting, however, because it can be shown that the limits obtained are identical with the limits obtained from Fourier analysis. In the next section the Fourier/Amplitude approach will be detailed. One should bear in mind that even though this method is rather new, it can formally be reduced to the likelihood difference with respect to infinity.

9.5 The Amplitude Method

The idea behind the amplitude method [55] is Fourier analysis, which is a well established method for analysis of periodic signals. The main advantages of the amplitude method is that:

- One can dispense of CPU-intensive Monte Carlo studies.
- It makes it easy to add systematic effects in a coherent and systematic way (something which is a problem in the likelihood approach).

- It also makes it easy to see the effect of statistical fluctuations and systematic biases.
- It makes it possible to combine different analyses, even from different experiments, in an easy way.

9.5.1 The Amplitude

Conceptually, this method is somewhat different from the likelihood methods since we now are looking for a peak in a frequency spectrum. The real part of the Fourier transformation of cos (ωt) is equal to $\sqrt{\frac{2}{\pi}}\,\delta(\nu-\omega)$, i.e. it has a peak at the true oscillation frequency $\nu = \omega$. Fluctuations of the background then correspond to noise. It becomes clear that contrary to the standard approach , the relevant variable is *not* $\Delta m_{B_s^0}$, but the Fourier amplitude at the frequency $\Delta m_{B_s^0}$.

In

$$\mathcal{P}_{B_s^0}^{\mathrm{mix}}(t) = \mathcal{P}\,(B_s^0 \longrightarrow \overline{B_s^0}) = \frac{1}{2\tau_{B_s^0}} e^{-\frac{t}{\tau_{B_s^0}}} \;(1 - A\cos{(\Delta m_{B_s^0}\;t)}) \tag{9.17}$$

$$\mathcal{P}_{B_s^0}^{\mathrm{unmix}}(t) = \mathcal{P}\,(B_s^0 \longrightarrow B_s^0) = \frac{1}{2\tau_{B_s^0}} e^{-\frac{t}{\tau_{B_s^0}}} \;(1 + A\cos{(\Delta m_{B_s^0}\;t)}) \tag{9.18}$$

a value of $\Delta m_{B_s^0}$ is assumed and the *amplitude* A is fitted. For $A = 1$ we recover the usual time distributions. For infinite statistics, the amplitude should be 1 for the true value of $\Delta m_{B_s^0}$ and 0 for values away from the true value. In a real experimental situation, the fit of the amplitude will take into account experimental conditions like resolutions, efficiencies, backgrounds, etc. The individual fitted values of the amplitude A_m for given values of $\Delta m_{B_s^0}$ are Gaussian distributed. It is this fact that makes averaging over different analyses/experiments so easy in this approach. One only has to average the different values of A_m taking into account their respective errors. The method for setting a limit on $\Delta m_{B_s^0}$ is that one can exclude a given value of $\Delta m_{B_s^0}$ with 95% C.L. if the amplitude satisfies [55]

$$A + 1.645\;\sigma_A < 1 \tag{9.19}$$

where σ_A is the uncertainty on the amplitude A. The procedure is to plot $A + 1.645\;\sigma_A$ as a function of $\Delta m_{B_s^0}$, and the intersection between the curve and the line $A = 1$ then determines the limit.

9.5.2 Exclusion Probability

In a given experiment one can be "lucky", meaning that the fitted values of the amplitude A fluctuate downwards, and one can set an "artificially" high limit. Of course, one can also be "unlucky", i.e. the amplitudes can fluctuate upwards (towards 1), which means that the obtained limit will decrease. One should, therefore, also give the *exclusion probability*, i.e. the probability of obtaining a 95% C.L. for a given value of $\Delta m_{B_s^0}$. One assumes that the true value of $\Delta m_{B_s^0}$ is beyond the sensitivity of the fit, so the expected value of the amplitudes A all are zero. One then calculates the probability that a fluctuation gives $A = 1$. One can then write

Sample characteristics	Sensitivity
Usual $\phi\pi$ characteristics	10.4
Increased statistics	11.1
Reduced background	11.6
Improved decay length resolution	12.7
Improved momentum resolution	11.2
Improved tagging	13.2

Table 9.1: *Variations of the sensitivity for a sample of 400 events of the* $D_s^- \longrightarrow \phi\pi^-$ *type with a 25% improvement of the different parameters.*

$$\mathcal{P}_{\text{exclusion}} = 1 - \int_{1-1.645\,\sigma_m}^{\infty} G(A, 0, \sigma_{A_m})\; dA \tag{9.20}$$

Here, one assumes that each fitted value of the amplitude A_m is described by a Gaussian probability density function $G(A, 0, \sigma_{A_m})$ which is centered at $A = A_m$ and has an uncertainty on the fitted amplitude equal to σ_{A_m}.

9.5.3 Sensitivity

Another useful quantity is the *sensitivity*. From reference [55] we have that the expected signal to background ratio is

$$S/B = \sqrt{\frac{n}{2}}\; f_s\; (1 - 2\eta)\; D(\omega, \sigma_l, \sigma_p) \tag{9.21}$$

where:

- n is the total number of events.
- f_s is the fraction of B_s^0 signal events in the sample.
- η is the mistag probability of the $b(\bar{b})$ identity at production time.
- $D(\omega, \sigma_l, \sigma_p)$ is a function taking into account the proper time resolution.

Table 9.1 gives the expected sensitivity for a sample of 400 events with a resolution and a background fraction corresponding to that in the 1994-1995 $D_s^- \longrightarrow \phi\pi^-$ mode (this is somewhat more optimistic than the total data sample). The table gives the improvement of the sensitivity one can expect with a 25% improvement in the different parameters. The main improvements come from the decay length resolution and the tagging.

The formula for the sensitivity depends only on the experimental characteristics and not on the data set itself. Graphically, the sensitivity can be plotted on the same plot as the amplitude. One assumes no sensitivity, i.e. all $A_m = 0$, and plot the curve $1.645\,\sigma_{A_m}$. The intersection with the line $A = 1$ then gives the sensitivity. The sensitivity corresponds to 50% exclusion probability.

9.5.4 The Amplitude Method in Case of a Measurement

Figures 9.7 and 9.8 show the amplitude fit and the corresponding likelihood difference with respect to an infinite oscillation frequency for a Monte Carlo simulation sample of 1000 B_s^0 signal events for $\Delta m_{B_s^0} = 3.0\,\text{ps}^{-1}$. There is a significant minimum of the likelihood at the generated frequency. Away from the generated value the amplitudes are compatible with 0, while the amplitude at the generated frequency is compatible with 1.

9.5.5 The Amplitude Method in Case of a Limit

Figures 9.9 and 9.10 show the amplitude fit and the corresponding likelihood difference for a Monte Carlo simulation B_s^0 signal sample of 250 events for $\Delta m_{B_s^0} = 11.25\,\text{ps}^{-1}$. The minimum in the likelihood is not significant enough to claim a measurement, even though the minimum is around the generated value. With the amplitude plot one can set a lower limit of $\Delta m_{B_s^0} > 11\,\text{ps}^{-1}$. The sensitivity is 12.5 ps^{-1}, i.e. just above the generated value. This is consistent with the minimum of the likelihood being around the generated value of $\Delta m_{B_s^0}$.

Figures 9.11 and 9.12 show the amplitude fit and the corresponding likelihood difference for the same sample in the case of only 100 events. Using the amplitude method, one can set a lower limit of $\Delta m_{B_s^0} > 6.5\,\text{ps}^{-1}$ while the sensitivity is 7.8 ps^{-1}. The likelihood has a local minimum around $\Delta m_{B_s^0} = 6.5$ ps^{-1} (i.e. far away from the true value) and then flattens out after $\Delta m_{B_s^0} = 11$ ps^{-1}.

9.6 Results

Using the amplitude method, the following excluded regions for the mass difference $\Delta m_{B_s^0}$ were obtained using the total event sample (see figure 9.13):

$$0 \leq \Delta m_{B_s^0} \leq 4.5\,\text{ps}^{-1} \text{ @ 95\% C.L.}$$

$$5.8\,\text{ps}^{-1} \leq \Delta m_{B_s^0} \leq 7.7\,\text{ps}^{-1} \text{ @ 95\% C.L.}$$

The sensitivity is $\Delta m_{B_s^0} = 6.9\,\text{ps}^{-1}$.

9.7 Systematic Uncertainties

Systematic uncertainties were estimated by varying the different parameters in the likelihood function which were kept constant in the fit. The parameters were varied according to their measured or expected uncertainties. To take into account that the variations may affect both the amplitude and the uncertainty on the amplitude, the systematic uncertainty was calculated according to reference [55]:

$$\sigma_A^{\text{syst}} = A^{\text{new}} - A + (1 - A) \times \frac{\sigma_A^{\text{new}} - \sigma_A}{\sigma_A} \tag{9.22}$$

where A, σ_A and $A^{\text{new}}, \sigma_A^{\text{new}}$ are the amplitude and its uncertainty before and after the variation of the parameter.

The following contributions were considered:

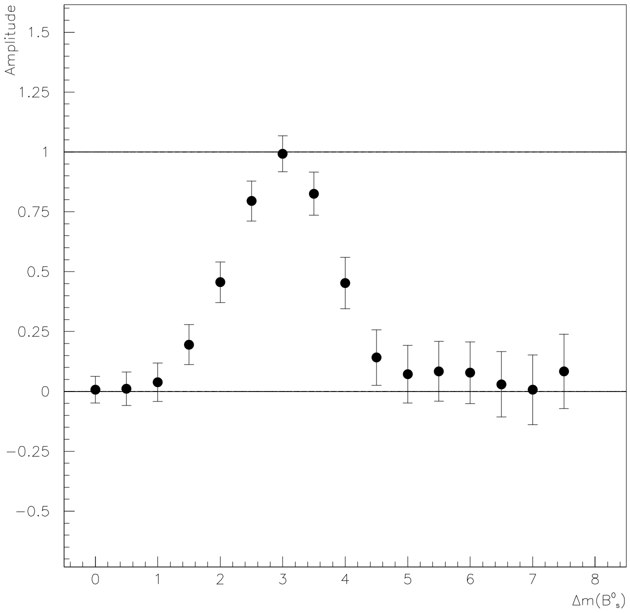


Figure 9.7: *Amplitude plot for a simulated sample of 1000 B_s^0 signal events for* $\Delta m_{B_s^0} = 3.0\,\mathrm{ps}^{-1}$

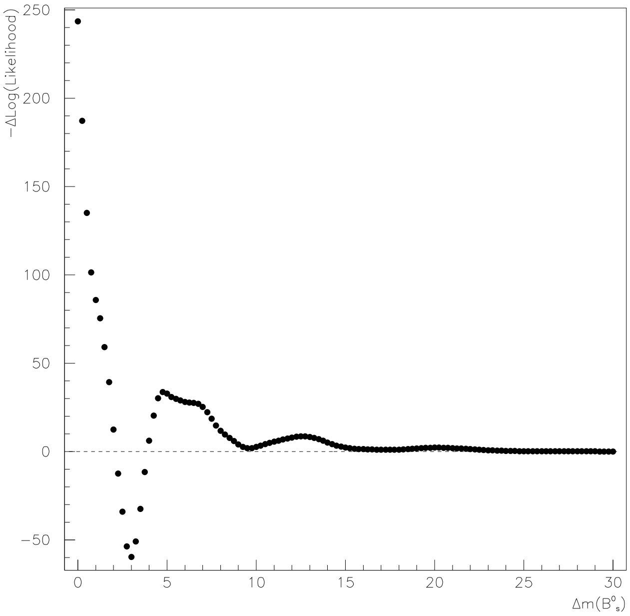


Figure 9.8: *Likelihood plot for a simulated sample of 1000 B_s^0 signal events for* $\Delta m_{B_s^0} = 3.0\,\mathrm{ps}^{-1}$.

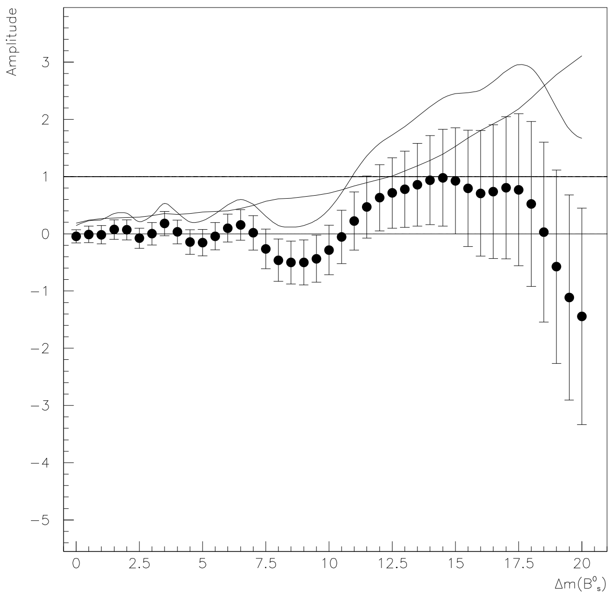


Figure 9.9: *Amplitude plot for a simulated sample of 250 B_s^0 signal events for* $\Delta m_{B_s^0} = 11.25\,\mathrm{ps}^{-1}$. *Also plotted are the sensitivity and A + 1.645σ_A, where A and σ_A are the amplitude and its uncertainty, respectively.*

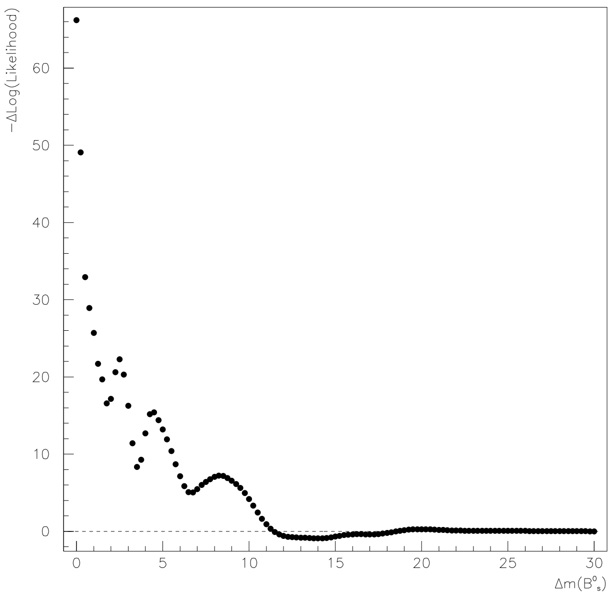


Figure 9.10: *Likelihood plot for a simulated sample of 250 B_s^0 signal events for* $\Delta m_{B_s^0} = 11.25\,\mathrm{ps}^{-1}$.

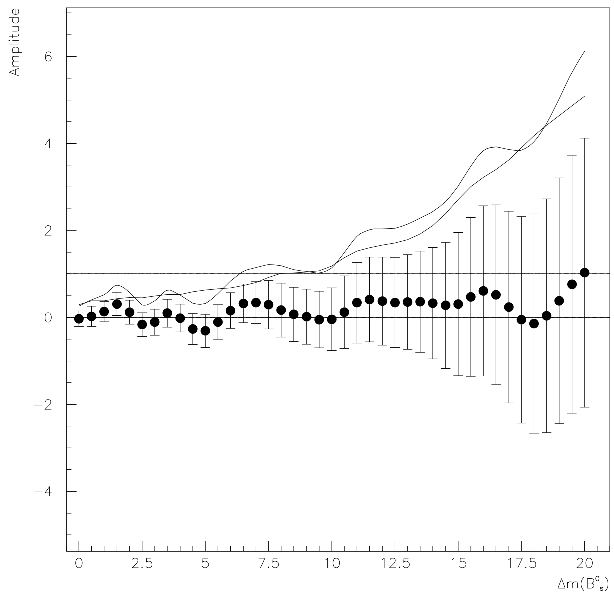


Figure 9.11: *Amplitude plot for a simulated sample of 100* B_s^0 *signal events for* $\Delta m_{B_s^0} = 11.25\,\mathrm{ps}^{-1}$. *Also plotted are the sensitivity and* $A + 1.645\sigma_A$, *where* A *and* σ_A *are the amplitude and its uncertainty, respectively.*

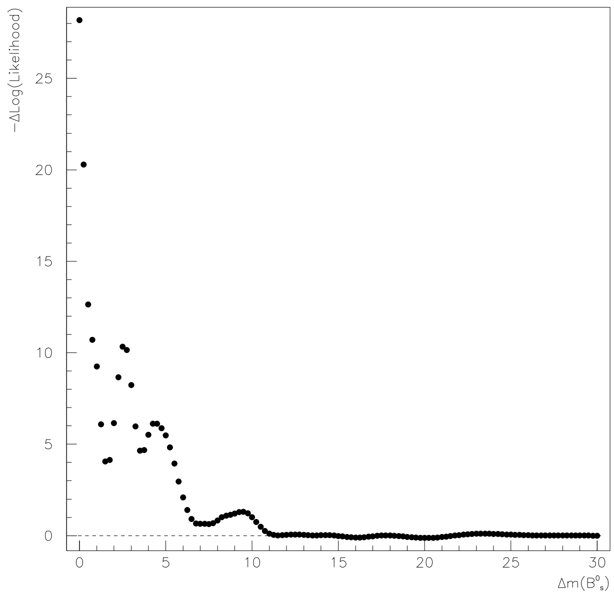


Figure 9.12: *Likelihood plot for simulated sample of 100* B_s^0 *signal events for* $\Delta m_{B_s^0} = 11.25\,\mathrm{ps}^{-1}$.

- $b/\bar{b}$ tagging purity:
  Studies done in reference [43] show that the data - Monte Carlo simulation agreement for the tagging purity in the opposite hemisphere of the B_s^0 candidate is 0.01 ± 0.02. The tagging purity was varied ±3.0%
- Signal proper time resolution function: Same procedure as for the B_s^0 lifetime fit.
- Background proper time parameterization: Same procedure as for the B_s^0 lifetime fit.
- B_s^0 purity and combinatorial background fractions: Same procedure as for the B_s^0 lifetime fit.
- Combinatorial background tag:
  The fraction of combinatorial background tagged as oscillating was varied within its statistical uncertainty.

The different contributions were assumed to be independent. They were added in quadrature to give the total systematic uncertainty.

Including the systematic uncertainty, the following excluded regions for the mass difference $\Delta m_{B_s^0}$ were obtained:

$$0 \leq \Delta m_{B_s^0} \leq 4.4\,\mathrm{ps}^{-1} \quad @\ 95\%\ \mathrm{C.L.}$$

$$6.0\,\mathrm{ps}^{-1} \leq \Delta m_{B_s^0} \leq 7.6\,\mathrm{ps}^{-1} \quad @\ 95\%\ \mathrm{C.L.}$$

The sensitivity decreases to

$$\Delta m_{B_s^0} = 6.7\,\mathrm{ps}^{-1}$$

The exclusion probability at $\Delta m_{B_s^0} = 7.6\,\mathrm{ps}^{-1}$ is 45% (see figure 9.15).

9.8 The asymmetry $N^{\mathrm{no\ osc}} - N^{\mathrm{osc}}$

Figure 9.16 shows the asymmetry $N^{\mathrm{no\ osc}} - N^{\mathrm{osc}}$. From equation (2.126) we know that this difference depends on cos $(\Delta m_{B_s^0} t)$. As can be seen from the figure, the oscillations are completely washed out by resolution effects and backgrounds and nothing can be deduced about $\Delta m_{B_s^0}$.

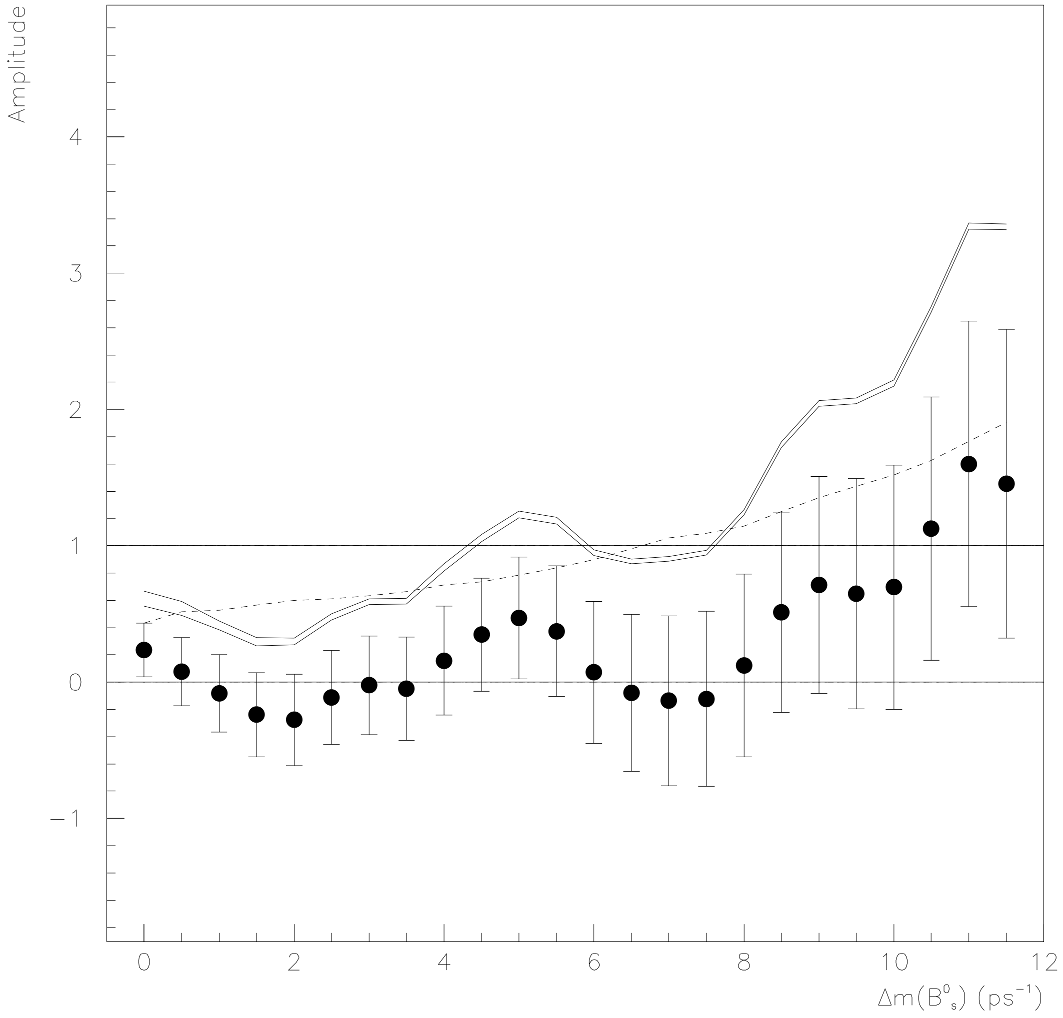


Figure 9.13: *Amplitude plot versus* $\Delta m_{B_s^0}$ *for the entire data sample: The points are the fitted amplitudes (A) with the corresponding statistical uncertainties* (σ_A). *The lower continuous line represents* $A + 1.645\,\sigma_A$. *The upper continuous line corresponds to the contribution from the combined statistical and systematic uncertainties. The excluded values of* $\Delta m_{B_s^0}$ *are given by the intersections between the upper continuous line and the line A = 1. The sensitivity is represented by the dashed line.*

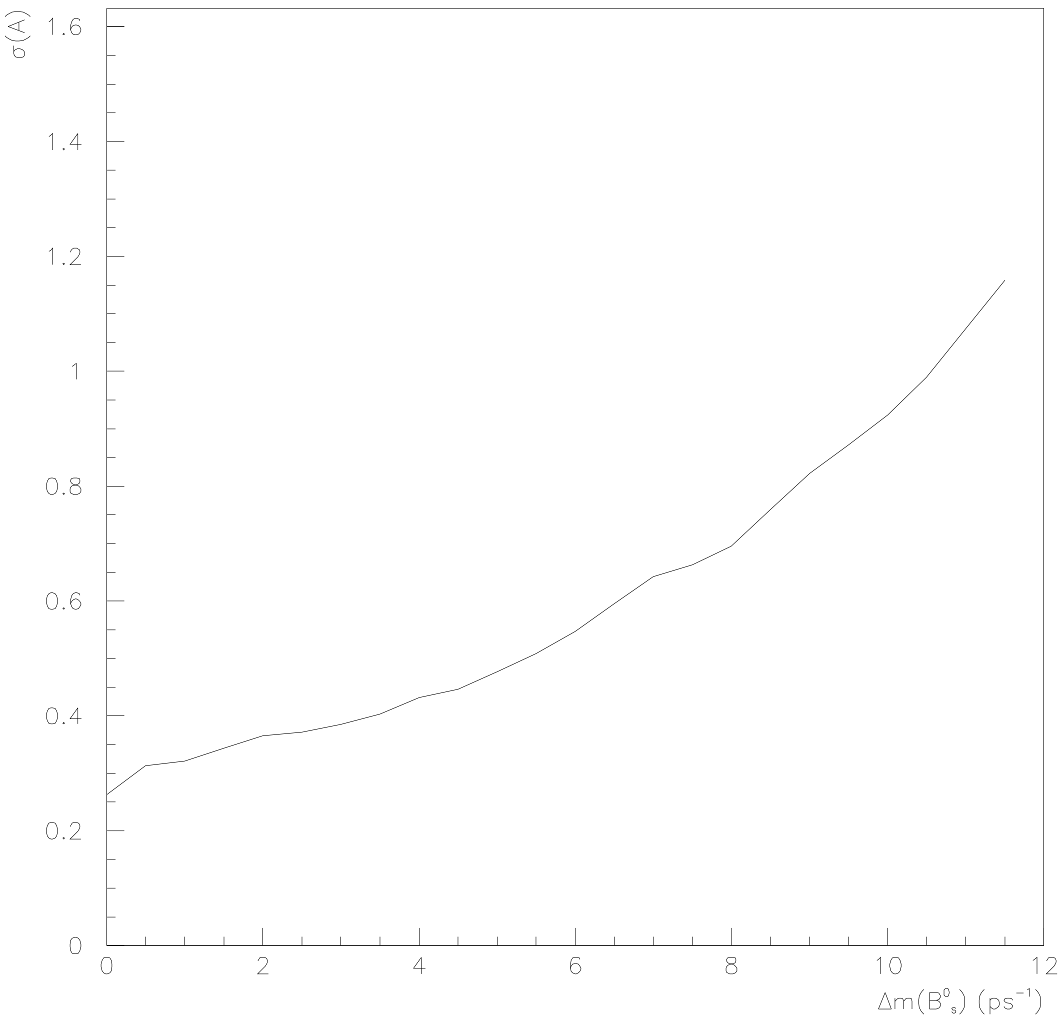


Figure 9.14: *Uncertainty on the oscillation amplitude versus the mass difference* $\Delta m_{B_s^0}$ *for the entire data sample.*

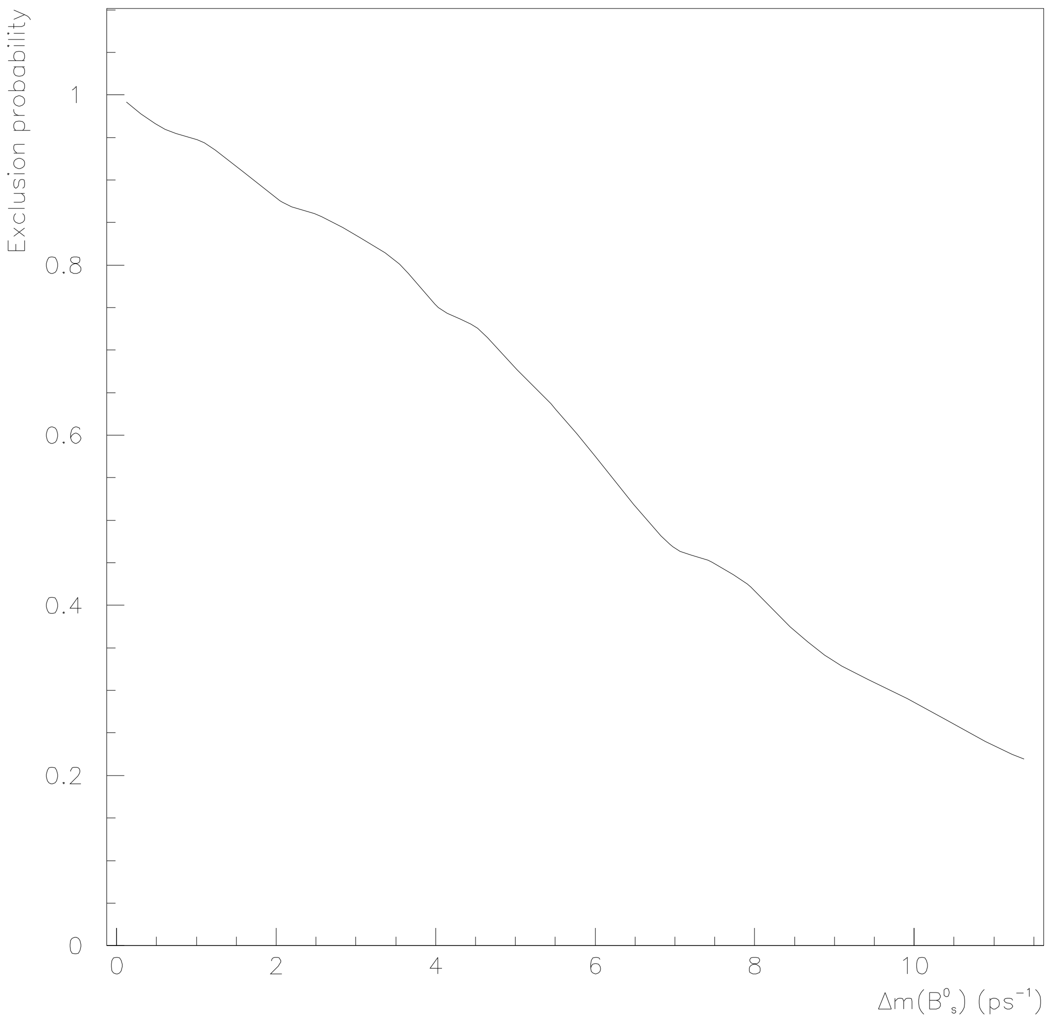


Figure 9.15: *Exclusion probability versus the mass difference* $\Delta m_{B_s^0}$ *for the entire data sample.*

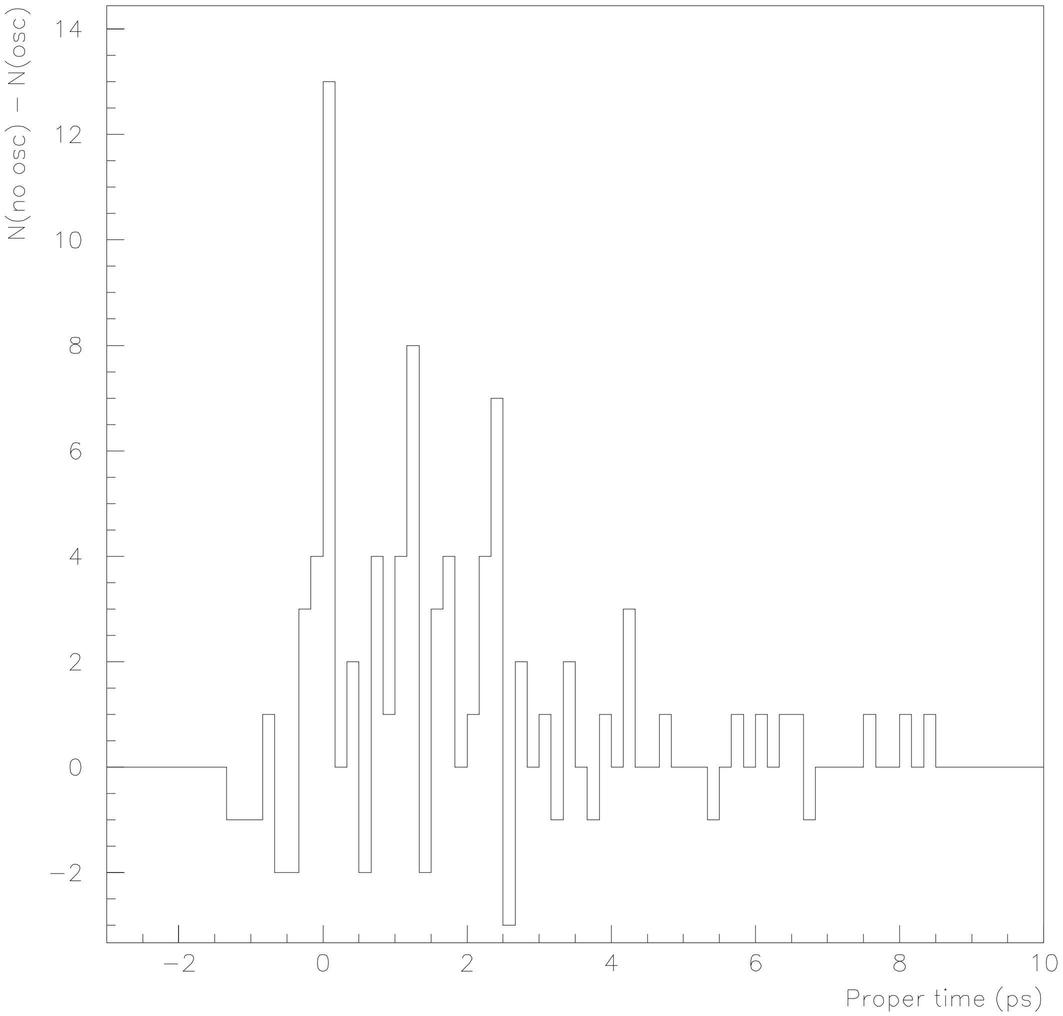


Figure 9.16: *The asymmetry* ($N^{\text{no osc}} - N^{\text{osc}}$) *versus the proper time for the entire data sample.*

Chapter 10

Conclusion

Using four million hadronic Z^0 decays recorded by the DELPHI detector at LEP between 1991 and 1995, B_s^0 mesons have been reconstructed in their semileptonic decays $D_s^{\mp} l^{\pm}$. A high transverse momentum lepton was combined with an oppositely charged D_s meson in the decays $D_s^- \longrightarrow \phi\pi^-$, $K^{*0}K^-$, $\phi\pi^+\pi^-\pi^-$, $\phi\pi^-\pi^0$, $K^{*0}K^{*-}$, $\phi e^-\bar{\nu}_e$ and $\phi\mu^-\bar{\nu}_\mu$. In a total of 392 reconstructed candidates, 197 were found to be B_s^0 signal events.

The advantage of this semi-exclusive analysis is a good proper time distribution and a low background, both combinatorial and from B_d^0 and B^+ mesons. One should also note that the combinatorial background is entirely characterized by the data itself. The B_s^0 lifetime was measured to be

$$\tau(B_s^0) = 1.48\,^{+0.16}_{-0.14}\ \text{(stat) ps}\ ^{+0.06}_{-0.06}\ \text{(syst) ps}$$

Figure 10.1 shows the measured B_s^0 lifetime for different types of DELPHI analyses. As can be seen, the $D_s^{\mp} l^{\pm}$ analysis gives a more precise measurement than both $D_s^{\mp} h^{\pm}$ and the more inclusive ϕl channels.

An upper limit on the lifetime difference between the two physical B_s^0 states was found to be

$$\frac{\Delta\Gamma_{B_s^0}}{\Gamma_{B_s^0}} \leq 56\% \quad @\ 95\% \quad \text{C.L.}$$

This limit takes into account statistical uncertainties only.

Using the amplitude method, the following regions of the mass difference between the physical B_s^0 states are excluded:

$$\begin{aligned} 0 &\leq \Delta m_{B_s^0} \leq 4.4\,\text{ps}^{-1} \quad @\ 95\%\ \text{C.L.} \\ 6.0\,\text{ps}^{-1} &\leq \Delta m_{B_s^0} \leq 7.6\,\text{ps}^{-1} \quad @\ 95\%\ \text{C.L.} \end{aligned}$$

This includes both statistical and systematic uncertainties. The sensitivity was found to be

$$\Delta m_{B_s^0} = 6.7\,\text{ps}^{-1}$$

and the exclusion probability at $\Delta m_{B_s^0} = 7.6\,\text{ps}^{-1}$ is 45%. Figure 10.2 shows the sensitivity and excluded regions for different types of semi-exclusive analyses.

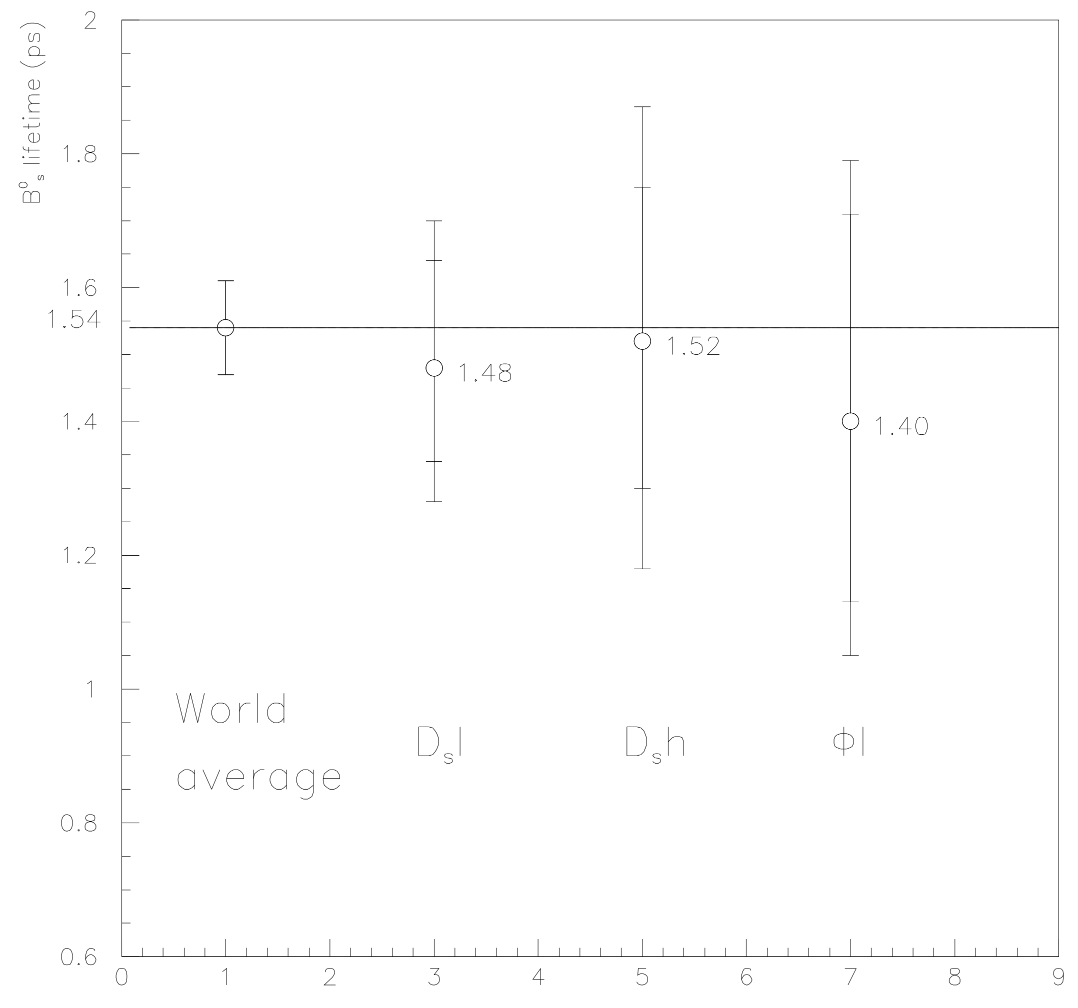


Figure 10.1: *Different B_s^0 lifetime measurements: world average [1], $D_s^{\mp}l^{\pm}$, $D_s^{\mp}h^{\pm}$ [42] and ϕl [56]. For each contribution the central value, the statistical and systematic uncertainties are shown. For the world average the statistical and systematic uncertainties are combined.*

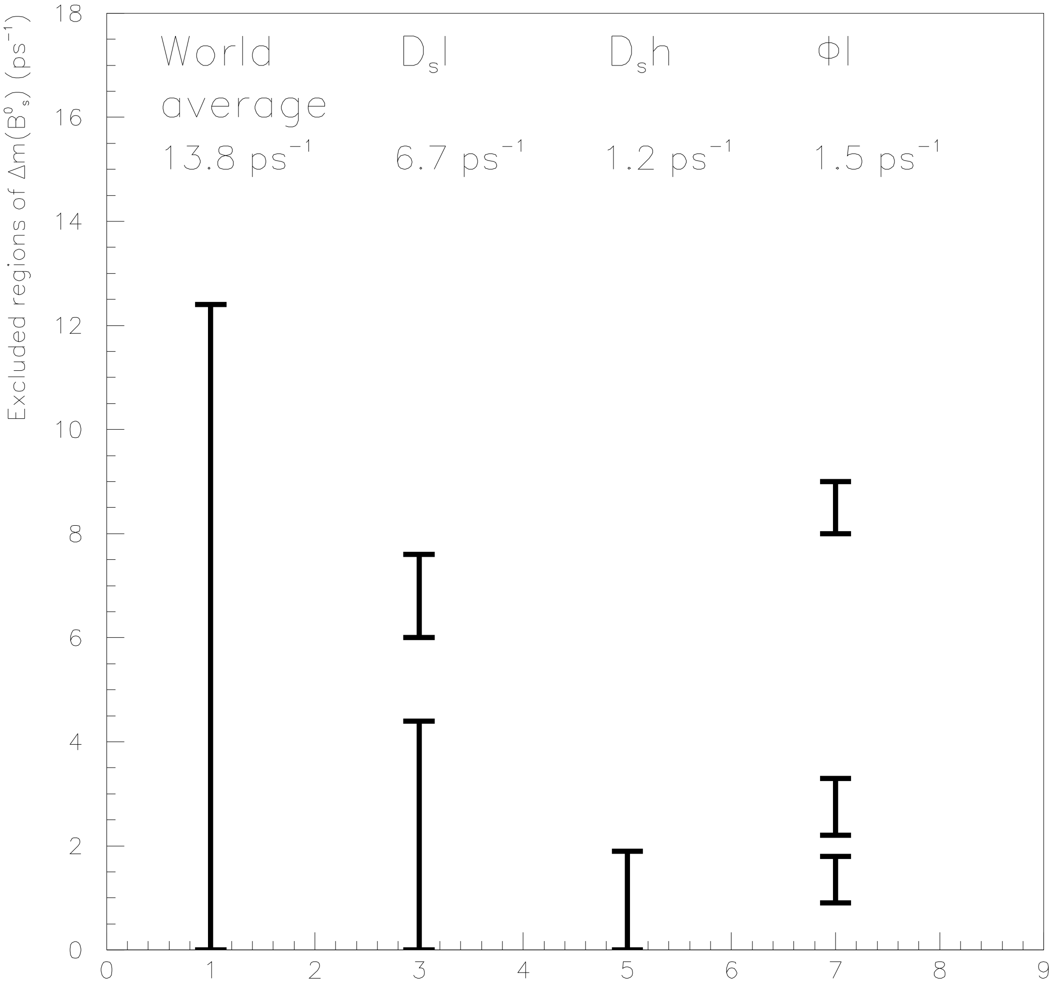


Figure 10.2: *Sensitivity and excluded regions (solid lines) of $\Delta m_{B_s^0}$ for different analyses: LEP Working Group [46], $D_s^{\mp}l^{\pm}$, $D_s^{\mp}h^{\pm}$ [42] and ϕl [56].*

Bibliography

[1] Review of Particle Physics, Particle Data Group, *The European Physical Journal*, Volume 3, Number 1-4, 1998.

[2] Invariante Variationsprobleme, E. Noether, Nachr. d. König, Gesellsch. d. Wiss. zu Göttingen, Math-Phys. Klassc. (1918), 235-257.

[3] *An introduction to Quantum Field Theory*, M. E. Peskin and D. V. Schroder, Addison-Wesley Publishing Company, 1995.

[4] S. L. Glashow, *Nuclear Physics* 22 (1961), 579;
S. Weinberg, *Physical Review Letters* 19 (1967), 1264;
A. Salam, *Elementary Particle Theory*, Ed. N. Svartholm, Almquist and Wiksell, 1968.

[5] Evidence for oscillation of atmospheric neutrinos, The Super-Kamiokande Collaboration, *Physical Review Letters* 81 (1998), 1562-1567.

[6] Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ = 540 GeV, UA1 Collaboration, *Physics Letters,* 122B (1983), 103;
Observation of single isolated electrons of high transverse momentum in events with missing transverse at the CERN $p\bar{p}$ collider, UA2 Collaboration, *Physics Letters,* 122B (1983), 476;
Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider, UA1 Collaboration, *Physics Letters,* 126B (1983), 398;
Evidence for $Z^0 \longrightarrow e^+e^-$ at the CERN $p\bar{p}$ collider, UA2 Collaboration, *Physics Letters,* 129B (1983), 130.

[7] L. Wolfenstein, *Physical Review Letters* 51 (1983), 1945.

[8] Measurement of $|V_{cs}|$ using W-decays at LEP2 (preliminary), DELPHI Collaboration, *DELPHI Internal Note* 98-107 CONF 174, July 1998.

[9] Unitary symmetry and leptonic decays, N. Cabibbo, *Physical Review Letters* 10 (1963), 531.

[10] *CP* violation in the renormalizable theory of weak interaction, M. Kobayashi and K. Maskawa, *Progress of Theoretical Physics* 49 (1973), 652.

[11] S. L. Glashow, J. Iliopoulos and L. Maiani, *Physical Review* D2 (1970), 1585.

[12] Experimental observation of a heavy particle J, J. J. Aubert *et al.*, *Physical Review Letters* 33 (1974), 1404;
Discovery of a narrow resonance in e^+e^- annihilation, J. E. Augustin *et al.*, *Physical Review Letters* 33 (1974), 1406;

[13] *CP* violation, R. Aleksan, Talk given at the *XXI International Meeting on fundamental Pysics*, Miraflores de la Sierra, Madrid, Spain, May 9-15, 1993.

[14] *CP* violation and beauty decays - a case study of high impact, high sensitivity and even high precision physics, I. I. Bigi, Invited lecture at the *First APCTP Hadron Physics Winter School*, February 23-27, 1998, APCTP, Seoul, Korea and UND-HEP-99-BIG01, January 1999.

[15] J. S. Bell, *Proc. Roy. Soc.* A231, (1955), 479;
G. L. Lüders, *Ann. Phys.* 2 (1957), 1;
R. Jost, *Helv. Phys. Acta* 30 (1957), 409.

[16] A. D. Sakharov, Pis'ma Zh. Ekps. Teor. Fiz. 5 (1967), 32.

[17] Heavy-quark physics and *CP* violation, Course by J. D. Richman, University of California, Santa Barbara, 1998 Elsevier Science B.V.

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Physical Review Letters* 13 (1964), 138.

[19] Behavior of neutral particles under charge conjugation, M. Gell-Mann and A. Pais, *Physical Review* 97 (1955), 1387.

[20] L. Wolfenstein, *Physical Review Letters* 13 (1964), 562;
L. Wolfenstein, *Communications in Nuclear Particle Physics* 21 (1994), 275.

[21] CERN NA31 Collaboration, *Physics Letters* 317 (1993), 233.

[22] Fermilab E731 Collaboration, *Physical Review Letters* 70 (1993), 1203.

[23] Fermilab E832 (KTeV) Collaboration, Presented by P. Shawhan at Fermilab, February 1999.

[24] CPLEAR Collaboration, CERN-EP/98-153, *Physics Letters* B444 (1998), 43.

[25] Overall determinations of the CKM matrix, *The BaBar physics book.*

[26] A measurement of $\sin 2\beta$ from $B \longrightarrow J/\psi K_s^0$ with the CDF detector, CDF Collaboration, CDF/PUB/BOTTOM/CDF/4845, Preliminary result for distribution, February 1999.

[27] Introduction to CP violation, in *CP violation*, Ed. C. Jarlskog, World Scientific, Singapore, 1989.

[28] CP violation in B decays and the ATLAS detector at LHC (CERN), A. W. Borgland, *Diploma thesis*, Norwegian Institute of Technology, 1993.

[29] The width difference in the $B_s^0 - \overline{B_s^0}$ system, M. Beneke, G. Buchalla and I. Dunietz, SLAC-PUB-7165, May 1996.

[30] Estimation of $\Delta\Gamma$ for the $B_s - \overline{B_s}$ system. Exclusive decays and the parton model. R. Aleksan, A. Le Yaouanc, L. Olovier and J.-C. Raynal, *Physics Letter* B316 (1993), 567-577.

[31] A top quark story: Quark mixing, CP violation and rare decays in the standard model, A. J. Buras and M. K. Harlander, *Heavy flavours*, edited by A. J. Buras and M. Lindner, World Scientific, Advanced series on directions in high energy physics - Volume 10, 1992.

[32] Heavy quark physics from lattice QCD, J. M. Flynn and C. T. Sachrajda, To appear in *Heavy Flavours* (2nd edition), edited by A J Buras and M Lindner, World Scientific, Singapore, 1998.

[33] An Updated Study of B Meson Oscillations using Dilepton Events, OPAL Collaboration, *Zeitschrift für Physik* C76 (1997), 417-423.

[34] A study of B meson oscillations using hadronic Z^0 decays containing leptons, OPAL Collaboration, *Zeitschrift für Physik* C76 (1997), 401.

[35] Limit on B_s^0 oscillation using a jet charge method, ALEPH Collaboration, *Physics Letters* B356 (1995), 409.

[36] Search for B_s^0 oscillations using inclusive lepton events, ALEPH Collaboration, *European Physical Journal* C7 (1999), 553.

[37] Study of B_s^0 oscillations and lifetime using fully reconstructed D_s^- decays, ALEPH Collaboration, *European Physical Journal* C4 (1998), 367.

[38] Study of the $B_s^0 - \overline{B_s^0}$ oscillation frequency using $D_s^- l^+$ combinations in Z decays, ALEPH Collaboration, *Physics Letters* B377 (1996), 205.

[39] Search for $B_s^0 - \overline{B_s^0}$ oscillations, DELPHI Collaboration, *Physics Letters* B414 (1997), 382.

[40] Study of $B_s^0 - \overline{B_s^0}$ oscillations using inclusive leptons with large P_t, *DELPHI Internal Note* 98-132 ICHEP'98 CONF 193, 1998.

[41] Study of $B_s^0 - \overline{B_s^0}$ oscillations using exclusively reconstructed B_s^0 events, *DELPHI Internal Note* 98-133 ICHEP'98 CONF 1994, 1998.

[42] Search for $B_s^0 - \overline{B_s^0}$ oscillations and measurement of the B_s^0 lifetime, *DELPHI Internal Note* 97-75 Jerusalem CONF 61, 1997.

[43] Study of $B_s^0 - \overline{B_s^0}$ oscillations and B_s^0 lifetime measurement using $D_s^- l^+$ events, *DELPHI Internal Note* 98-131 ICHEP'98 CONF 192, 1998.

[44] A Search for $B_s^0 - \overline{B_s^0}$ oscillations using the semileptonic decay $B_s^0 \longrightarrow \phi l^+ X \nu_l$, CDF Collaboration, FERMILAB-PUB-98/401-E.

[45] Time dependent $B_s^0 - \overline{B_s^0}$ mixing using inclusive and semileptonic B decays at SLD, SLD Collaboration, SLAC-PUB-788, contributed paper 181 to ICHEP 98, Vancouver.

[46] Combined results on B_s^0 oscillations: Results for summer 1998 conferences, The LEP B oscillations working group, LEPBOSC 98/2, September 1998.

[47] Measurement of the B_s^0 meson lifetime using semileptonic decays, CDF Collaboration, *Physical Review* D59 (1999), 032004.

[48] A Study of the width difference in the $B_s^0 - \overline{B_s^0}$ system, ALEPH contribution to ICHEP'98, Vancouver, 1998.

[49] The DELPHI detector at LEP, DELPHI Collaboration, *Nuclear Instruments and Methods in Physics Research* A303 (1991), 233-276.

[50] Performance of the DELPHI detector, DELPHI Collaboration, *Nuclear Instruments and Methods in Physics Research* A378 (1996), 57.

[51] P. Billoir, *Nuclear Instruments and Methods in Physics Research* 225 (1984), 352.

[52] Longitudinal shower development in the DELPHI electromagnetic calorimeter HPC, C. Kreuter, *Diplomarbeit*, University of Karlsruhe (TH), 21. June 1993.

[53] Particle Identification using the RICH detectors based on the RIBMEAN package, M. Battaglia, P.M. Kluit, *DELPHI Internal Note* 96-113 RICH 90, September 1996.

[54] Study of charged kaon production in three-prong tau decays, W. Hao, *PhD thesis*, NIKHEF-H, December 1995.

[55] H.G. Moser and A. Roussarie, *Nuclear Instruments and Methods in Physics Research* A384 (1997), 491.

[56] Decay parameters of the B_s^0 meson in the inclusive ϕl channel. O. Røhne, *PhD thesis*, University of Oslo, August 1998.

[57] PAW - Physics Analysis Workstation: http://wwwcn.cern.ch/asd/paw/index.html

[58] ARGUS Kinematical Analysis Language:
http://www.physik.uni-dortmund.de/ARGUS/Kal/kal.html.

[59] Kinematic Analysis Language - A user manual, A.W. Borgland, G. Eigen, *DELPHI Internal Note* 97-61 PROG 220, June 1997.

[60] Measurement of the charged particle multiplicity of weakly decaying B hadrons, DELPHI Collaboration, *Physics Letters* B425 (1998), 399.

[61] T. Sjőstrand, *Computer Physics Communications* 39 (1986), 347;
T. Sjőstrand and M. Bengtsson, *Computer Physics Communications* 43 (1987), 367.

[62] Tuning and Test of Fragmentation Models Based on Identified Particles and Precision Event Shape Data, DELPHI Collaboration, *Z. Phys.* C73 (1996), 11.

[63] Heavy quark tagging with leptons in the ALEPH detector, ALEPH Collaboration, *Nuclear Instruments and Methods in Physics Research* A346 (1994), 461.

[64] Measurement of $\pi^{\pm}$, $K^{\pm}$, p and $\bar{p}$ production in $Z^0 \longrightarrow q\bar{q}$, $Z^0 \longrightarrow b\bar{b}$ and $Z^0 \longrightarrow u\bar{u}, s\bar{s}, c\bar{c}$ (Particle Identification with the DELPHI Barrel Ring Imaging Cherenkov Counter), E. Schyns, *PhD Thesis*, University of Wuppertal, WUP-DIS 96-22.

[65] *Statistical methods for research workers*, R. A. Fisher, London, 1958.

[66] Observation of the Isospin-Violating decay: $D_s^{*-} \longrightarrow D_s^- \pi^0$, CLNS 95/1346, CLEO 95-10, July 1995.

[67] Vertex routine VDF2ND in DSTANA - DELPHI Analysis Tools Library: http://infodan.in2p3.fr/delphi/analysis/dstana/dstana.html

[68] Measurement of the beam position in DELPHI, *DELPHI Internal Note* 96-6 PHYS 590, 1996.

[69] *MINUIT - Function Minimization and Error Analysis*, Reference Manual, Version 94.1, F. James, CERN.

[70] DELPHI tagging routine: jchcmb.car. Author: G.V. Borisov (DELPHI).

[71] Measurement of $B_d^0 - \overline{B_d^0}$ oscillations, DELPHI Collaboration, *Physics Letters* B398 (1997), 207.

[72] Search for $B_s^0 - \overline{B_s^0}$ oscillations and measurement of the B_s^0 lifetime, *DELPHI Internal Note* 97-75 CONF 61, July 1997.

Appendix A

KAL Script

In an ordinary Fortran based analysis program the user needs to have detailed knowledge of the data structures. In addition, the code is not easily adaptable to other analyses. In Kinematic Analysis Language (KAL) the user does not have to know anything about the underlying data structures. Instead she writes the analysis in the form of a KAL script, using intuitive commands like **identify** π^+, **select** π^+, etc. Default values are given for all standard selection criteria, but can easily be changed by the user via the **cut** command.

This example shows how one would reconstruct the decay $D_s^- \longrightarrow \phi\pi^-$, where $\phi \longrightarrow K^+K^-$, access the generator information and put everything into an Ntuple to be analyzed by PAW [57]. The chosen selection criteria are for pedagogical purposes only.

```
kal on                      ; KAL script to reconstruct the decay Ds --> phi(K+K-) pi
                            ; and put the information in an ntuple

cut kasigric -5             ; Disable particle ID for kaons (i.e. all
                            ; particles will be considered kaons).
cut pisigric 1              ; Want only particles tagged by the RICH
                            ; as pions.

identify k+ k+              ; Load charged tracks into KAL under
                            ; the kaon hypothesis (K+ and K-).
identify pi+ pi+            ; Load charged tracks into KAL under
                            ; the pion hypothesis (pi+ and pi-).

select k+ k-                ; Make K+K- combinations.
  save phi dmass 0.025      ; Save combination as a phi if within 25 MeV
                            ; of the nominal phi mass.
endsel                      ; Defines the end of the K+K- loop.

selcc phi pi+               ; Select phi pi+ and phi pi- combinations.
                            ; 'selcc' means charge conjugated also.
  save ds+                  ; Save combination as Ds+ if phi pi+ and
                            ; Ds- if phi pi-.
```

```
endsel                     ; End of phi pi+/phi pi- loop.

selcc ds+                  ; Access all Ds+ and Ds-.
  dsp = p                  ; Keep Ds momentum and mass in user
  dsm = mass               ; defined variables 'dsp' and 'dsm'.

  seldaug phi              ; Access the phi daughter of the Ds.
    phim  = mass           ; Keep the phi mass and the
    phipx = px             ; x component of phi momentum.

    seldaug k+             ; Access positive kaon daughter of phi.
      krichtag = khaidika ; Keep the RICH kaon tag.
    endsel                 ; End of K+ loop.
  endsel                   ; End of phi loop.

  seldaug pi+              ; Access the pion daughter (pi+ if Ds+,
                           ; pi- if Ds-).
    pivdhit = nasshit      ; Keep number of VD hits for pion.
  endsel                   ; End of pion loop.

                           ; Put everything in an ntuple (yes, KAL
                           ; calls it an 'ntupel'.

  ntupel dsp dsm phim phipx krichtag pivdhit @
  text  'Ds information' @
  tags  'dsp dsm phim phipx krichtag pivdhit'
endsel                     ; End of loop over all Ds+/Ds-.

frame mc                   ; Access MC (generator) information.
  selcc ds+                ; Get Ds+ and Ds-.
    dsp = p                ; Keep Ds momentum.

    seldaug phi            ; Access phi daughter.
      phim = mass          ; Keep phi mass.
    endsel                 ; End of phi loop.

    ntupel dsp phim @             ; Put Ds momentum and
    text  'MC information' @      ; phi mass in ntuple.
    tags  'dsp phim'
  endsel                   ; End of loop over all generator Ds+/Ds-.

endframe                   ; Back to FRAME RECO (default), i.e.
                           ; reconstructed tracks.

return                     ; This returns job control to the usual
                           ; DELPHI program structure
```

Appendix B

Likelihood Normalization

The Maximum log likelihood fit is a powerful fitting method, but one must take great care in the normalization. Let us here give some short explanations of the different likelihoods and the normalizations that are used in this analysis. More formal and mathematical details about the likelihood fit itself can be found in [69].

The starting point of the likelihood fit is that one assumes that the data is distributed according to a known probability distribution. It is only one or several parameters in this distribution that are unknown and needs to be determined by the fit. The problem is thus to find the distribution from the underlying physics processes and then normalize it properly so it becomes a probability density. One example will make this clearer.

The proper times of a particle are distributed according to an exponential distribution. If we were interested in finding the lifetime τ, we would have the likelihood

$$\mathcal{L} = \prod_{\mathrm{i}=1}^{\mathrm{N}} \frac{1}{\tau}\, \mathrm{e}^{\frac{-\mathrm{t_i}}{\tau}} \tag{B.1}$$

i.e. we would calculate the probability for each data point t_i and the total likelihood would just be the product of all the probabilities. The fitted value τ would just be the value of the lifetime that maximizes this product of probabilities. It will be the *most likely* value, thus the name *Maximum likelihood.* This simple example shows two important things. Multiplying lots of small numbers will soon become a numerical problem. Using the *Log* of the likelihood is, therefore, usually preferable. Here 'Log' means the natural logarithm. The second thing is the normalization "constant" $\frac{1}{\tau}$. It is a constant over the data set, but it changes as the fitted value changes, i.e. from one iteration over the data set to another. The reason it is needed is to make the distribution a true probability density and in particular to make sure that the sum of all probabilities are constant and equal to one. If the Log of the likelihood is used, the actual value of the normalization is of no importance. It can be any constant. It must, however, in no case depend on the parameter that one tries to fit. The maximum of $e^{\frac{-t_i}{\tau}}$ is 1 which is obtained for $\tau = \infty$, i.e. independently of the data set the fit would always return the largest possible value for the lifetime. The factor $\frac{1}{\tau}$ is, therefore, needed to "balance" this tendency. More formally one must do[1]:

[1]Here any resolution and acceptance effects are neglected, which means that proper times are observed between 0 and infinity. In a real experimental situation this would be $t_{\min}$ and $t_{\max}$ where $t_{\min}$ can be negative.

$$\mathcal{L} = \prod_{\mathrm{i}=1}^{\mathrm{N}} \frac{\mathrm{e}^{\frac{-\mathrm{t_i}}{\tau}}}{\int_{\mathrm{t}=0}^{\infty} \mathrm{e}^{-\frac{\mathrm{t}}{\tau}} \mathrm{dt}} \tag{B.2}$$

where the integral is over the whole "observation space", i.e. we take the distribution and then normalize it by dividing by the integral.

In the likelihood fit of the decay width difference we want to fit two separate lifetimes τ_1 and τ_2 in the sample. Using equation (B.1), one is then tempted to construct the following likelihood:

$$\mathcal{L} = \prod_{\mathrm{i}=1}^{\mathrm{N}} \left(\frac{1}{\tau_1} \mathrm{e}^{\frac{-\mathrm{t_i}}{\tau_1}} + \frac{1}{\tau_2} \mathrm{e}^{\frac{-\mathrm{t_i}}{\tau_2}} \right) \tag{B.3}$$

However this is wrong. Here each term is normalized independently, while what we must do is the following:

- Find the distribution that the data are distributed according to. Here this is the sum of two exponentials
$$f_1 e^{\frac{-t_i}{\tau_1}} + (1.0 - f_1) e^{\frac{-t_i}{\tau_2}} \tag{B.4}$$
where f_1 is the fraction of CP-even events. The decay $B_s^0 \longrightarrow D_s^- l^+ \nu_l$ contains equal mixtures of the two CP eigenstates, which gives (having dropped the common factor 0.5)
$$e^{\frac{-t_i}{\tau_1}} + e^{\frac{-t_i}{\tau_2}} \tag{B.5}$$
- Then we normalize the distribution by dividing by the integral
$$\int_{t=0}^{\infty} \left(e^{\frac{-t}{\tau_1}} + e^{\frac{-t}{\tau_2}} \right) dt = \tau_1 + \tau_2 \tag{B.6}$$

i.e. the correct likelihood is

$$\mathcal{L} = \prod_{\mathrm{i}=1}^{\mathrm{N}} \frac{1}{\tau_1 + \tau_2} \left(\mathrm{e}^{\frac{-\mathrm{t_i}}{\tau_1}} + \mathrm{e}^{\frac{-\mathrm{t_i}}{\tau_2}} \right) \tag{B.7}$$

The main difference with respect to the previous one is that here both *exponentials* have the *same weight*, while in the other case they were weighted accordingly to their respective lifetime. Besides giving the wrong overall normalization, this would have the additional effect of introducing different fractions of the two CP eigenstates.

In the case of flavour oscillations this has a surprising effect. The un-normalized likelihood is just

$$\mathcal{L} = \prod_{i=1}^{N(\mathrm{mix})} P_{\mathrm{i}}^{\mathrm{mix}} \prod_{i=1}^{N(\mathrm{unmix})} P_{\mathrm{i}}^{\mathrm{unmix}}$$

where P^{mix} and P^{unmix} are the un-normalized (probability) densities for oscillating and non-oscillating events, respectively. The normalization "integral" should be over the whole observation space, i.e. one must divide by the integral of the sum of oscillating and non-oscillating contributions:

$$\int_{t=0}^{\infty} \left(P^{\mathrm{mix}} + P^{\mathrm{unmix}}\right) dt \tag{B.8}$$

Since the oscillations cancel in the sum, the normalization in this case do not depend on $\Delta m_{B_s^0}$.

Appendix C

Conference Contributions

Three of the modes of this analysis:

$$D_s^- \longrightarrow \phi\pi^+\pi^-\pi^-$$

$$D_s^- \longrightarrow K^{*0}K^{*-}$$

$$D_s^- \longrightarrow \phi\pi^-\pi^0$$

have been merged with another $D_s^{\mp}l^{\pm}$ analysis in DELPHI. This common B_s^0 lifetime, lifetime difference and mixing analysis has been presented at the *International Europhysics conference on high-energy physics (HEP 97)* in Jerusalem, 19-26 August 1997 [72], and HEP 98 in Vancouver, July 23-29, 1998 [43]. The two articles are written by the DELPHI *B lifetimes / oscillation team.* The parts in section 3 and figures 4 and 5 (HEP 97) and section 4 and figure 2 (HEP 98) concerning the three above-mentioned modes are the author's work.

Erratum

Equation (2.10) on page 7 should be:

$$j^{\mu}(x) = \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\phi)}\Delta\phi$$

Equation (2.26) on page 11 should be:

$$\psi(x) \longrightarrow \psi^{'}(x) = e^{ig'Yf(x)}\,\psi(x)$$

Equation (2.101) on page 31 should be:

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix}$$

Page 51: ... it would mean, with a high probability, that it was a kaon or a proton.