CERN-THESIS-2002-070

Inklusive Analyse der Oszillationen neutraler B-Mesonen mit dem DELPHI-Detektor bei LEP

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik der Universität
Karlsruhe

genehmigte

DISSERTATION

von

Dipl. phys. Thomas Allmendinger
aus Vaihingen/Enz

Tag der mündlichen Prüfung 3.5.2002
Referent: Prof. Dr. M. Feindt, Institut für Experimentelle Kernphysik
Korreferent: Prof. Dr. G. Quast, Institut für Experimentelle Kernphysik

Abstract

Transitions between particles and antiparticles, so-called 'oscillations', were first observed in the neutral Kaon system in 1956 [L$^+$56] ($\bar{s}d \rightleftharpoons s\bar{d}$). Oscillations were also observed in the neutral B_d^0 meson system in 1987 ($\bar{b}d \rightleftharpoons b\bar{d}$) [A$^+$87]. The time dependence of the B_d^0 oscillations was resolved in later experiments at LEP [Abr93, B$^+$93, A$^+$96c, G$^+$96]. Many efforts were made during the years resolving the much faster oscillations in the B_s meson system. The latest world average for the B_s oscillation frequency is $\Delta m_s > 14.9\,\mathrm{ps}^{-1}$ at 95 % CL compared to the value from the B_d^0 system $\Delta m_d = 0.496 \pm 0.007\,\mathrm{ps}^{-1}$ [OG01].

In a novel approach using the data set of the DELPHI detector taken at the Z^0 pole from 1992 to 2000 a fully inclusive analysis was performed. The main aim of the analysis unlike previous analyses was to use the large statistics of fully inclusive b decays. This was only achieved by exploiting the full capabilities of the DELPHI detector and the application of advanced analysis techniques, such as neural networks and quasi-random point sets used for resolution integrations.

After all cuts including the removal of all high p_T lepton events, which are treated in a separate analysis [KPP$^+$01], a sample of 408k events was available. Dedicated production and decay tag algorithms, momentum and decay length reconstruction algorithms were developed, partly based on the BSAURUS package [ABF$^+$01].

The result was extracted using a maximum likelihood fit and the following value for the mass difference of the two physical B_d^0 states was obtained:

$$\Delta m_d = 0.501 \pm 0.013\,(\mathrm{stat}) \pm 0.018\,(\mathrm{sys})\ \mathrm{ps}^{-1}$$

The total error was therefore $0.022\,\mathrm{ps}^{-1}$. Systematic effects were mostly compensated by additional free fitting parameters.

Utilising the 'Amplitude Method' [HGA97] a likelihood fit searching for B_s oscillations was performed on the same data set. The time dependence of the B_s oscillations was not resolved but a limit on the mass difference of the two physical B_s states was set at:

$$\Delta m_s > 4.7\,ps^{-1} \quad \text{at 95 \% CL.}$$

Using the error on the amplitude one can extract the expected limit ('Sensitivity') on Δm_s at 95% CL:

$$\mathrm{Sensitivity} = 7.4\,\mathrm{ps}^{-1}$$

Contents

Introduction

I may not have gone where I intended to go,
but I think I have ended up where I intended to be.
Douglas Adams

Heavy flavour physics, especially B physics, has become more and more important for tests of the so-called 'Standard Model' and measurements of its parameters. In particular, B physics offers the most direct way to determine the parameters of the quark mixing matrix. This matrix connects the measured mass eigenstates to the flavour eigenstates, which couple to the weak interaction. As this matrix has non-diagonal entries transitions between different quark families are possible, e.g. b decays: $b \to c$. Determining the matrix elements of the so-called 'Cabbibo-Kobayashi-Maskawa (CKM) Matrix' [KM73,Cab63] is an important task in modern particle physics, not least because the matrix is a source of CP violation. After the first observation of CP violation in the neutral kaon sector in 1964 [CCFT64], the observation of CP violation in the B sector has now also been experimentally established by BABAR [A$^+$01b, Pre01, Wea01], BELLE [Haz, A$^+$01a] and CDF [Blo].

Transitions between particles and antiparticles, so-called 'oscillations' or 'particle mixing', were first observed in the neutral Kaon system in 1956 (quark content: $\bar{s}d \rightleftharpoons s\bar{d}$) [L$^+$56]. Oscillations were also observed in the neutral B_d^0 meson system in 1987 ($\bar{b}d \rightleftharpoons b\bar{d}$) [A$^+$87]. A first time dependent measurement of the B_d^0 meson oscillations was possible at LEP in 1993 after the development of high resolution silicon vertex detectors [Abr93, B$^+$93]. Oscillations in the neutral B_s meson ($\bar{b}s \rightleftharpoons b\bar{s}$) system were also observed at LEP [Mos93, For93], although their time dependence was not resolved. The ratio of mixed to unmixed B_s mesons was measured to be $\chi_s = 0.46 \pm 0.11$. All oscillations are driven by the mass difference Δm of the participating states. The recent world averages for the oscillation frequencies in the B meson sector are $\Delta m_d = 0.496 \pm 0.007\,\mathrm{ps}^{-1}$ for the B_d^0 meson system and $\Delta m_s > 14.9\mathrm{ps}^{-1}$ at 95 % CL for the B_s system [OG01]. Clearly B_s oscillations are expected to be much faster than B_d^0 oscillations and thus represent an enormous experimental challenge. While the contribution to the CKM matrix

parameters is nearly the same in both B meson cases, a measurement of the B_s frequency would allow a dramatic reduction of the errors from theoretical uncertainties [Bur01]. Therefore many efforts have been made over the years to resolve the faster B_s oscillations. A more detailed introduction to the theory of B oscillations can be found in chapter one.

Various B meson oscillation analyses were made at the DELPHI experiment at LEP [A+96a,Lip] most of them based on an exclusive reconstruction of single decay chains or on B decays which contain a lepton with high transverse momentum. In common to all these analyses is the statistical limitation imposed by ignoring most of the hadronic decay modes and thus using only a few percent of the available statistics. Over the years the understanding of the detector improved together with the quality of the simulated events. These conditions allowed the development of inclusive techniques for the approach to B physics, where the shape of more general variables is used, i.e jet-charge and vertex charge. In the Karlsruhe working group a lot of inclusive tools for B physics are collected in the BSAURUS [ABF+01] package, which provides the basic inputs to this analysis. The package makes extensive use of all parts of the DELPHI experiment especially of its unique particle identification capabilities, together with modern methods of statistical data analysis, e.g. neural networks. Chapter two gives a brief description of the DELPHI detector and its components. The first half of chapter three is dedicated to neural networks - as they are the preferred method used inside BSAURUS.

The inclusive approach of the analysis presented in this work allows the recovery of the large statistics of hadronic B decays for an inclusive B meson oscillation analysis based on the data in the years 1992 to 1995 (LEP I phase) and the Z^0 data from LEP II taken in the years 1996 to 2000. The result is a new, independent contribution to B_s oscillation measurements and additionally a very precise measurement of B_d^0 oscillation frequency. Also novel to this kind of analysis, is the treatment of the proper time reconstruction necessary to resolve the time dependent oscillations. The B meson decay time is reconstructed from the decay length and the momentum of the B meson. Both quantities are reconstructed in the detector with limited resolution. The fitting method, which is later applied for the extraction of the results, takes this finite resolution effects into account by convolution of the resolution functions with the theoretical expectations. Therefore a numerical integration over these resolution functions is necessary. The statistical power of the analysis is improved by the separation of well reconstructed events with good proper time resolution from others, where the resolution is not as good. In previous analyses the measurement of the decay length and the B meson momentum was combined and the classification was done in terms of

the proper decay time, where e.g. five different resolution classes were chosen. In this analysis the resolutions for decay length and momentum were kept separate in all routines, which improves the analysis granularity and thus its statistical power. The analysis consists of 15 classes for the momentum and 5 for the decay length resolution of the B meson, which leads to 75 possible different combinations. The consequence of this improvement is a two-dimensional integration, compared to a one-dimensional integration for the standard approach. The standard approaches for numerical integration (e.g. trapezoid rule) are found to be too slow to be of practical use when large statistics are involved. Therefore a two step solution towards a very fast and efficient integration routine was developed.

A first improvement was found in the point set covering the two-dimensional integration space of the numerical integration. Although one might think that a grid of points is the best choice one can make a special kind of random number, so-called 'Quasi Random Numbers', provide smaller errors for the same amount of points used for the integration. The reason for this can be found in the 'smoothness' of the point set, which is better for quasi random numbers. Another improvement of the integration routine is derived from the knowledge of the resolution functions. These functions determine largely the shape of the integrand, which leads to the possibility of using a so-called 'importance sampling method'. The aim of the method is to increase the density of points in phase space regions where the value of the integrand is expected to be large. The importance sampling method is realized in the form of an analytic variable transformation, since the shape of the resolution functions are known. A detailed description of the integration routine is provided in the second part of chapter three.

Chapter four briefly describes the basic properties of the BSAURUS package, which are important for this analysis. The reconstruction of the secondary vertex and the B meson momentum are described, which are the essential variables for the proper time measurement. The algorithms for tagging events as 'mixed' or 'unmixed' by their b-quark flavour (b or $\bar{b}$) at production and decay time of the B meson are explained. Several neural networks starting at the track level were developed and finally combined to form the desired tagging quantities for the production and the decay of the B meson.

In chapter five, the analysis itself is presented starting with a description of the event selection. The analysis uses data collected on the Z^0 in the LEP I phase from 1992 to 1995, and additionally the calibration data of the LEP II phase, also taken on the Z^0. Events containing an identified lepton with high transversal momentum are removed, which reduces the statistical overlap with other existing B meson oscillation analyses [Lip,A$^+$96a,B$^+$]. Finally

over 400000 candidate events were selected, which is approximately ten times the statistics of the analysis, based on semi-leptonic B decays [KPP$^+$01]. In combination with a second inclusive vertices B_s analysis [P$^+$02, TKPA01], the combined number for the DELPHI experiment is $\Delta m_s > 14.5\,\mathrm{ps}^{-1}$ (Sensitivity: $12.0\,\mathrm{ps}^{-1}$) [KPP$^+$01, TKPA01, P$^+$00].

The chapter also contains descriptions of the algorithms for separating the different classes of decay length and momentum resolution, as well as the calibration of the flavour tagging algorithms on the data. This last point is very important for the correct modelling of the data inside the analysis. Finally, after discussion of the fitting methods, which are slightly different between the B_d^0 ('Likelihood Approach') and the B_s case ('Amplitude Approach'), the final results are presented, together with a breakdown of the systematic effects.

Chapter 1

Theoretical Background

Particle physics studies the constituents of matter and their interactions. Present knowledge has built up from many years of experimental discoveries and theoretical developments. From these discoveries emerged a simple model of matter, the so-called 'Standard Model'. There are two kinds of fundamental, structureless, spin $\frac{1}{2}$ fermions - quarks and leptons. In total there are six known types ('flavours') of quarks, three charged leptons and three neutral leptons ('neutrinos'). There are four forces governing the universe: the strong force binds quarks to form hadrons and nucleons to form nuclei, the electromagnetic force acts on all charged particles, the weak force is responsible for processes like beta decay and the gravitational force act on all bodies with mass. Classically, forces are described in terms of fields through which particles interact. In quantum theory the equivalent view describes the force as being transmitted by the exchange of field quanta ('gauge bosons'), of integral spin. The first part of this chapter provides a brief introduction to the Standard Model, listing the basic properties of the fundamental fermions and interactions.

The second part of the chapter describes in more detail the phenomena of transitions between particle and antiparticles, so-called 'oscillations' or 'mixing'. They were predicted in 1955 [GMP55] and observed in 1956 in the $K^0\bar{K}^0$ system [L$^+$56]. The quark content of a K^0 is $d\bar{s}$ and $\bar{d}s$ for its antiparticle $\bar{K}^0$. Transitions between the K^0 and $\bar{K}^0$ have $\Delta S = 2$ and proceed via a second order weak interaction, not conserving the strangeness S of the system. Oscillations were also observed in 1987 in the $B^0_d\bar{B}^0_d$ system [A$^+$87], where the B^0 has quark content $\bar{b}d$ and $b\bar{d}$ for the $\bar{B}^0_d$. The last neutral $q\bar{q}'$ system where oscillations have been established is the B_s ($=\bar{b}s$) system [Mos93, For93]. All the particle to antiparticle transitions are driven

by the mass difference ΔM of the two states[1], which is directly proportional to the rate or oscillation frequency of the transitions. The whole phenomena of quark mixing is well described within the Cabbibo-Kobayashi-Maskawa (CKM) matrix formalism [KM73, Cab63]. A brief introduction to the CKM formalism and the mathematical treatment of particle to antiparticle transitions is given in section 1.2.

Simulated events are crucial for an inclusive analysis and thus modelling the process $e^+e^- \rightarrow$ hadrons is essential. Therefore a short description of the simulation chain from a theoretical point of view is provided in the last part of the chapter.

Fopr further details on the individual subjects, more extensive reviews can be found in [Per, Kan, Sto94]

1.1 The Standard Model

The Standard Model of particle physics is based on gauge theories. The elementary particles are fermions - quarks and leptons. These fermions can be ordered into 3 generations, all of them containing 2 quarks and 2 leptons. The main difference between the individual generations are the masses of the individual particles. The quantum numbers Q^{em} (electromagnetic charge), Y (weak hypercharge), I^W (weak isospin) and I_3^W (third component of the weak isospin) of the 3 generations are the same (see table 1.1).

A brief summary of the interactions, the Standard Model is based on, is given in the following subsections.

1.1.1 Quantum Electro Dynamics (QED)

Quantum Electro Dynamics (QED) is the gauge theory describing electromagnetic interactions between electric charged fermions via the exchange of a single gauge boson, the photon γ. The photon γ has no electric charge and so does not self interact. This fact is represented in the Abelian character of the underlying gauge group $U(1)$. The coupling strength is proportional to the fermion electric charge, and this charge is a conserved quantity in all QED processes. Both properties can be deduced from the $U(1)$ gauge theory character. In figure 1.1 (a) the basic process $e^+e^- \rightarrow \mu^+\mu^-$ is illustrated. [Fey49].

Higher order corrections to the Born level diagrams performed in perturbation theory lead to closed-loop processes, shown in the diagrams 1.1 (b)

[1]The decay width difference $\Delta\Gamma$ of the two states also influences the oscillations (i.e. Kaons K_s, K_L). In the case of B meson their contribution is supressed.

Generation			Quantum Numbers			
I.	II.	III.	Q^{em}	Y	I^W	I_3^W
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$\begin{matrix} 0 \\ -1 \end{matrix}$	-1	$+\frac{1}{2}$	$\begin{matrix} +\frac{1}{2} \\ -\frac{1}{2} \end{matrix}$
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$\begin{matrix} +\frac{2}{3} \\ -\frac{1}{3} \end{matrix}$	$+\frac{1}{3}$	$+\frac{1}{2}$	$\begin{matrix} +\frac{1}{2} \\ -\frac{1}{2} \end{matrix}$

Table 1.1: The different elementary fermions in the Standard Model. Quarks and leptons are separated into 3 generations on the left, the associated quantum numbers are stated on the right. Doublets are marked with a L to denote, that only the left-handed components take part in the weak interaction. The prime of the lower quark doublets members denote the weak eigenstate (see section 1.1.2). The different quantum numbers are, Q^{em} the electromagnetic charge, Y the weak hypercharge, I^W the weak isospin and I_3^W the third component of the weak isospin.

and (c). A priori these diagrams lead to divergent integrals, with an infinite contribution to any amplitude calculation beyond the Born level. The solution to this problem, the redefinition of the electric charges and masses to their physical values, is called *renormalisation* [Lee69]. As a consequence of this technique all amplitude and cross-section calculations are finite. An impressive proof of this approach is the $g-2$ factor, where agreement between experiment and theory is of the order of 1 part in 10 billion [CM99, B$^+$01].

The success of the QED higher order corrections being finite, due to renormalisation, and the calculation of physical predictions with high precision, leads to the requirement that a candidate theory for particle interactions has to be renormalisable, otherwise the perturbative expansion approach would fail. In 1971 it was shown by 't Hooft [tH71] that all gauge theories are renormalisable, which marks gauge theories as a preferred starting point.

1.1.2 The Electroweak Interaction

The gauge group of the weak interaction is the $SU(2)_L$ group and the $U(1)$ group for the electromagnetic interaction (see section 1.1.1). A separate treatment of the two gauge theories is not possible, because leptons in the $SU(2)_L$ doublet have different electromagnetic charges. This leads to a theory, which is invariant with respect to a $SU(2) \times U(1)$ gauge transformation.

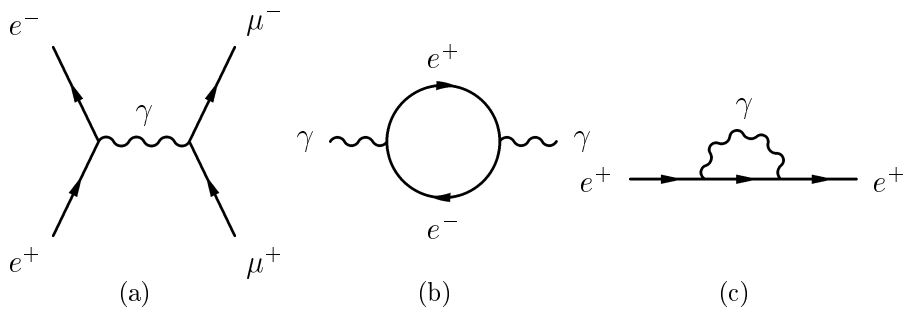

Figure 1.1: (a) The lowest order (or Born level) QED Feynmann diagram. Two different QED loop diagrams, the creation and re-annihilation of an e^+e^- pair (b), and the emission and absorption of a photon by an electron (c).

The unification of the electromagnetic and weak interaction was first performed by Glashow, Salam and Weinberg (GWS) [Gla61, Sal, Wei67]

The mediators of the non-abelian $SU(2)_L$ group are the gauge bosons W^1, W^2 and W^3. The gauge boson of the $U(1)_Y$[2] group is the B. The real physical mediators of the weak charged currents are the $W^\pm$ bosons, a linear combination,

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2) \tag{1.1}$$

of the W^1, W^2 gauge bosons. The gauge boson fields of the photon (γ) and the Z^0 can be obtained via mixing in the following way,

$$\begin{aligned} A_\mu = &\quad \cos\Theta_w \cdot B_\mu &&+ \sin\Theta_w \cdot W^3_\mu \\ Z_\mu = &- \sin\Theta_w \cdot B_\mu &&+ \cos\Theta_w \cdot W^3_\mu \end{aligned} \tag{1.2}$$

introducing a single parameter Θ_w known as the weak mixing angle or Weinberg angle.

The electroweak theory, as described in the section above, does not contain any mass terms in its Lagrangian either for the fermions nor for the gauge bosons. Any introduction of mass terms would destroy the local gauge symmetry and as a consequence the possibility of renormalising the theory.

The favoured solution of this problem is the introduction of a new complex scalar field Φ, termed the Higgs field [Hig64]. It allows the assignment of masses to the $W^\pm$ and Z bosons by coupling them to the Higgs field Φ, conserving the gauge invariance of the whole system. The fermion masses

[2] Y denotes the weak hypercharge which is connected via the Gell-Mann-Nishijama formula $Q^{em} = I_3^W + \frac{1}{2}Y$

are achieved in a similar way by introducing Yukawa couplings into the Lagrangian. As a consequence of the Higgs mechanism a neutral physical boson H^0 with unknown mass is predicted[3].

1.1.3 Quantum Chromo Dynamics (QCD)

The gauge theory of the strong interaction (QCD) [FGML73] is based on the non-Abelian gauge group $SU(3)_C$, introducing three colour degrees of freedom (called red, green and blue). The colour degree of freedom was introduced to resolve the so called spin statistics problem with some baryonic states, such as the Ω^-, which contains three identical s quarks in a symmetric spin state. As a solution to the conservation of the Pauli principle each quark was assigned a different colour resulting in a colour singlet state. All experimentally observed states are colour singlets. Baryons are built on colour-singlets of red, green and blue and mesons of colour and anti-colour combinations.

The gauge invariance under $SU(3)_C$ transformations leads to eight different gauge fields, termed gluon fields. As a consequence of the non-abelian character of the $SU(3)_C$ group, in agreement with experimental results, the gluons carry colour themselves leading to self-interaction, illustrated in figure 1.2 (b) and (c). Self-interactions are responsible for the confinement and the asymptotic freedom of quark states.

Confinement is the explanation of the fact that only colour singlets can be stable. If one tries to remove a single quark out of a colour singlet, the amount of energy needed is that high, that finally a new quark-antiquark is created, leading to two colour singlets. The reason for this behaviour is

[3]The current experimental lower limit of the Standard Model Higgs boson from the LEP2 collaborations is 114 GeV [Gro01]

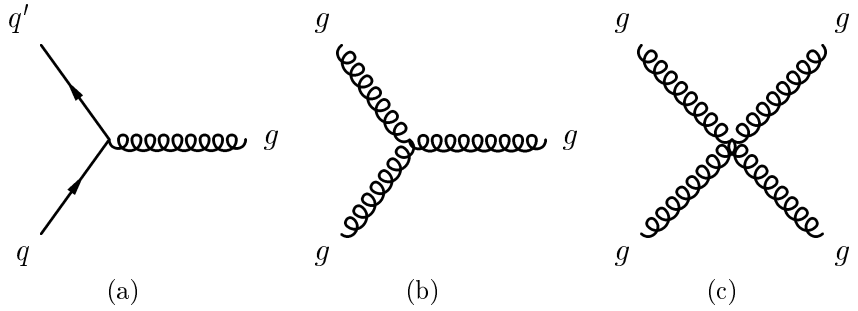

Figure 1.2: The fundamental vertices in QCD are (a) the quark-gluon-coupling b) three-gluon-coupling c) four-gluon-coupling

the self-interaction of the gluons, filling the space between two quarks with a gluon field, the strength of which is proportional to the quark-antiquark distance.

The opposite behaviour is observed when going to very small quark-quark distances. The coupling becomes smaller and smaller for shorter distances resulting in a quasi-free state. This phenomenon is called asymptotic freedom and is an important requirement for heavy quark theories, discussed later in section 1.3.4

1.2 The Cabbibo Kobayashi Maskawa (CKM) Matrix

Based on experimental observations it became clear in the early sixties, that the quark sector for the weak-charged current did not behave as simple as the lepton sector. Up till then the only known quarks where the u, d and s quarks. The transition strength for $e \to \nu_e$ and $\mu \to \nu_\mu$ are the same, while the decay rate difference for $\pi^- \to \mu\nu_\mu$ and $K^- \to \mu\nu_\mu$ transitions is a factor of approximately twenty. An explanation, proposed by Cabibbo [Cab63], was the introduction of a weak eigenstate d', which is a superposition of the mass eigenstates d and s in the following way:

$$d' = d\cos\Theta_C + s\sin\Theta_C \tag{1.3}$$

The introduced mixing angle is called Cabbibo angle, preserving universality of the weak interaction by the existence of only one coupling constant.

A problem of the Cabbibo modifications was the appearance of Flavour Changing Neutral Currents (FCNC), which were experimentally not ob-

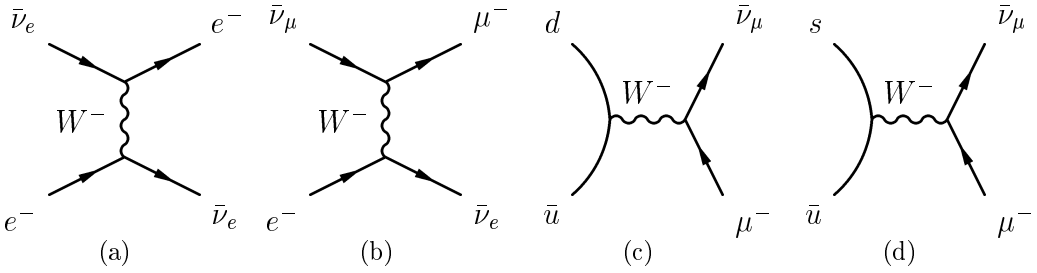

Figure 1.3: The processes (a) and (b) illustrate neutrino scattering. They have the same cross section. The decay rate of the π^- ($\bar{u}d$) (c) is approximately twenty times higher than the K^- ($\bar{u}s$) rate (d). An explanation of this behaviour is given by the Cabibbo theory.

served. The solution to this problem was the modification of the s state into the weak eigenstate s',

$$s' = s\cos\Theta_C - d\sin\Theta_C. \tag{1.4}$$

The demand of equivalence of the coupling of the s' and d' states to the Z^0, leads to the introduction of a new doublet for the s' state and the proposal of a new unknown quark state c (charm). This solution of solving the FCNC problem by the introduction of weak eigenstates together with the charm hypothesis is called the GIM (Glashow, Iliopoulos and Maiani) mechanism [GIM70]. The charm quark was established experimentally in the mid 1970's by the discovery of $c\bar{c}$ states [A$^+$74].

In a later development Kobayashi and Maskawa [KM73] proposed the extension of the four-quark pattern to a six-quark pattern to accommodate the possibility of CP violation including also the possibility of a third quark generation. Evidence for the fifth quark was not long in coming, with the discovery of $b\bar{b}$ states [H$^+$77, B$^+$81].

The interaction Lagrangian $\mathcal{L}$ for the weak charged current is now written in terms of mass eigenstates as follows,

$$\mathcal{L}_{int} = -\frac{g}{\sqrt{2}}(\bar{u}_L, \bar{c}_L, \bar{t}_L)\gamma^\mu V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_\mu^\dagger + h.c., \tag{1.5}$$

introducing the CKM matrix V_{CKM}, which has the form:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.6}$$

The CKM matrix is in general a 3×3 matrix consisting of 9 complex numbers, thus leading to 18 real parameters. The requirement of unitarity ($V^\dagger V = 1$) provides 9 constraints, so that only 9 real parameters remain. By an appropriate redefinition of relative quark phases another 5 parameters can be removed, ending up with 4 real parameters. Of these, 3 correspond to angles (the number of independent rotations in three-dimensional space) and a complex phase, a potential source of CP violation (Asymmetries in the decays of particles and antiparticles).

1.2.1 The Wolfenstein Parametrisation

For phenomenological applications it is more convenient to use an approximate parametrisation of the CKM matrix, which makes explicit the strong

hierarchy observed experimentally[4]. The favoured parametrisation, proposed by Wolfenstein [Wol83], has the parameters A, ρ, η and λ. An expansion in powers of λ leads to the following representation:

$$V_{CKM} \simeq \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.7}$$

In practice, $|V_{ud}|$ and $|V_{us}|$ are known to an accuracy of better than 1 %, $|V_{cb}|$ is known to 5 %. Hence, the Wolfenstein parameters λ and A are rather well determined experimentally:

$$\lambda = |V_{us}| = 0.2205 \pm 0.0018,\ A = \left|\frac{V_{cb}}{V_{us}^2}\right| = 0.80 \pm 0.04. \tag{1.8}$$

The experimental uncertainties on ρ and η are at the 20 % level.

A beautiful way to visualise the implications of unitarity is provided by the so-called unitarity triangles, using the fact that the unitarity equations

$$V_{ij}V_{ik}^* = 0\ (j \neq k) \tag{1.9}$$

can be represented as the equation of a closed triangle in the complex plane. Of the six possible triangles, the most useful from the phenomenological point of view is the relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,, \tag{1.10}$$

since it is the triangle expected to have large angles [HQ]. In the standard form, $V_{cd}V_{cb}^*$ is real, and the unitarity triangle has the form shown in figure 1.4 (a). It is useful to rescale the triangle by dividing all sides by $V_{cd}V_{cb}^*$. The rescaled triangle has the coordinates $(0,0)$, $(1,0)$, and $(\bar{\rho},\bar{\eta})$, where

$$\bar{\rho} = \left(1-\frac{\lambda^2}{2}\right)\rho,\ \bar{\eta} = \left(1-\frac{\lambda^2}{2}\right)\eta \tag{1.11}$$

are related to the Wolfenstein parameters ρ and η (see figure 1.4 (b)).

Unitarity amounts to the state that the triangle is closed, and CP is violated when the area of the triangle does not vanish, i.e. when the angles are different from zero or 180^0.

[4]The diagonal elements are close to unity, V_{us} and V_{cd} are of the order 0.2, V_{ts} and V_{cb} are around 0.04

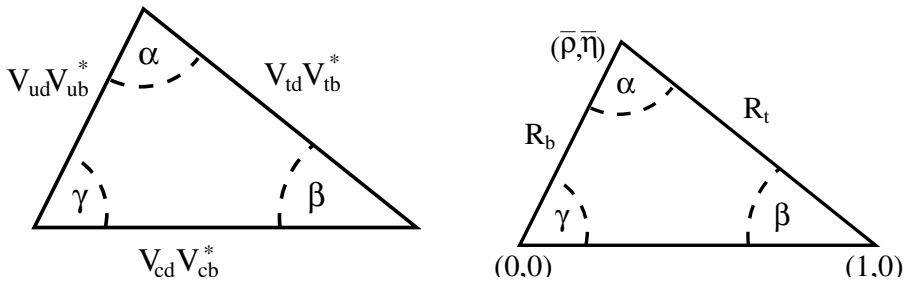

Figure 1.4: (a) The unitarity triangle in the standard form, and (b) in its rescaled form in the $\bar{\rho} - \bar{\eta}$ plane.

1.2.2 The Experimental Status

Determination of the shape of the triangle with high precision is one of the major tasks of particle physics of the recent years. Available recent measurements are e.g. the angle $\sin 2\beta$ from BELLE [A$^+$01a, Haz], BABAR [A$^+$01b, Wea01] and CDF [A$^+$00] via the process $B_d^0 \to J/\psi K_s^0$. The side R_b and its relation to V_{ub} and V_{cb} comes from various B decay measurements at LEP [Haw01] and CLEO [Zha01] and the ratio ϵ'/ϵ is measured by NA48 [Tat01] and KTEV [Gla].

To determine R_t, one needs information on $|V_{td}|$, which can be extracted from neutral $B^0 - \bar{B}^0$ mixing. As the measurement in this thesis is about B meson mixing a quick theoretical review on the topic is given in the next section.

A recent scan in the $(\bar{\rho}, \bar{\eta})$ plane [C$^+$01] based on the best knowledge of the experimental and theoretical inputs with a method called "*95 % C.L. scanning*" [HQ] leads to the following result:

$$\bar{\rho} = 0.224 \pm 0.038, \qquad \bar{\eta} = 0.317 \pm 0.040 \tag{1.12}$$

Figure 1.5 illustrates the different experimental inputs determing the shape of the unitarity triangle, together with the result of the fit based on currently available measurements.

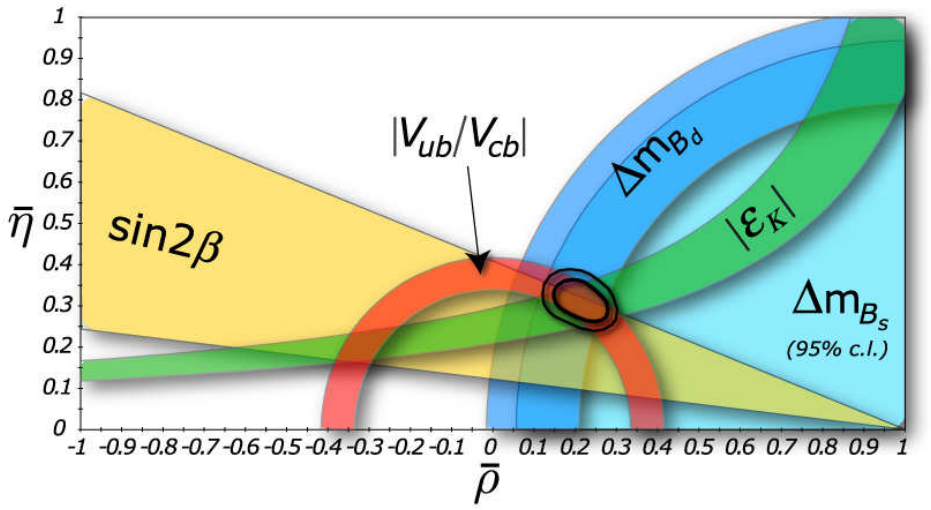

Figure 1.5: The experimental status of the CKM triangle. One can see the different experimental constraints leading to the best fit result in the $(\bar{\rho}, \bar{\eta})$ plane with $\bar{\rho} = 0.224 \pm 0.038$ and $\bar{\eta} = 0.317 \pm 0.040$.

1.2.3 Mixing of $B^0 - \bar{B}^0$ Mesons

The $b\bar{q}$[5] flavour eigenstates can mix via the box diagrams illustrated in figure 1.6. The mass eigenstates are denoted by

$$|B_H\rangle = p|B^0\rangle + q|\bar{B}^0\rangle, \qquad |B_L\rangle = p|B^0\rangle - q|\bar{B}^0\rangle \tag{1.13}$$

where

$$p = \frac{1+\epsilon_B}{\sqrt{2(1+|\epsilon_B|^2)}}, \qquad q = \frac{1-\epsilon_B}{\sqrt{2(1+|\epsilon_B|^2)}}, \tag{1.14}$$

with ϵ_B analogous to ϵ in the $K^0 - \bar{K}^0$ system. "H" and "L" denote *Heavy* and *Light* respectively. In the $B^0 - \bar{B}^0$ system one has $\Delta\Gamma \ll \Delta M$[6] and it is more suitable to distinguish the mass eigenstates by their masses than the corresponding life-times.

The strength of the $B^0_{d,s} - \bar{B}^0_{d,s}$ mixing is described by the mass difference

$$\Delta M_q = M^q_H - M^q_L, \qquad q = d, s. \tag{1.15}$$

ΔM_q can be expressed in terms of the off-diagonal elements of the neutral B-meson mass matrix, similar to the Kaon system,

$$\Delta M_q = 2|M^{(q)}_{12}|, \qquad q = d, s \tag{1.16}$$

[5] q can either denote a d or s quark.

[6] The ratio $\frac{\Delta\Gamma}{\Delta m}$ is expected of the order $\sim \left(\frac{m_b^2}{m_t^2}\right)$

with,

$$2m_{B_q}|M_{12}^{(q)}| = |\langle \bar{B}_q^0|\mathcal{H}_{eff}(\Delta B = 2)|B_q^0\rangle| \tag{1.17}$$

The Hamiltonian $\mathcal{H}_{eff}$ includes only the contribution from the top sector, due to $m_{u,c} \ll m_t$, with the following form,

$$\begin{aligned}\mathcal{H}_{eff}^{\Delta B=2} &= \frac{G_F^2}{16\pi^2} M_W^2 (V_{tb}^* V_{tq})^2 \eta_B S_0(x_t) \times \\ &\times \left[\alpha_s^{(5)}(\mu_b)\right]^{-6/23} \left[1 + \frac{\alpha_s^{(5)}(\mu_b)}{4\pi} J_5\right] \mathcal{Q}(\Delta B = 2) + h.c. \end{aligned} \tag{1.18}$$

where $\eta_B = 0.55 \pm 0.01$ denotes a QCD correction [BJW90, UKJS98] factor, the Wilson coefficient $S_0(x_t)$ and the relevant operator $\mathcal{Q}(\Delta B = 2) = (\bar{b}q)_{V-A}(\bar{b}q)_{V-A}$ [Bur01].

Finally one finds using 1.18

$$\Delta M_q = \frac{G_F^2}{6\pi^2} \eta_B m_{B_q} (\hat{B}_{B_q} F_{B_q}^2) M_W^2 S_0(x_t) |V_{tq}|^2, \tag{1.19}$$

where F_{B_q} is the B_q-meson decay constant and $\hat{B}_{B_q}$ is the renormalisation group invariant parameter. The calculation of these parameters represents the biggest uncertainty of about 20 % for the formula above, the result either obtained by QCD lattice calculations [LL00] or with the help of QCD sum rules [BBBD92].

From $B_d^0 - \bar{B}_d^0$ mixing parametrised by ΔM_d the side $BA = R_t$ of the unitarity triangle can be determined:

$$R_t = \frac{1}{\lambda} \frac{|V_{td}|}{|V_{cb}|}. \tag{1.20}$$

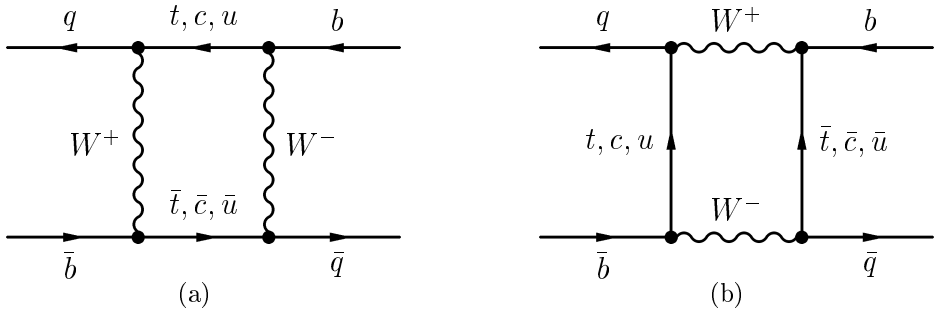

Figure 1.6: Box diagrams contributing to the $B^0 - \bar{B}^0$ mixing. q can either denote a s or a d quark.

A solution of the big theoretical uncertainties from F_{B_q} and $\hat{B}_{B_q}$ is the measurement of B_s^0 and B_d^0 mixing at the same time.

The additional measurement of $B_s^0 - \bar{B}_s^0$ mixing parametrised by ΔM_s together with ΔM_d allows the determination of R_t in a different manner. One finds

$$\frac{|V_{td}|}{|V_{ts}|} = \xi \sqrt{\frac{m_{B_s}}{m_{B_d}}} \sqrt{\frac{\Delta M_d}{\Delta M_s}}, \qquad \xi = \frac{F_{B_s}\sqrt{\hat{B}_{B_s}}}{F_{B_d}\sqrt{\hat{B}_{B_d}}} \tag{1.21}$$

and together with,

$$|V_{td}| = |V_{cb}|\lambda R_t, \qquad |V_{ts}| = |V_{cb}|(1 - \frac{1}{2}\lambda^2 + \bar{\rho}\lambda^2), \tag{1.22}$$

R_t can be derived. The theoretical uncertainties are now reduced to the parameter $\xi = 1.15 \pm 0.06$ [C$^+$01, LL00], which can be controlled much better than the elements ΔM_d and ΔM_s separately, due to cancellations of hadronic matrix element contributions.

1.2.4 Temporal Evolution

In the non relativistic approximation, the temporal evolution of any system described by a wavefunction $\psi(\vec{x}, t)$ will obey the Schroedinger equation describing the evolution of the physical vector,

$$i\frac{\partial \psi(\vec{x}, t)}{\partial t} = \hat{H}_{tot}\psi(\vec{x}, t) \tag{1.23}$$

The evolution in time of the states $|B^0(t)\rangle$ and $|\bar{B}^0(t)\rangle$ is given by

$$|B_{H,L}(t)\rangle = e^{-i(M_{12} - i\Gamma_{12}/2)t}|B_{H,L}\rangle \tag{1.24}$$

where $|B_{H,L}\rangle$ are given by equation 1.13. These are the physical mass eigenstates which possess a definite lifetime. At time $t = 0$, the states $|B_{H,L}(t)\rangle$ and $|\bar{B}_{H,L}(t)\rangle$ are produced as either pure B^0 or $\bar{B}^0$.

The evolution in proper time of a B meson created initially as either pure B^0 or $\bar{B}^0$ can be derived from equation 1.24 and is given by

$$|B^0(t)\rangle = f_+(t)|B^0\rangle + \eta f_-|\bar{B}^0\rangle \tag{1.25}$$

$$|\bar{B}^0(t)\rangle = \frac{f_-(t)}{\eta}|B^0\rangle + f_+|\bar{B}^0\rangle \tag{1.26}$$

where the deviation from 1 of the CP violation parameter $\eta = q/p$ is expected to be small. In mixing studies it is set to 1. Using $\Delta M = (M_H - M_L)/2$, and assuming that $\Gamma = (\Gamma_H + \Gamma_L)/2$, the $f_\pm(t)$ are given by

$$f_\pm(t) = \frac{1}{2}e^{\Gamma t/2}e^{-iMt}(e^{-(\Delta\Gamma/2 - \Delta M)t} \pm e^{(\Delta\Gamma/2 - \Delta M)t}) \tag{1.27}$$

The $f_+(t)$ represents the transition amplitudes for a pure $|B^0\rangle$ ($|\bar{B}^0\rangle$) state at time $t = 0$ to propagate as a $|B^0\rangle$ ($|\bar{B}^0\rangle$) and the $f_-(t)$ the transition amplitude for a pure $|B^0\rangle$ ($|\bar{B}^0\rangle$) state at time $t = 0$ to propagate as a $|\bar{B}^0\rangle$ ($|B^0\rangle$).

The probability density function $p_+(t)$ ($p_-(t)$) for a initially pure B^0 meson to decay as a B^0 ($\bar{B}^0$) as a function of the proper time is given by

$$p_\pm(t) = e^{-\Gamma t}\frac{\Gamma^2 - (\Delta\Gamma)^2}{2\Gamma}\left[\cosh\left(\frac{\Delta\Gamma}{2}t \pm \cos(\Delta m\, t)\right)\right] \tag{1.28}$$

For experimental purposes $\Delta\Gamma$ is set to zero and equation 1.28 will reduce to

$$p_\pm(t) = \frac{1}{2}\Gamma e^{-\Gamma t}(1 \pm \cos(\Delta m\, t)) \tag{1.29}$$

1.3 The Process $e^+e^- \to$ Hadrons

The description of $e^+e^- \to$ hadrons from high energetic e^+e^- collisions is theoretically well understood except the hadronisation process, which is not accessable to perurbation theory. The lack of knowledge is compensated by a combined model consisting of exactly calculable electroweak processes, phenomenological algorithms and the application of experimentally known hadron properties.

The model can be divided into four time ordered phases, illustrated in figure 1.7. The process starts with e^+e^- annihilation into a quark anti-quark pair via γ or Z exchange (Phase I). The high energetic strong interaction part can be calculated exactly in perturbative QCD (Phase II). The third part describing low energy QCD and hadron formation is modelled in a phenomenological way. Perturbative calculations are not applicable due to the large strong coupling constant α_s (Phase III). In the last phase the decay of the primary hadron resonances produced in phase III into the final state particles observed in the detector is modelled. In the case of heavy quarks (b and c quarks) produced in the Z decay special models like HQET[7] are used.

1.3.1 The Electroweak Phase

The process $e^+e^- \to q\bar{q}$ is governed by the Standard Model of electroweak interaction (see section 1.1). It can be described in first order (Born approximation) by the exchange of a photon or a Z^0 boson. Neglecting fermion

[7] Heavy-Quark-Effective-Theory

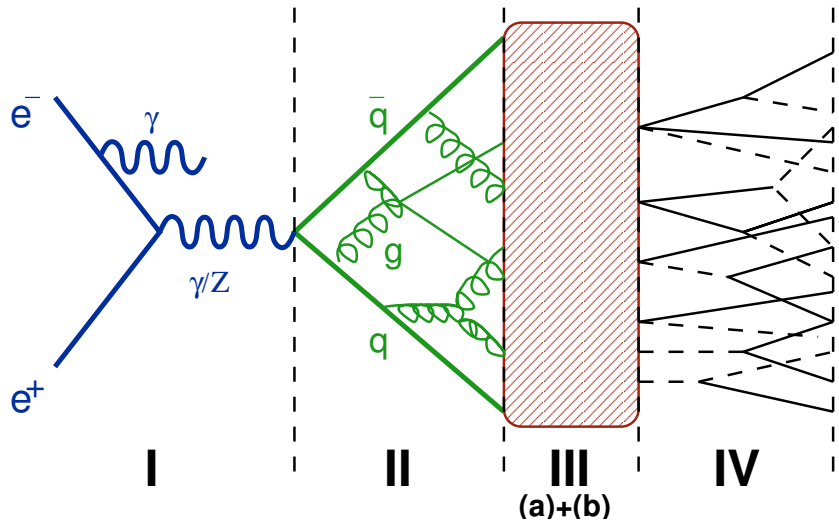

Figure 1.7: The four phases of e^+e^- annihilation: I. electroweak phase with initial state radiation, II. perturbative QCD phase based on gluon radiation and splitting as elementary processes, III. confinement and hadron formation phase and IV. decay into final state particles.

mass effects the total cross section can be derived as [Per]:

$$\begin{aligned}\sigma(e^+e^{\rightarrow}q\bar{q}) = N_C \cdot \frac{4\pi}{3} \cdot \frac{\alpha_s^2}{s} \quad &\cdot \quad [Q_e^2 Q_q^2 \\ &+ \quad (V_e^2 + A_e^2) \cdot (V_q^2 + A_q^2) \cdot |\chi|^2 \\ &+ \quad 2Q_q Q_q V_q V_e \cdot Re(\chi)], \end{aligned} \tag{1.30}$$

where

$$\chi = \frac{1}{4\sin^2\Theta_w \cdot \cos^2\Theta_w} \cdot \frac{s}{s - M_Z^2 + iM_Z\Gamma_Z}. \tag{1.31}$$

The origin of the different terms are the γ exchange for the first, the Z^0 exchange for the second and the γ/Z^0 interference for the third. V_f and A_f are the vector and axial vector couplings of the fermion to the Z^0. $\sin^2\Theta_w$ denotes the weak mixing angle, α the fine structure constant, N_C the number of colours and s the center of mass energy squared.

Formula 1.30 is modified by electroweak and strong interaction corrections of higher order. These corrections are virtual loop diagrams, illustrated in figure 1.1 and radiative corrections of photons in the initial and final states.

Based on the behaviour of quarks (see section 1.1.3) at high energies the evolution of the initial quark anti-quark pair can be modelled in two different phases dependent on their characteristic energies, as described in the next section.

1.3.2 The Perturbative Phase

Two different approaches are commonly used to calculate the quark and gluon cross sections in phase II. The first method is the complete second order QCD matrix element (ME) calculation, and the second is the parton shower (PS) approach in leading log approximation.

QCD Matrix Elements

In this approach all QCD Feynmann diagrams are calculated up to a certain order. Currently calculations exist up to second order, which allows a description of final states with four partons at maximum. Third order matrix elements are under construction by different groups, but not yet available in Monte Carlo programs.

Parton Shower Model

The exact calculation of multi-parton final states requires a huge amount of Feynmann diagrams, which have to be taken into account. As this was impossible in former times, an alternative approach, the Leading Log Approximation (LLA), was developed[8]. In LLA only the first order terms are taken into account and a multi-parton configuration is reached by subsequent application of the basic QCD processes $q \to qg$, $g \to q\bar{q}$ and $g \to gg$. The probability of such a process $a \to bc$ taking place is described by the Altarelli-Paresi equation [GAA+00]:

$$\frac{d\mathcal{P}_{a\to bc}}{dt} = \int dz \frac{\alpha_s(Q^2)}{2\pi} \mathcal{P}_{a\to bc}(z). \tag{1.32}$$

where $t = \ln(Q^2/\Lambda^2_{QCD})$ denotes the evolution parameter and $\mathcal{P}_{a\to bc}(z)$ are the Altarelli-Paresi-Splitting functions,

$$\mathcal{P}_{q\to qg}(z) = C_F \frac{1+z^2}{1-z}, \tag{1.33}$$

$$\mathcal{P}_{g\to gg}(z) = N_C \frac{(1-z(1-z))^2}{z(1-z)}, \tag{1.34}$$

$$\mathcal{P}_{g\to q\bar{q}}(z) = T_R(z^2+(1-z)^2), \tag{1.35}$$

with $C_F = 4/3$, $N_C = 3$ and $T_R = N_F/2$. z or $(1-z)$ denotes the momentum fraction of the initial parton a transfered to the final state parton b (c). The

[8] Latest developments indicate progress in the development of QCD matrix element based models [CKKW01, Pot01]

evolution of the shower via parton splitting is stopped when an energy scale $Q_0 \approx 1$ GeV is reached. The transition to colour neutral hadrons is described in the next section.

1.3.3 The Fragmentation Phase

The perturbative approach fails in calculating the formation of bound hadron states, when the value of the coupling constant α_s becomes large i.e. at a Q^2 scale of the order of Λ^2_{QCD}. Therefore only phenomenological models motivated by QCD are feasible. Two commonly used models are discussed in this section: the *string fragmentation*, which is the default model used in the DELPHI Monte Carlo generator [Sjo94], and the *cluster fragmentation* [MW84].

The String Fragmentation

In the simplest model of string fragmentation [Mor89, AGIS83] a pair of massless quark and anti-quark move in opposite directions with velocity c. Based on "confinement" the self coupling of the gluons leads to a linear colour field between the two quarks, called a string. The energy density within the colour field is a constant of about $\kappa \approx 1GeV/fm$, leading to a linearly growing potential between the quarks. If the energy in the colour field grows beyond the $q\bar{q}$ creation threshold the string breaks up and a new meson is formed. This process is shown schematically in figure 1.8.

The two $q\bar{q}$ systems share energy and longitudinal momentum according to z and $(1-z)$. The probability distribution of z is given by the longitudinal fragmentation function $f(z)$. Based on experimental data a good description

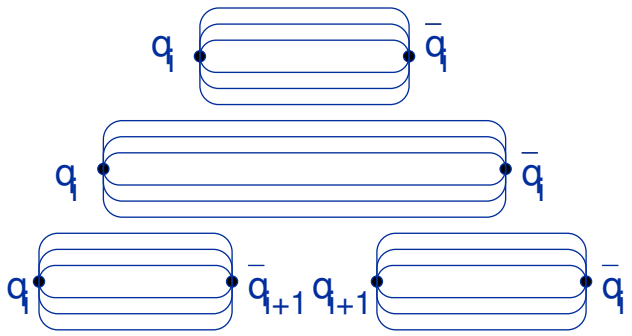

Figure 1.8: The basic picture of string fragmentation is illustrated. A $q\bar{q}$ breaks up in two $q\bar{q}$ systems sharing energy and longitudinal momentum.

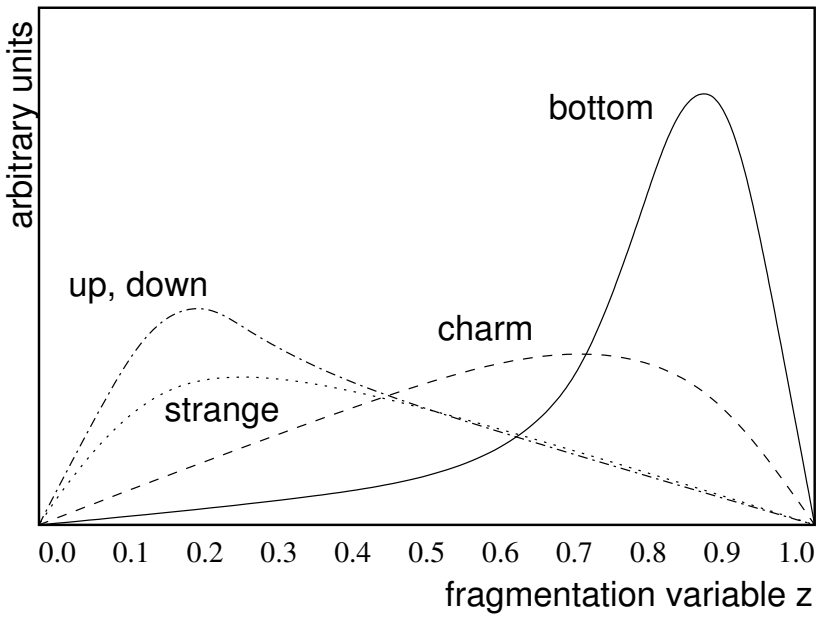

Figure 1.9: The fragmentation function for different quark types. The heavier quarks are located at higher z represented by their "hard fragmentation".

for heavy quarks (b and c) can be achieved by the Peterson function [PSSZ83],

$$f(z) = \frac{1}{z(1 - 1/z - \epsilon_q/(1-z))^2}. \tag{1.36}$$

were ϵ_q denotes a free parameter, which is in principle the squared ratio of the masses of heavy to light quarks. However as the light quark masses are not precisely known, the value of ϵ_q is obtained from data measurements in practice.

The light flavours (u, d and s) can be described by the symmetric Lund function,

$$f(z) = \frac{(1-z)^a}{z} \cdot e^{-bm_\perp^2/z} \tag{1.37}$$

with two free parameters, which have to be fitted to data. In figure 1.9 one can see a schematic view of the fragmentation functions for the individual flavours. It can be seen clearly that the maximum of the functions for heavy quarks is located at high z. This behaviour is called "hard fragmentation" and leads to the effect that primary heavy quarks keep most of their initial energy, when combining up with other quarks to form hadrons.

Gluons are incorporated quite easily in the string fragmentation picture. They are treated as transverse excitations of the colour field. This allows the modelling of soft and hard gluons without additional parameters.

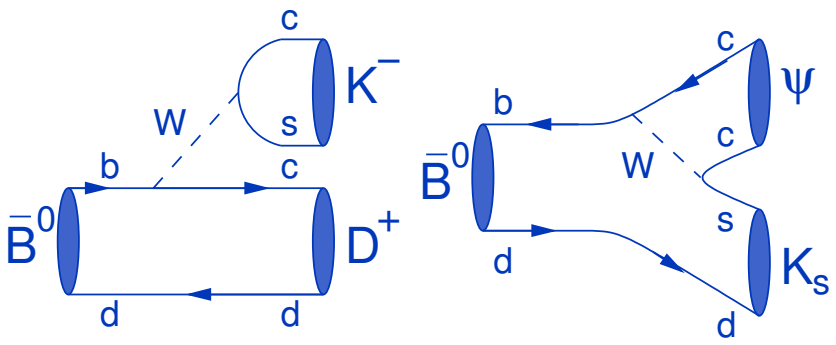

Figure 1.10: Two examples for b decays in the spectator model. The left figure illustrates a weak decay of a $\bar{B}^0$ meson in a D^+ and a K^-. On the right side the weak decay of a B^0 meson in a J/Ψ and a K_s is visible.

Cluster Fragmentation

Cluster fragmentation models have been used in some simulation programs since the mid 80's by different groups [Web84, MT85, GM87]. The principle of the model is, that after the perturbative phase which developed up to an energy scale Q_0, all gluons split into $q\bar{q}$ pairs. In the next step all quarks and anti-quarks are collected in colour-neutral clusters. These clusters then decay preserving certain colour-, spin- and phase space constraints into primary hadrons.

Cluster models profit from the fact that they do not need a fragmentation function and the number of free parameters is smaller compared to string models. Detailed comparisons between data and monte carlo have shown in recent years, that cluster fragmentation models cannot describe the data as well as the sting models.

1.3.4 Heavy Quarks and Particle Decays

The final simulation process after the fragmentation phase is the decay of all primary mesons and baryons into stable final particles. For this purpose tables containing all measured lifetimes and branching ratios are implemented. For the case of heavy hadrons, containing b quarks, the experimental knowledge about their decay properties is limited. On the other hand the high mass of the b quark of around 5 GeV which is much larger than the typical strong interaction scale $\Lambda_{QCD} \approx 0.2$ GeV, opens up the possibility of a theoretical description.

A general principle of the theoretical approaches is the dominance of the heavy b quark compared to the light *spectator* quark. The influence of the

light quark is ignored in the spectator model. Spectator effects are expressed in terms of higher order perturbative corrections to the pure b quark behaviour. In figure 1.10 two examples for B meson decays are illustrated.

A dedicated theoretical framework for the description of B hadron properties is provided by Heavy Quark Effective Theory (HQET), Heavy Quark Symmetry (HQS) and Heavy Quark Expansion (HQE) [Mar, Vai95, Dav97].

Chapter 2

The Experiment

2.1 LEP

The LEP[1] e^+e^- collider started in 1989, after a construction time of over 6 years. It is located at the european research center CERN[2], Geneva, between the Jura mountains and the airport (see figure 2.1). The main ring consists of eight straight sections, the interaction zones, and eight arcs giving a total circumference of 26.7 km, constructed at an average depth of approximately 100 m underground. At four interactions zones the detectors ALEPH[3], DELPHI [4], L3[5] and OPAL[6] are sited.

In the so called LEP-I phase from 1989 to 1995 the e^+e^- collisions had an energy of 91.2 ± 3 GeV, which corresponds to the rest mass of the Z^0 boson. The cross section at this energy is increased due to resonant production, leading to an enormous data statistic. This huge statistic is the key for high precision tests of the Standard Model, the main goal of the LEP-I phase.

In the LEP-II phase the center of mass energy was increased from 130 GeV up to 208 GeV. For this purpose all 128 copper RF acceleration cavities were replaced by super-conducting Niob cavities. One of the main goals of this phase was the exact measurement of the W^+W^- production cross section, also the mass, width and the coupling of the W^+ and the W^- at an energy beyond 160 GeV. Further goals were the search for new physics, such as the Higgs boson and supersymmetric particles, and energy dependence measurements of Standard Model observables, e.g. the forward backward

[1] Large Electron Positron
[2] Conseil Europe´en pour la Reserche Nucle'aire
[3] A detector for LEP PHysics
[4] DEtector with Lepton, Photon and Hadron Identification
[5] Named after: 3. Letter of intent
[6] Omni Purpose Apparatus for LEP

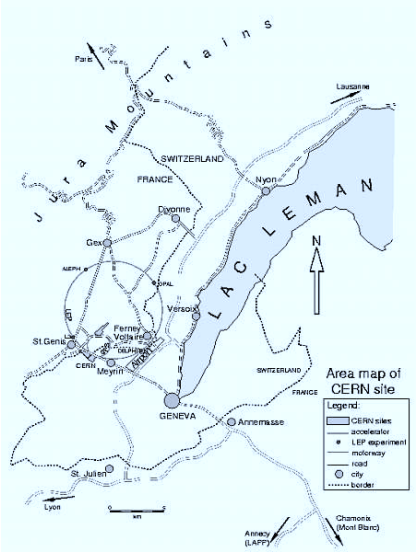

Figure 2.1: Geographical location of the LEP collider

asymmetry A_{FB}.

The experimental LEP activities ended in November 2000. The experiments were dismantled and the ring was removed. The LEP tunnel will be used for the new proton antiproton collider ring LHC[7] with a designed center of mass energy of 14 TeV, which will be installed up to the year 2007. Main goal of the LHC is the discovery of the Higgs boson and the search for new physics, for which two experiments will be installed (ATLAS and CMS). A dedicated B-factory (LHC-B) and a heavy ion experiment (ALICE) are also under construction for the LHC ring. First data of these experiments can be expected in the year 2008.

[7]Large Hadron Collider

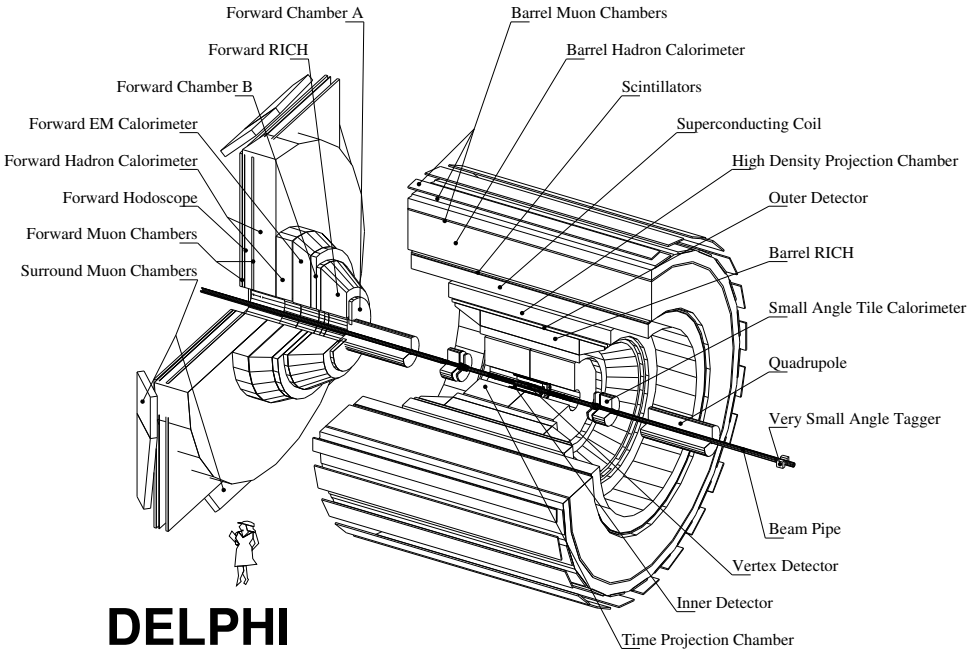

Figure 2.2: Schematic view of the DELPHI detector (Status year 2000).

2.2 The Delphi Detector

DELPHI was designed as an omni-purpose detector covering the whole solid angle of 4π. It was constructed for precise reconstruction of charged tracks and a calorimetry with fine granularity to provide good reconstruction of neutral particles. DELPHI was not constructed on well proven technology only, like the OPAL detector. There was also room for new developments in detector technology like the HPC[8] and the RICH[9] detectors, which gave DELPHI its unique possibility for particle identification.

The individual tracking subdetectors were embedded in a magnetic field of 1.23 T from a superconducting magnetic solenoid, forcing charged tracks to a curvature, which allowed the measurement of their momentum. The detectors were arranged in a cylindrical barrel region and two forward endcaps on either side (see figure 2.2). The coordinate system of the detector was defined by the barrel axis starting at the interaction point. This z-axis followed the electron beam, the x-axis points to the center of LEP and the y-axis to the top direction. The polar angle with respect to z was called Θ.

[8]High density Projection Chamber
[9]Ring Imagine Cherencow cHambers

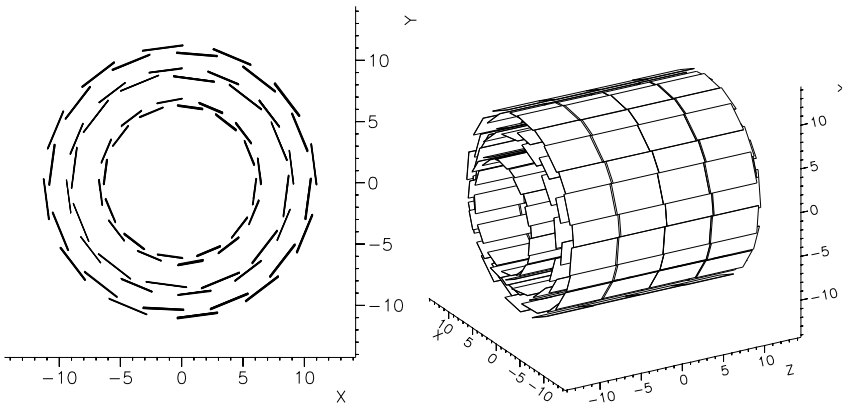

Figure 2.3: Layout of the Delphi microvertex detector.

The azimuthal angle around the z-axis was called Φ. An overview of the individual acceptance regions and the resolutions in space of the subdetectors is given in the tables 2.1 and 2.2. In the following sections brief descriptions of the most important detector parts are provided. A detailed description of the performance of the DELHI detector can be found in [A$^+$91, A$^+$96b].

2.2.1 Tracking System

The **Vertex Detector (VD)** is one of the most important detector components for B physics. Mounted closely to the interaction point, it allowed high precision track measurements, necessary for the reconstruction of secondary vertices (e.g. B-, D- mesons). Three cylindrical shells of silicon micro-strip detectors surrounding the interaction point parallel to the beam pipe at radii of 6.3, 9.0 and 10.9 cm, build the VD. Each layer consisted of 24 modules, which were mounted with small overlap to fully cover the azimuthal angle fully (see figure 2.3). During the winter break of 1994 the VD was modified by equipping two layers of the VD with double sided micro strip detectors which provided measurements in $R\Phi$ and z direction. The acceptance for the different years can be found in table 2.1. The intrinsic resolution of the VD is 7.6 μm in $R\Phi$ and 9 μm in Z (for perpendicular tracks).

The **Inner Detector (ID)** followed the VD with an inner radius of 11.8 cm up to 28.0 cm outer radius. It covered in its cylindrical volume the polar angle from 17° to 163° and was mainly used for fast trigger decisions ($\approx 3\,\mu$s) and gave some redundancy in the track reconstruction for the VD. The ID consisted of two different parts. The inner part was a jet-chamber architecture pointing to the interaction point, the outer was a Multi-Wire-

VD layer	Year: 1992/93	Year: 1994/95	Year: 1996/00
closer	30° $R\Phi$ readout	27° $R\Phi$, z readout	23° $R\Phi$, z readout
inner	35° $R\Phi$ readout	35° $R\Phi$ readout	23° $R\Phi$ readout
outer	40° $R\Phi$ readout	40° $R\Phi$, z readout	23° $R\Phi$, z readout

Table 2.1: The polar angle acceptance of the VD for the different years.

Proportional-Chamber (MWPC) starting at a radius of 23 cm. The jet chamber is divided into 24 sectors in Φ with 24 wires each. The MWPC consisted of five layers with 192 signal wires in each layer, which served to resolve the left-right ambiguity of the jet-chamber and allowed measurements in z with a precision of 0.5 to 1 mm depending on Θ. The achieved resolution of the ID was 50 μm in $R\Phi$ and 1.5 mrad in Φ.

In 1995 the MWPC layers were replaced by Straw-Tube detectors, which delivered no z information, but covered an increased polar angle region from 15° to 165°.

The **Time Projection Chamber (TPC)** was the main tracking device of the DELPHI detector. Its active cylindrical volume ranged from an inner radius of 35 cm to a outer radius of 111 cm with a length of 134 cm. It was filled with an argon-methane gas mixture. As one can see in the schematic view of figure 2.4, the TPC was divided into two hemispheres with six sectors in Φ. It was read out at the end-caps by 16 concentric pad rows and 192 anode wires. A track passing through the gas left a tube of ionized gas along its way. A homogeneous electric field along the z direction lead to a drift of the ionisation electrons to the end-caps. The $R\Phi$ resolution was governed by the segmentation of the end-cap, while the z information was delivered by the drift time. This lead to the possibility of measuring up to 16 space points per track. The single-point resolution for tracks was 250 μm in $R\Phi$ and 880 μm in z.

In addition to the 3D space measurement, the TPC allowed the readout of the signal height of each point. Due to this measurement the specific energy loss $\frac{dE}{dx}$ gave a contribution to the particle identification system for low velocity particles ($< 1\,GeV$).

The **Outer Detector (OD)** was a five layered drift tube system mounted at radii ranging from 197 to 206 cm, surrounding the central cherenkov detectors (see section 2.2.3). The detector with an approximate length of 460 cm, covered a polar angle from 42° to 138°. A track resolution of 110 μm in $R\Phi$ and by drift time measurement, 3.5 cm in z was reached. The OD was used

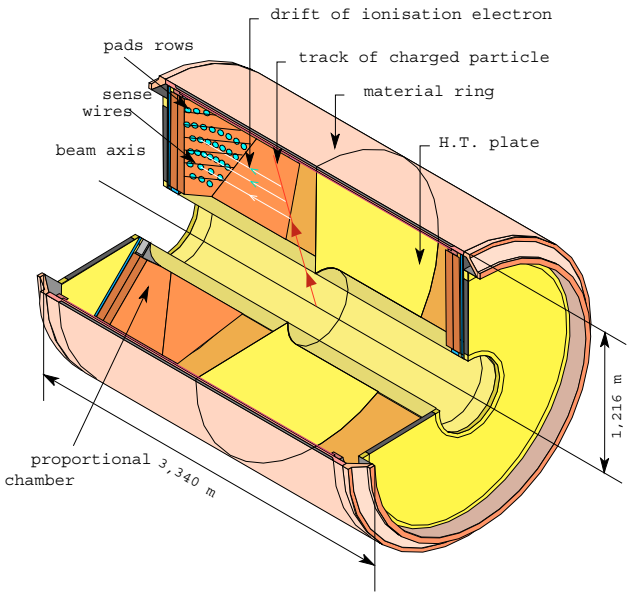

Figure 2.4: The TPC was the main tracking device of the DELPHI detector.

mainly for the matching of particle tracks to signals in the Cherenkov counters. It also improved the curvature measurement of high energetic tracks and delivered a fast trigger signal.

The **Forward Chambers A+B (FCA, FCB)** were track detectors for the forward region. The FCA was mounted on the end-caps of the TPC at $z = \pm 160$ cm, based on a straw tube technique. A polar angle region from 11° to 33° was covered and a track resolution of 290 μm for the x and y direction was reached. The FCB was a 12 layered multi-wire proportional chamber mounted on the FCA. The chamber was used for the trigger decisions of the forward region, covering a polar angle of 11° to 35°, while a track resolution of 150 μm in x and y direction was reached.

The optimal **performance of the tracking** of the DELPHI detector is given by a track reconstruction including all of the different components. The main limitations came from the fact that a rather big part of the space within the magnetic field was used by the Cherenkov detectors for particle identification (see section 2.2.3). By using $Z^0 \rightarrow \mu^+\mu^-$ events the total momentum resolution in the barrel part was determined to be

$$\sigma(1/p) = 0.57 \times 10^{-3}\,(\mathrm{Gev/c})^{-1}, \tag{2.1}$$

combining VD, ID, TPC and OD track elements. In the forward region from

20° to 35° the momentum resolution was,

$$\sigma(1/p) = 1.31 \times 10^{-3}\,(\mathrm{Gev/c})^{-1}, \tag{2.2}$$

requiring at least VD and FCB.

A summary of the different polar acceptance regions and the resolution of the measured space points for the individual detectors is given in table 2.2

2.2.2 Calorimetry

The **High density Projection Chamber (HPC)** was the electromagnetic barrel calorimeter mounted directly on the inner side of the magnet, covering a polar angle region from 42° to 137°. The HPC consisted of 144 modules arranged in 6 rings with 24 modules in each ring. It was a lead-gas sampling calorimeter based on the TPC principle, delivering a good angular resolution of 1° in Φ direction and 4 mm in z direction.

The **Electromagnetic Forward (EMF)** calorimeter covered a polar angle from 10° to 36.5°. It consisted of 2×4532 leadglas blocks with approximately 20 radiation lengths, covering an angle of $1° \times 1°$ each.

The **Hadronic Barrel (HAB)** calorimeter was installed within the iron yoke of the DELPHI solenoid. The end-cap regions were covered by the Hadron Calorimeter Forward **(HAF)**. The thickness of the forward and the barrel calorimeters was about 6 radiation lengths, covering a polar angle from 10° to 170°. The achieved granularity was 3.75° in Φ direction and 3° (2.6°) in the barrel (end-cap) Θ direction.

The **luminosity measurement** was covered by the **SAT**[10] and **VSAT**[11] calorimeters installed in the extreme forward region. The measurement of the luminosity delivered by LEP was done by counting Bhabba events with low Q^2. The SAT consisting of lead and scintillating fibers was mounted at $z = \pm 2.35$ m. It was replaced in the winter break of 1993 by the **STIC**, a lead-scintillator sampling calorimeter with a relative error of 0.2 %, achieving a better resolution compared to the SAT. Another independent measurement is provided by the tungsten-sampling calorimeter VSAT mounted at $z = \pm 7.5$ m.

[10] Small angle tagger
[11] Very Small angle tagger

Detector	Acceptance region			Points	Resolution/point
	R [cm]	$\|z\|$ [cm]	Θ [$\circ$]		[σ, mm]
VD(1994)	6.3, 9.0 10.9	12.3, 12.9 13.0	27, 35 40	max. 5 2	$R\Phi$:0.007-0.011 z : 0.013
ID(jet) (MWPC)	11.8-22-3 23.0-28.0	40 50	17-163 30-150	24 5	$R\Phi$: 0.05 z : 0.5-1.0
TPC	35.0-111.0	≤ 134	20-160	16 192	$R\Phi$: 0.23 z : 0.88
OD	197.0-206.0	≤ 232	42-138	$5 \times R\Phi$ $3 \times z$	$R\Phi$: 0.11 z : 35
MUB	≈ 445 ≈ 485	≤ 185	52-128	2(+2)	$R\Phi$: 1.5 z : 10
FCA	30.0-103.0	155-165	11-33	$2 \times (x, u, v)$	x : 0.29 u,v$\rightarrow$y: 0.24
FCB	53.0-195.0	267-283	11-35	$4 \times (x, u, v)$	x : 0.15 u,v$\rightarrow$y: 0.15
MUF	70.0-460.0	463	9-43	$(4) \times (x, y)$	x,y : 1.0

Table 2.2: The acceptance region and the resolution for individual sub-detectors

2.2.3 Identification

A speciality distinguishing DELPHI from the other three LEP detectors was its dedicated particle identification framework besides the ionisation loss measured in the TPC.

The **Ring Imaging Chrenokov (RICH)** detectors were two detector components using the effect of Cherenkov light production for charged particles traversing a dielectric medium with a velocity larger than the speed of light in that medium. The emission angel Θ_{CH} depends on the mass m and the momentum p via the relation

$$\cos\Theta_{CH} = \frac{\sqrt{1+m^2/p^2}}{n} \tag{2.3}$$

where n is the refractive index of the radiator medium. The information about the Cherenkov angle and the number of photons emitted, which is proportional to $\sin\Theta_{CH}$, are used to evaluate masses of charged particles.

The DELPHI RICH contained two radiator systems of different refractive index. A liquid radiator was used for particle identification in the momentum range from 0.7-4.0 GeV/c and a gas radiator was used from 2.5-25.0 GeV/c. Two independent detectors were available to cover the full barrel and endcap regions. In figure 2.5 the Cherenkov angle and the ionisation loss $\frac{dE}{dx}$ is illustrated for different particle types.

The **muon chambers** in the forward (MUF) and in the barrel (MUB) consisted of double-layered drift chambers. One of the layers was mounted on the outer side of the hadron calorimeter, building the surface detector component of DELPHI. The other was assembled inside the hadron calorimeter. The structure had a geometry of 18.8 cm width, 2.2 cm height and a few meters length. The resolution in $R\Phi$ was 1.5 mm in the barrel region. The z coordinate was measured by the signal time difference of either end of the chamber, lead to a precision of 1 cm. The polar angle region was covered in the region from 9° to 43°, 52° to 128° and 137° to 171°. The remaining gaps were filled by the Surrounding Muon Chamber (MUS), installed in the winter break of 1994.

2.2.4 The Delphi Analysis Chain

The following section gives a short overview about the DELPHI analysis chain. Starting with the trigger, slow control and data acquisition online system, then the DELANA, the DST and DELSIM offline system. For a detailed discussion see [del89, Sac94, Q$^+$89, B$^+$95, A$^+$92].

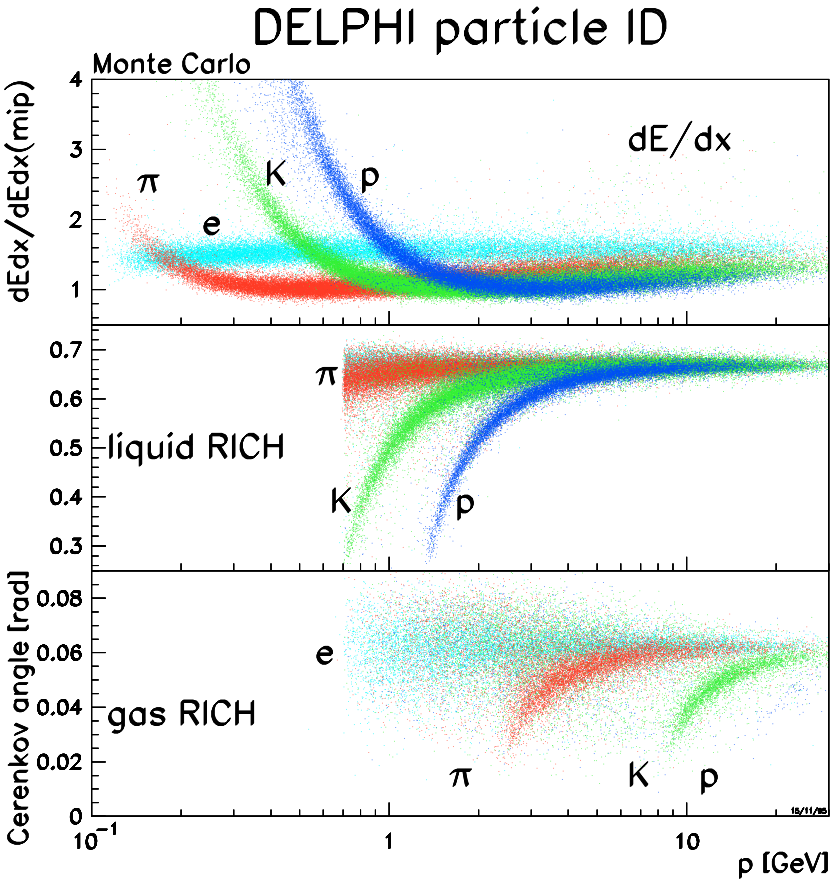

Figure 2.5: dE/dx and RICH information for particle id taken from simulation. The energy loss dE/dx inside the TPC (top), and the cherenkov angles in the liquid (middle) and gas (bottom) radiator for various particles.

The **slow control system** monitored and controled the status of the detector, e.g. high voltages of all subdetectors, gas supplies, temperature, etc. All relevant parameters were stored in a database, used later for the calibration of all the detectors.

The **Trigger** was build on four steps (T1-T4), with the purpose to separate the events with physical interest from background events. The steps T1 and T2 were hardware based systems, synchronised with the beam cross over (BCO). The time difference between two beam crossings was $22\mu s(11\mu s)$ in runs with 4(8) electrons packets per beam. Thus the T1 trigger based on the detectors with a very fast readout (ID,OD,FCA,FCB) and the scintillators. In the case of a positive decision of T1, the additional information of T2 (e.g. TPC, HPC) was used, leading to a decision after $39\mu s$ for T1 and T2. The next steps T3 and T4 were software based triggers, using already a partial event reconstruction and particle identification. The trigger rates were 400 Hz for T1 and 5 Hz for T2. T3 and T4 reduced the background by an additional factor 1.5. The trigger efficiency of the system was extremely high and for multihadronic events close to 100 % over the whole polar angle.

The **data acquisition system (DAS)** had the purpose to read out the digitised data of all detectors and store them on magnet tapes. All events accepted by the second trigger step T2 were stored on tape. The mean size of a Z^0 multihadron event stored as 'raw' event on tape was around 150 kB. The events were identified by a 'FILL', which means a LEP filling, a 'RUN' number, which means a status of stable conditions for the detector and a 'EVENT' number.

The **DELANA** program was the main reconstruction program of DELPHI [del89]. It contained a module for every subdetector which performed the necessary alignment, calibration of the raw data and a local track reconstruction, if possible. The result of this first reconstruction step were track elements, which were connected by a global search algorithm, starting at the TPC and then extrapolating to the inner and outer side of the detector. After assigning vertex detector hits to these track candidates, the whole track fit was redone and a matching of tracks and neutral clusters was done. Finally, based on the fitted tracks, a primary vertex was calculated.

The result of this procedure is a reconstructed event, stored in the DST[12] format, including all relevant information like momentum vectors, energy deposition and others. The size of a multihadronic Z^0 event is at this point around 80 kB.

[12]Data summary tape

An additional reduction of the data size is achieved by the production of the **Short DST (SDST)** format. Aim of this step is the running of some standard reconstruction software lead to savings in the CPU usage of the final analysis. The most important are:

- The tagging of $Z^0 \to b\bar{b}$ events.
- Identification of electrons, photons and neutral pions.
- Myon identification.
- Reconstruction of V^0 and Λ^0 decays
- Reconstruction of VD only tracks.
- Reconstruction of interactions with the detector material.

The removement of some individual detector information lead to a mean size of 20 Kb per multihadronic Z^0 decay for the SDST format. The SDST format was replaced in 1996 by the XSDST, which contained some additional detector information.

The **detector simulation of Delphi (DELSIM)** is a program which produces 'data' for a particular reaction which are as close as possible to the real raw data from the detector. These 'data' are then processed through the reconstruction program DELANA and the subsequent analysis programs in exactly the same way as for real data. DELSIM can be split into three different parts.

The first part starts with the physical generation of a primary physics process. Usually external MonteCarlo generators were used for this, e.g. JETSET [Sjo94], ARIADNE [Lon92], tuned with physical input from the latest available data. In the second step all generated particles were tracked through the detector up to the point where they hit an active detector component. This is done by stepping through the magnetic field including the possibility that these particle can undergo a secondary interaction. This step requires a detailed material description for the different detector components. In the last step the detector response of each component is modelled.

Chapter 3

Mathematical Methods

A very common problem in a high energy physics analysis is the separation between a certain type of event (signal) from a large amount of background. In an inclusive analysis, like the one described in this work, the separation power of a single physical quantity ('input variable') is usually quite weak and so effective tagging power is only achieved by the combination of many such variables. Various methods can be used to solve the problem of optimal variable combination e.g. Fisher discriminant analysis [Fis36], Likelihood Method [Cow98] or Neural Networks [HKP91]. There are two main reasons why Neural Networks are preferred in this analysis: First is a Neural Network contains a set of free parameters, the weights w, which naturally contain information about the correlations between the physical input quantities and the desired 'target'. These weights can be iteratively obtained by presenting examples of signal and background events ('training') to the Neural Network together with the desired target. The second desirable feature of Neural Networks is the robustness of the network against statistical correlations between the input variables, which are not correlated with the desired target value, in contrast to e.g. the Likelihood Method. A brief mathematical introduction to Neural Networks is provided in the first part of this chapter.

In the second part the numerical integration procedure will be motivated and explained (see section 3.2). It is based on a special set of random numbers, so-called quasi random numbers, and a dedicated variable transformation algorithm.

3.1 Neural Networks

The development of Neural Networks started in the 1960's and was partially inspired by Neuroscience [Ros58]. After a promising start it took over 20

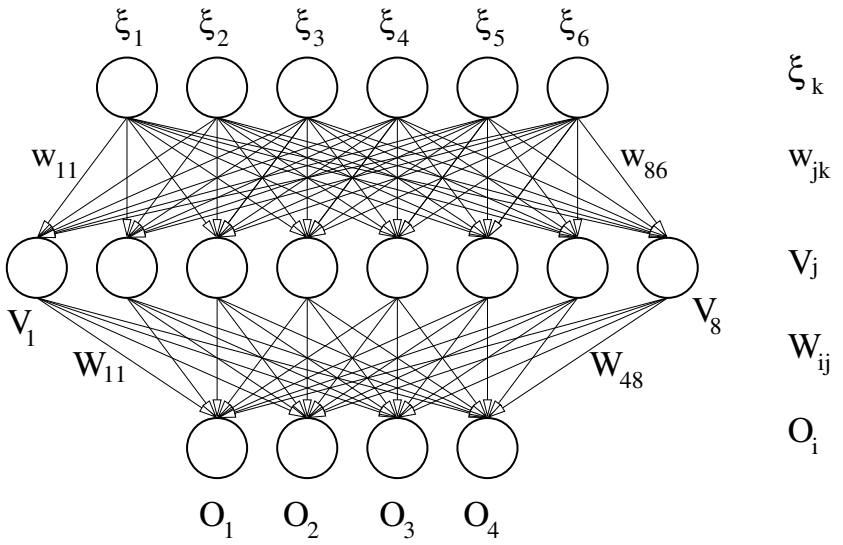

Figure 3.1: A two layer Feed Forward Neural Network, illustrating the notation for units and weights. ξ_k denotes the input units, V_j the hidden layer units, O_i the output units, W_{ij} and w_{jk} the weights of the connections.

years until the development of multi-layer networks, in combination with an effective learning algorithm, created a usable tool for science [MR81,RHW86].

The different networks used in the analysis are all of the same type, the so called Multilayer-Feed-Forward Network [HKP91]. They are optimal for the separation of Signal from Background on the basis of different physical variables.

3.1.1 Feed Forward Neural Networks

The basic properties of a neural network are units (knots, neurons) and connections with associated weights. Figure 3.1 illustrates the topology of a two-layered Feed-Forward Network, showing the notation for units and weights used in the section. The information flow is explained on an example pattern μ. It starts at the inputs ξ_k^μ and processes to the hidden unit V_j by

$$V_j^\mu = g(h_j^\mu) = g(\sum_k w_{jk}\xi_k^\mu) \tag{3.1}$$

where $g(h)$ is a special function ('activation function') computed by the units. It is usual to use a sigmoid function for the activation function $g(h)$, which has the form:

$$g(h) = \frac{1}{1+e^{-c\cdot h}}, \tag{3.2}$$

where c is a parameter ('temperature') put to a fixed value. Output unit i thus produces the final output:

$$O_i^\mu = g(\sum_j W_{ij} V_j^\mu) = g(\sum_j W_{ij} g(\sum_k w_{jk} \xi_k^\mu)) \tag{3.3}$$

An error measure can be found by comparison of the final output O_i^μ with the desired output ζ_i^μ via,

$$\begin{aligned} E(\omega) &= \frac{1}{2} \sum_{\mu,i} (\zeta_i^\mu - O_i^\mu)^2. \\ &= \frac{1}{2} \sum_{\mu,i} [\zeta_i^\mu - g(\sum_j W_{ij} g(\sum_k w_{jk} \xi_k^\mu))] \end{aligned} \tag{3.4}$$

As function 3.4 is a continuous differentiable function[1] of every weight it is possible to apply a gradient descent algorithm to learn appropriate weights.

3.1.2 Training a Neural Network

The gradient descent algorithm ('gradient descent rule') suggests changing of each w_{ij} by an amount Δw_{ij} proportional to the gradient of E at the present location, which is equivalent to sliding downhill on the surface in ω space. For the hidden-to-output connections the gradient descent rule gives

$$\begin{aligned} \Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} &= \eta \sum_\mu (\zeta_i^\mu - O_i^\mu) g'(h_i^\mu) V_j^\mu \\ &= \eta \sum_\mu \delta_i^\mu V_j^\mu \qquad (3.5) \\ \text{with} \quad \delta_i^\mu &= g'(h_i^\mu)(\zeta_i^\mu - O_i^\mu) \qquad (3.6) \end{aligned}$$

For the input-to-hidden connections Δw_{jk} the gradient is more compli-

[1] Only true if function g is also continuous and differentiable

cated as they are more embedded in 3.4. One obtains

$$\begin{aligned}
\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} &= -\eta \sum_{\mu} \frac{\partial E}{\partial V_j^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}} \\
&= \eta \sum_{\mu,i} (\zeta_i^{\mu} - O_i^{\mu}) g'(h_i^{\mu}) W_{ij} g'(h_j^{\mu}) \xi_k^{\mu} \\
&= \eta \sum_{\mu,i} \delta_i^{\mu} W_{ij} g'(h_j^{\mu}) \xi_k^{\mu} \\
&= \eta \sum_{\mu} \delta_j^{\mu} \xi_k^{\mu} \qquad (3.7)
\end{aligned}$$

$$\text{with} \quad \delta_j^{\mu} = g'(h_j^{\mu}) \sum_i W_{ij} \delta_i^{\mu} \qquad (3.8)$$

Equation 3.8 determines the δ for a given unit V_j in terms of the δ's of the unit O_i that it feeds. The coefficients are just the usual W_{ij}'s, but they are propagating errors (δ's) backwards instead of signals forward: hence the name error back-propagation.

3.1.3 Application in High Energy Physics

As briefly discussed in the introduction the network is trained by a set of simulated events. The simulated event set is split into a training and a test sample. The test sample is necessary to control the error behaviour over different iterations, making sure that the generalisation ability is not lost due to overtraining. During one training epoch all events of the training set are presented to the network and the back propagation algorithm is applied. After each epoch the network error is calculated from the test set and compared to the error of the training set. The repeated procedure of training and testing the network is stopped, if the test error reaches a plateau. The training error would drop even further, but only due to statistical fluctuations learned by the network. If this is the case a network is overtrained and it's generalisation ability is weakened. Usually a network is trained after 100 to 400 epochs depending on the problem the network is used for.

3.2 Numerical Integration

A common problem in a physics analysis is related to limitations of the measurement of a certain physics quantity, i.e. the momentum of a B meson is only partially reconstructed due to undetected particles and detector effects.

Using simulated events it is possible to extract resolution functions, providing a connection between the measured and the true value. Starting from a theoretical prediction for a physics quantity, one can extract an estimate of the expected measured distribution, by the convolution of the theory function with the resolution functions.

The convolution is realized by numerical integration in this analysis. The separate resolution functions of two physical observables (decay length and momentum of the B meson) requires a two-dimensional integration for every event, because the resolution functions change event by event. Due to the large statistics available of approximately half a million events, computing power is a limiting factor if a standard integration algorithm is used. Therfore two approaches were developed to optimise the speed of the integration while keeping a high numerical precision.

The first approach is the point set which is used for the integration. Usually on would think, that a normal grid is the optimal choice in two dimension, where in general the error drops with the inverse of N where N denotes the number of points used for the integration (In fact for normal random numbers the error converges with $1/\sqrt{N}$). However, one can find point sets ('Quasi Random Numbers') where the error converges faster compared to points on a grid. This can be understood if one does a projection of the points onto a coordinate axis. A projection of the grid onto e.g. the x-axis leads to a distribution with a lot of spikes, while the same projection is flat for quasi random numbers (In fact any arbitrary projection axis can be chosen to get a flat distribution for a quasi random point set.) Therefore the phase space is more uniformly filled by the quasi random numbers (the uniformity is mathematically based on the definition of a 'discrepancy' [JJR97]). Two different quasi algorithms were tested for this analysis, called 'Niederreiter' and 'Hammersly' [H.92, J.00] and are briefly described in section 3.2.3.

An additional improvement to the integration is achieved by smoothing the function of the integration i.e. reducing the 'variance' of the integrand. As the integration function is fixed by the problem, an equivalent procedure is to change the distribution of the point set within the phase space. The aim is to increase the density of points in phase space regions where the value of the integrand is expected to be large. This method is called 'importance sampling'. The importance sampling method is realized in the form of an analytic variable transformation in this analysis, made possible by the fact that the shape of the resolution functions is known to be Gaussian. The variable transformation is described in detail in section 3.2.4.

3.2.1 Monte Carlo Integration

The considered problem is the estimation of the multidimensional integral

$$J = \int_K f(x) d\mu(x), \tag{3.9}$$

where $x = x^\mu = (x^1, x^2, ..., x^s)$ denotes a point in the s-dimensional integration region K[2]. The integrand is $f(x)$, and $d\mu(x)$ denotes a measure on K. Therefore

$$\int_K d\mu(x) = 1, \tag{3.10}$$

and $d\mu(x)$ is positive. A series of points x can be found that their probability density is given by $d\mu(x)$:

$$\text{Prob}(x \in A) = \int_A d\mu(x), \tag{3.11}$$

for all small rectangular regions A. The integral J is estimated using the sum

$$S = \frac{1}{N} \sum_k f(x_k), \tag{3.12}$$

where the x_k $(k = 1, 2, ..., N)$ are a finite set of N points obtained in one or another way. The error made in using this estimate is

$$\eta = S - J. \tag{3.13}$$

One can derive for a sample of random points, independent and identically distributed with probability density $d\mu(x)$ from the Central Limit Theorem: The error η is normally distributed around zero, with standard deviation $\sqrt{V/N}$, while V denotes the variance of the integrand. The convergence of the error to zero by Monte Carlo integration is as $1/\sqrt{N}$ and additional smoothness properties only lead to a smaller error inasmuch as they lead to a smaller V (see 3.2.4).

3.2.2 Classical discrepancy

In one dimension the trapezoid rule method, where the points are equidistantly distributed, gives a $1/N^2$ convergence, better than the classical Monte

[2]Greek indices denote individual coordinates

Carlo method from the previous section. The reason for this lies in the uniformity of the distribution for the equidistant point set compared to the random point set. A measure for deviations of uniformity is the *classical discrepancy*, which allows to quantify irregularities, in any dimension.

The following counting function is defined for a point y in an s-dimensional hypercube K:

$$\chi(y;x) = \prod_{\mu} \Theta(y^{\mu} - x^{\mu}), \tag{3.14}$$

simply checks if the point x is inside the hyper-rectangle defined by the origin and y. The local discrepancy at y, for the point set x_k is then

$$g(y) = \frac{1}{N} \sum_{k} \chi(y;x_k) - \prod_{\mu} y^{\mu}. \tag{3.15}$$

The function $g(y)$ counts the fraction of the point set below y, and compares it with the volume of K that is below y. It is obvious that the more uniform the point set is, the smaller is $g(y)$. A global discrepancy, or the deviation of $g(y)$ from zero, is defined by:

$$D_m \equiv \int_K g(y)^m dy. \tag{3.16}$$

Useful measures of the global discrepancy are D_1 (linear),D_2 (quadratic) and the Kolmogorov discrepancy:

$$D_\infty = \lim_{k\to\infty} (D_{2k})^{1/2k} = \sup_{y\in K} |g(y)|. \tag{3.17}$$

For truly random point sets the following results can be derived:

$$\langle D_1 \rangle = 0 \quad , \quad \langle D_2 \rangle = \frac{1}{N}\left(\frac{1}{2^s} - \frac{1}{3^s}\right) \quad , \quad \langle D_\infty \rangle \overset{s=1}{=} \sqrt{\frac{\pi}{2N}} \log 2 \tag{3.18}$$

A common measure of quality for a certain point set is the quadratic discrepancy normalised to the quadratic discrepancy for truly random numbers, called the normalised quadratic discrepancy[3].

3.2.3 Niederreiter and Hammersly Generator

In a dedicated analysis different types of quasi random number generators were tested [J.00]. Based on these results the Hammersly [J.00] method

[3]This ratio is equal to one for random points; Any smaller value is equivalent to faster convergence

was chosen, which is a mixture of a linear grid and the Niederreiter [H.92] generator. A brief description of the generator and its performance is given here.

Starting point is the Halton generator [Hal60] in one dimension, by choosing a base, an integer b. Any integer n can be written in base b:

$$n = n_0 + n_1 b + n_2 b^2 + n_3 b^3 + ... \tag{3.19}$$

The so-called 'radical-inverse transform' (to base b) is defined by

$$\Phi_b(n) = n_0 b^{-1} + n_1 b^{-2} + n_2 b^{-3} + n_3 b^{-4} + ... \tag{3.20}$$

The Halton sequence to base b is then:

$$x_k = \Phi_b(k) \quad , \quad k = 1, 2, ... \tag{3.21}$$

A upper bound for the discrepancy D_∞ can be derived as follows:

$$D_\infty \leq C_b \frac{\log N}{N} \quad , \quad C_b = \begin{cases} \frac{b^2}{4(b+1)\log b} & \text{when b is even} \\ \\ \frac{b-1}{4\log b} & \text{when b is odd} \end{cases} \tag{3.22}$$

A generalisation of equation 3.21 to dimension s is achieved by choosing several bases $b_1, b_2, ..., b_s$ and

$$x_k^\mu = \Phi_{b_\mu}(k) \quad , \quad k = 1, 2, ... \quad , \quad \mu = 1, 2, ..., s \tag{3.23}$$

One clearly runs into trouble if any of two bases have common factors, so the common choice to avoid this, is to take the first s prime numbers as bases.

The Niederreiter generator is an improvement of the Halton generator by making use of additional functions $m_\mu(k), \mu = 1, 2, ..., s$ and defining a new sequence

$$x_k^\mu = \Phi_b(m_\mu(k)). \tag{3.24}$$

The functions $m_\mu(k)$ are cleverly chosen, amongst other things in such a way that when $N = b^m$, the numbers $m_\mu(1), m_\mu(2), ..., m_\mu(N)$ are just a permutation of $(1, 2, ..., N)$. For details see [H.92].

The Niederreiter generator is the optimal choice for a wide range of numbers N. In special cases, for certain values of N a slightly better discrepancy can be obtained. The Hammersly algorithm is constructed by a grid as the first component ($x_n^1 = \frac{n-1}{N}$) and Niederreiter sequences for all other components. It is obvious that this only works for $n = N$ set of points.

N	Halton D_2	Niederreiter D_2	Hammersly D_2
512	$1.119 \cdot 10^{-2}$	$5.815 \cdot 10^{-3}$	$4.844 \cdot 10^{-3}$
1024	$1.119 \cdot 10^{-2}$	$3.136 \cdot 10^{-3}$	$2.647 \cdot 10^{-3}$
2048	$5.544 \cdot 10^{-3}$	$1.896 \cdot 10^{-3}$	$1.453 \cdot 10^{-3}$
4096	$4.066 \cdot 10^{-3}$	$9.888 \cdot 10^{-4}$	$7.807 \cdot 10^{-4}$
8192	$2.076 \cdot 10^{-3}$	$5.493 \cdot 10^{-4}$	$4.220 \cdot 10^{-5}$

Table 3.1: The normalised quadratic discrepancies for the Halton, Niederreiter and the Hammersly generator. The discrepancy is calculated for dimension $s = 2$, lead to optimal point sets for $N = 2^l, l = 1, 2, 3....$

Table 3.1 illustrates the performance of the Halton, Niederreiter and Hammersly generator on the basis of the normalised quadratic discrepancy. The chosen dimension is $s = 2$ lead to an optimal point set for $N = 2^l, l = 1, 2, 3.....$ The numbers are taken from [J.00].

Detailed systematic studies on integration problems in multi-dimensions with quasi random numbers were done, including bias and correlation effects, in a separate work. A detailed discussion would go beyond this chapter. Further details can be found in [J.00].

3.2.4 Importance Sampling

It is derived from the Central Limit Theorem introduced in section 3.2.1, that the reduction of the integration error is achieved by using larger point sets or by the reduction of the variance of the integrand. As already stated in the introduction, the chosen technique for variance reduction used in this work is Importance Sampling as the shape of the integrand is basically known. The basic principle of importance sampling is illustrated in figure 3.2. The distribution of the points in the integration phase space is changed according to the 'importance' or magnitude of the integration function.

The problem in the later physics analysis (section 5.5) is of the following form:

$$I = \int_0^{l_{max}} \int_0^{p_{max}} f(l,p) R_l(l - l_m) R_p(p - p_m) dl dp \tag{3.25}$$

R_l and R_p are Gaussian resolution functions with widths σ_l, σ_p small compared to the whole phase space $[0..l_{max}, 0..p_{max}]$. The variable transformation optimized for this problem is described in detail in the next section.

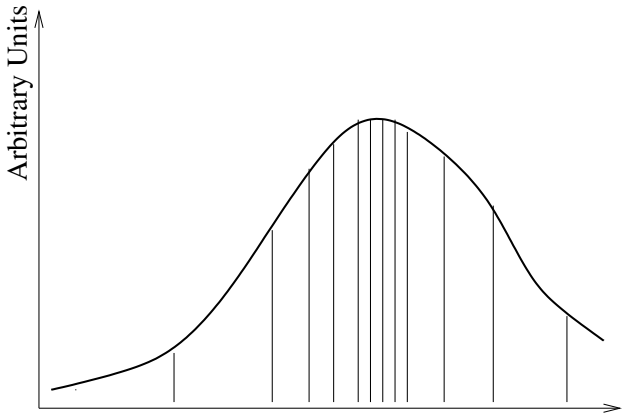

Figure 3.2: The basic principle of the Importance Sampling method. The points used for the integration are not distributed uniformly in the phase space. They follow the magnitude ('importance') of the integration function.

3.2.5 Variable Transformation

The transformation can be divided into two parts. The first is a shift to l_m and p_m as center of the phase space and an additional normalisation to the expected widths σ_l and σ_p. Applying the transformations,

$$v = \frac{l - l_m}{\sigma_l}, \quad \text{and} \quad w = \frac{p - p_m}{\sigma_p} \tag{3.26}$$

leads to the integral,

$$I = \int_{-l_m/\sigma_l}^{(l_{max}-l_m)/\sigma_l} \int_{-p_m/\sigma_p}^{(p_{max}-p_m)/\sigma_p} f(v\sigma_l + l_m, w\sigma_p + p_m) R(v\sigma_l) R(w\sigma_p) \sigma_l \sigma_p dv dw \tag{3.27}$$

with new integration variables v and w. The second step consists of a transformation following the basic shape of the phase space distribution. The chosen function is a 'Breit-Wigner' function of the form:

$$f(k) = \frac{1}{1 + k^2} \tag{3.28}$$

The reason of this choice is the similarity of a Breit-Wigner function to a Gauss function, as the Breit-Wigner has a peak shape with certain width and in contrast to the Gauss function an analytically inverse function exists for the Breit-Wigner. The inverse function of a Breit-Wigner is

$$f^{-1}(k) = \arctan k. \tag{3.29}$$

Finally the transformation has the form,

$$x = \frac{1}{\pi}(\arctan(v) + \frac{\pi}{2}), \quad \text{and} \quad y = \frac{1}{\pi}(\arctan(w) + \frac{\pi}{2}), \tag{3.30}$$

leading to the following equations for v and w,

$$v = \tan(\pi(x - \frac{1}{2})) \quad \text{and} \quad w = \tan(\pi(y - \frac{1}{2})). \tag{3.31}$$

After this transformation one gets an integral of the form,

$$\begin{aligned} I \;=\; & \int_{L_{down}}^{L_{up}} \int_{P_{down}}^{P_{up}} f(\tan(\pi(x - \frac{1}{2}))\sigma_l + l_m, \tan(\pi(y - \frac{1}{2}))\sigma_p + p_m) \cdot \\ & R(\tan(\pi(x - \frac{1}{2}))\sigma_l) \cdot R(\tan(\pi(y - \frac{1}{2}))\sigma_p) \cdot \\ & \sigma_l \sigma_p \pi^2 (1 + \tan(\pi(x - \frac{1}{2}))^2)(1 + \tan(\pi(y - \frac{1}{2}))^2) dxdy \end{aligned} \tag{3.32}$$

with integration bounds,

$$\begin{aligned} L_{down} &= \frac{1}{\pi}(\arctan(-l_m/\sigma_l) + \frac{\pi}{2}) \\ L_{up} &= \frac{1}{\pi}(\arctan((l_{max} - l_m)/\sigma_l) + \frac{\pi}{2}) \end{aligned} \tag{3.33}$$

$$\begin{aligned} P_{down} &= \frac{1}{\pi}(\arctan(-p_m/\sigma_p) + \frac{\pi}{2}) \\ P_{up} &= \frac{1}{\pi}(\arctan((p_{max} - p_m)/\sigma_l) + \frac{\pi}{2}) \end{aligned} \tag{3.34}$$

An illustration of the integration principle is visible in figure 3.3. The left plot shows the distribution of 512 points based on pseudo random numbers generated with the RANLUX [Jam94] generator, the middle plot 512 points from a quasi-random sequence based on a Hammersly-Niederreiter algorithm and the right plot the transformation result of these points using the described Importance Sampling method.

3.2.6 Test of the Integration Algorithm

The quality of the integration algorithm was calculated by a test function similar in shape and behaviour to the later physics problem. The following integral,

$$I = \int_{l=0.0}^{5.0} \int_{p=0.0}^{50.0} e^{\frac{-lm}{\tau cp}} \cos(b\frac{lm}{cp})\frac{m}{cp} R_p((p - p_m)/p) R_l(l - l_m) dpdl \tag{3.35}$$

was chosen with constants $m = 5.2789$, $\tau = 1.5$ and $c = 0.02997$. The parameter b is a frequency and was varied during the tests. The functions R_l

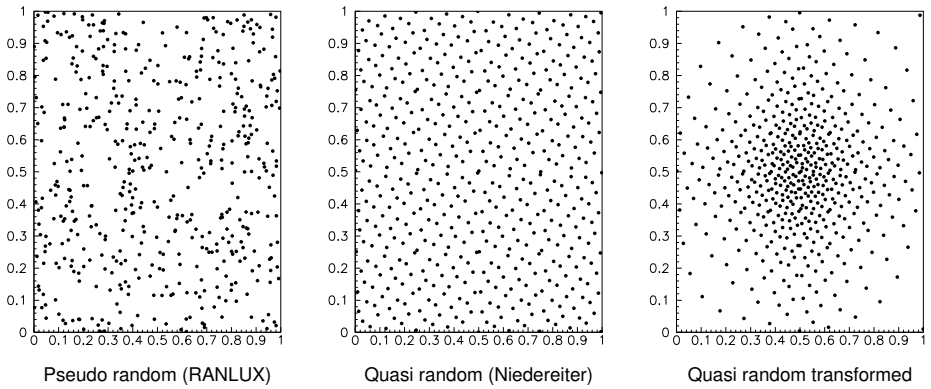

Figure 3.3: The left plot illustrates a two-dimensional distribution of 512 Pseudo Random points, the middle plot 512 points based on the Hammersly-Niederreiter algorithm and the right plot the variable transformation result of the Importance Sampling method.

and R_p are normalised Gaussian with width $\sigma_l = 0.02$ for R_l and $\sigma_p = 0.05$ for R_p. The true value of the integral was calculated with a trapezoid rule based algorithm using $6.4 \cdot 10^7$ points.

The integration procedure was tested using quasi random points sets with 512, 2048 and 8192 points based on the Hammersly algorithm and a normal grid of 8281 points. To every point set the importance sampling algorithm was applied. By permutation of the first component for the Hammersly sets it was possible to generate 500 individual point sets. The origin of the normal grid was shifted 500 times. For every point set k the difference D_k of the integral I_k with the truth I_{true} was calculated. A measure of the bias behaviour is given by the mean $\hat{D}$ of the difference D_k. The quality of the fit procedure itself is given by the variance σ_D^2 of D_k.

In table 3.2 the result of the test integrations is given for the different point sets with different parameters b. In the second column the mean $\hat{D}$, in the third column σ_D the square root of the variance and in the last column the relative error $\sigma_{rel} = \sigma_D / I_{true}$ is given. One can clearly see in the table that all point sets are bias free and that the Hammersly set has a smaller error than a grid set with the same size.

$B =$	Point Set	Mean $\hat{D}$	σ_D	σ_{rel}
1.0	$I = -1.38163 \cdot 10^{-1}$			
	512	$6.22 \cdot 10^{-7}$	$2.65 \cdot 10^{-5}$	$1.92 \cdot 10^{-4}$
	2048	$2.73 \cdot 10^{-6}$	$8.11 \cdot 10^{-6}$	$5.87 \cdot 10^{-5}$
	8192	$6.31 \cdot 10^{-7}$	$1.57 \cdot 10^{-6}$	$1.14 \cdot 10^{-5}$
	GRID	$2.62 \cdot 10^{-6}$	$5.77 \cdot 10^{-5}$	$4.18 \cdot 10^{-4}$
10.0	$I = -3.76541 \cdot 10^{-1}$			
	512	$1.84 \cdot 10^{-6}$	$1.11 \cdot 10^{-4}$	$2.93 \cdot 10^{-4}$
	2048	$3.22 \cdot 10^{-5}$	$5.45 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$
	8192	$4.69 \cdot 10^{-6}$	$8.12 \cdot 10^{-6}$	$2.16 \cdot 10^{-5}$
	GRID	$5.25 \cdot 10^{-7}$	$2.63 \cdot 10^{-5}$	$6.98 \cdot 10^{-5}$
20.0	$I = -1.90064 \cdot 10^{-2}$			
	512	$4.31 \cdot 10^{-6}$	$4.28 \cdot 10^{-4}$	$2.25 \cdot 10^{-2}$
	2048	$5.01 \cdot 10^{-6}$	$9.75 \cdot 10^{-5}$	$5.13 \cdot 10^{-3}$
	8192	$6.09 \cdot 10^{-6}$	$7.77 \cdot 10^{-6}$	$4.10 \cdot 10^{-4}$
	GRID	$6.05 \cdot 10^{-6}$	$1.29 \cdot 10^{-4}$	$6.79 \cdot 10^{-3}$
25.0	$I = -1.56606 \cdot 10^{-2}$			
	512	$5.14 \cdot 10^{-6}$	$9.21 \cdot 10^{-4}$	$5.90 \cdot 10^{-2}$
	2048	$4.24 \cdot 10^{-5}$	$2.52 \cdot 10^{-4}$	$1.62 \cdot 10^{-2}$
	8192	$1.78 \cdot 10^{-5}$	$1.10 \cdot 10^{-5}$	$6.90 \cdot 10^{-4}$
	GRID	$2.69 \cdot 10^{-6}$	$6.36 \cdot 10^{-5}$	$4.10 \cdot 10^{-3}$

Table 3.2: In the left column the different values for b are given. The associated top row provides the exact result of the integral I_{true}. In the second column the mean D_k, in the third column the deviation σ_D and in the last column the relative error $\sigma_{rel} = \sigma_D / I_{true}$ is given.

Chapter 4

B Hadron Reconstruction

The following chapter provides an overview of B hadron reconstruction. Most of the basic work is done by the BSAURUS program package developed in Karlsruhe over the last 5 years [ABF+01]. The chapter concentrates on the BSAURUS parts which are important for this analysis. The BSAURUS package is optimised for B hadron reconstruction but the enrichment of multihadronic decays containing a b-quark is provided by the official DELPHI b-tagging package (AABTAG) [BM96,BM95].

Due to the finite lifetime of B hadrons ($\approx$ 1.6 ps) together with their boost, a B hadron can fly a few millimetres before its decay ($< l > \approx$ 3) mm, which produces a secondary vertex displaced in space from the primary interaction vertex of the event. A data event ($Z \to b\bar{b}$) is plotted in figure 4.1, where both secondary vertices of the B hadrons are clearly visible. The b-tagging makes use of this by defining a so-called 'impact parameter' for every track, which is the distance between the point of closest approach of the track to the primary vertex. The impact parameter can be signed as either positive or negative depending on the projection of the impact parameter to the axis of the jet the track is assigned to. The impact parameter distribution for uds-quark events is centred around zero, while it is shifted towards positive values for b quark events (c quark events are also slightly shifted due to the finite lifetime of c hadrons). In combination with additional event information i.e. secondary vertex mass, a good estimator for b quark identification is constructed. A more detailed description of the b-tagging package is given in the first part of the chapter.

Two basic problems have to be solved by the BSAURUS package for this analysis. The first is the reconstruction of the B meson lifetime, which is achieved by the measurement of the B meson decay length and its momentum. The second is the determination of the B meson flavour (b or $\bar{b}$) at production and decay time.

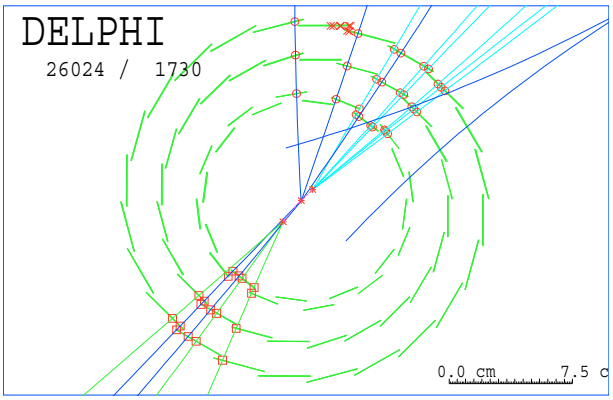

Figure 4.1: A $Z \to b\bar{b}$ event in the Delphi vertex detector. Both secondary vertices are significantly separated from the primary vertex in the middle.

The starting point for the Bsaurus package is the selection of multihadronic events, for which at least 5 charged tracks and some energy deposit in the detector is required. These events are split into two hemisphere, following their two main jets. A first estimate of the b hadron energy is obtained using the so-called 'Rapidity Algorithm'. In addition a sample of candidate tracks is provided by the Rapidity Algorithm, which are used for the reconstruction of a secondary vertex. Based on these quantities better estimators are calculated for the energy and the decay length of the B hadron candidate using neural network techniques. The first part of the chapter describes the different tools in detail.

The last part is dedicated to the problem of determining the B meson flavour at production and decay time. The basic idea is rather simple, but very effective: Every single track contains information about the quark flavour at its decay point by its electric charge, which is either correlated or anti-correlated to the B meson flavour (A positively charge track is defined correlated to a $\bar{b}$ quark ($Q = +1/3$) and anti-correlated to a b quark ($Q = -1/3$). Therefore different track level neural networks were developed, individually to recognize production flavour, decay flavour and B hadron species type. A tag at the level of the B hadron is finally achieved by combination of the track level quantities, using a likelihood ratio construction.

4.1 B Tagging

Starting point of the Delphi B tagging is a successful reconstruction of the primary vertex. The size and the position of the electron and positron packets are used as boundary condition. Together with tracks from the fragmentation process the position is estimated in a χ^2 fit with a precision of 10 μm in x

and 5 μm in y direction (see [B.96,PFM85] for details).

With the help of this primary vertex one can calculate the impact parameter significance S for each individual track. S is defined as closest approach of the track to the primary vertex in the $R\phi$ plane divided by its error. The significance S is signed positive (negative) if the projection $\vec{P}$ of the impact parameter on the track associated jet axis is in positive (negative) jet direction. This leads to positive impact parameters for long lived particles. Fragmentation tracks have a Gaussian distribution around $S = 0$ due to the limited detector resolution. An N-track probability P_N can be derived from the track significances S. The probability P_N is smaller for events with a long lived particle (e.g. b quark and c quark decays), compared to uds production.

The b tagging algorithm can be improved by making use of the expected jet structure of the b quark and additional discriminating variables [Bor]. Therfore the jet structure of the event was resolved using the JADE [Sjo94] algorithm with jet resolution parameter $y_{cut} = 0.01$. The following variables are used:

Impact parameter significance The most important variable is the N track probability P_N (P_j^+ in the plot) based on the impact parameter significance described in the previous section.

Invariant mass The four-momentum sum of all tracks associated to a jet is calculated. The invariant mass M_s is derived from the jet four-momentum. For c quarks a sharp drop at $M_s = 1.8\,Gev/c^2$ is expected, while the distribution extends up to $M_s = 5\,Gev/c^2$ for b quarks.

Rapidity The rapidity R_s^{tr} is expected to be bigger for c quarks compared to b quarks due to the lower mass and the lower multiplicity.

Charged energy The charged energy fraction X_s^{ch} of tracks from the secondary vertex compared to all tracks in the jet is usually higher for b quarks compared to light quarks.

These variables x_i are combined to a single b quark estimator y in the following way:

$$y = \prod_{i=1}^{n} \frac{f_i^B(x_i)}{f_i^S(x_i)} = \prod_{i=1}^{n} y_i \tag{4.1}$$

where $f_i^B(x_i)$, $f_i^S(x_i)$ are probability density functions for background and signal for variable x_i. The events with $y < y_0$ are tagged as a signal. The value y_0 can be varied to select a desired purity or efficiency of the tagging.

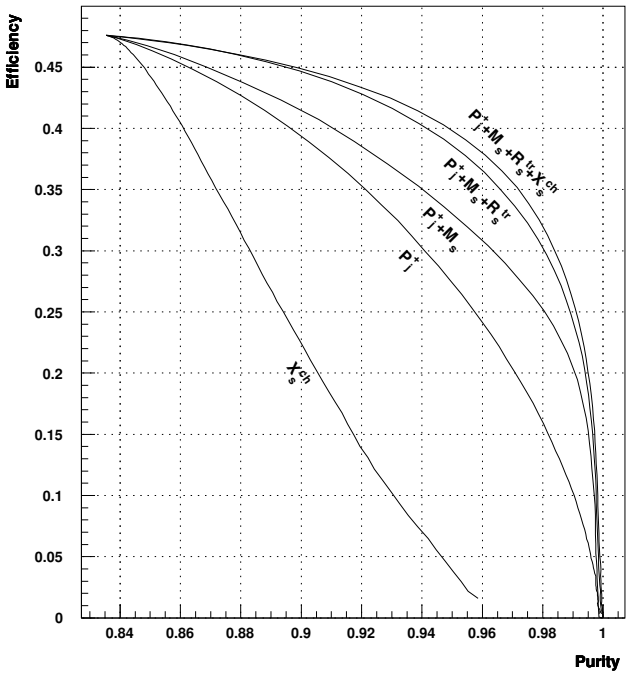

Figure 4.2: The left plot illustrates the purity efficiency behaviour of the combined DELPHI b tagging calculated for quark jets. By adding extra information from invariant mass M_s, rapidity R_s^{tr} and charged energy X_s^{ch} to the pure impact parameter tag, the purity is improved dramatically.

In figure 4.2 the purity vs. efficiency curve for the individual contributions is plotted. One can see that especially the inclusion of the rapidity and the invariant mass improves the sample at very high purities. All plots in the following part of the chapter were made on a event sample with a b-tagging precut, which guarantees a 90 % $b\bar{b}$ event purity.

4.2 Bsaurus

B hadrons may have, due to their large mass, thousands of decay channels, all with small branching fractions. The exclusive reconstruction of dedicated decay channels is the standard procedure for B hadron physics. This however results in very limited statistics and hence to a severe limitation for studying B physics. The idea of BSAURUS is the inclusive reconstruction of as many properties of b-jets as possible with high efficiency and good purity. This is achieved by exploiting the capabilities of the DELPHI detector to their maximum, applying wherever possible physics knowledge about B production and decays and combining different information sources with modern tools - mainly neural networks.

The following section gives a short overview of the event processing steps which are done by BSAURUS.

4.2.1 Hadronic Selection

The event selection of BSAURUS is very loose to provide enough freedom for individual analyses to apply their aim selection. However a minimal hadronic selection is necessary to make sure that all quantities can be calculated successfully. Multihadronic Z^0 decays were selected using the following requirements:

- at least 5 reconstructed charged particles,
- the energy sum of charged particles (with momentum > 0.2 GeV/c) has to be larger than 12% of the center-of-mass energy,
- at least 3% of the charged energy sum has to be in each of the forward and backward hemispheres defined with respect to the beam axis.

The same multihadronic selection criteria are required for the Monte-Carlo sample, based on fully simulated $Z \to q\bar{q}$ events using JETSET 7.3 [Sjo94].

4.2.2 Event Hemispheres

Each event is split into two hemispheres using the plane perpendicular to the thrust axis. In addition, each hemisphere is assigned a reference axis defined in the following way:

- In the case of a two jet event, the reference axis for the hemisphere is the jet axis in that hemisphere.
- If a hemisphere contains 2 or more jets (i.e. an event with 3 or more jets):
  - if one of the jets is the highest energy jet in the event, that jet axis forms the reference axis.
  - The combined b-tag probability is calculated for the jets in the hemisphere. The jet that is most 'B-like' is then selected to form the reference axis.

The rapidity of a track is defined as follows,

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) \tag{4.2}$$

for E the track energy and p_L the longitudinal momentum component along the reference axis for the hemisphere-

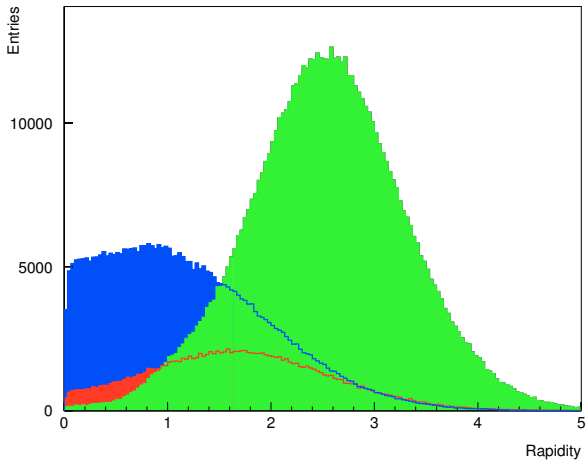

Figure 4.3: The B hadron rapidity distribution. The three different shaded areas indicate particles origin from a weakly decaying B hadron (light grey), fragmentation partner of the B hadron (dark grey) and from the fragmentation (black).

A first estimate of a B-candidate four-momentum vector is returned by the so-called 'Rapidity Algorithm'. This is defined to be simply the sum of individual track momentum vectors in a hemisphere for tracks with rapidity greater than 1.6. The cut at rapidity 1.6 selects tracks from the B hadron decay with an efficiency of 89 % and a sample purity of 77 %. The rapidity distribution for fragmentation, leading fragmentation and B hadron decay tracks in the simulation is shown in figure 4.3.

4.2.3 B Energy Reconstruction

The B hadron energy reconstruction is mainly based on a correction function for the raw energy E_{raw}, derived from the rapidity algorithm, compensating energy loss due to physical and detector effects. Separate correction functions are derived for 2-jet and more than 2-jet events.

To be used in the correction procedure, hemispheres must pass the following cuts:

- an initial minimum reconstructed B-candidate energy E_{raw} of 20 GeV
- the initial reconstructed B-candidate mass m_{raw} derived from the initial B hadron four-momentum estimate, lies within two standard deviations of the total data sample median value
- the ratio x_h of the hemisphere energy E_{hem} to beam energy E_{beam} lies in the range, $0.6 < x_h < 1.1$.

The starting point is an estimate of the B energy and mass, E_{raw}, m_{raw}. From Monte Carlo studies, these estimates are chosen to be from the rapidity algorithm for events with $>$ 2-jets and to be derived from the sum of 'B-weighted' four-vectors for the 2-jet case. This involves weighting (via a sigmoid threshold function) the momentum and energy components of charged tracks. The value used by the weighting is derived from a neural network, which separates fragmentation tracks from B hadron decay tracks (see section 4.2.4). Neutral particles are weighted by their rapidity. In this way the effect of tracks from the B decay are enhanced and tracks from the primary vertex are suppressed in the summation.

The correction procedure is motivated by the observation in Monte Carlo of a correlation between the energy residuals $\Delta E = E_{raw} - E_B^{truth}$ and m_{raw} (which is approximately linear in m_{raw}) and a further correlation between ΔE and x_h resulting from neutral energy losses and inefficiencies. The correction proceeds in the following way: The data are divided into several samples according to the measured ratio x_h and for each of these classes the B energy residual ΔE is plotted as function of m_{raw}. The median values of ΔE in each bin of m_{raw} are calculated and their m_{raw} dependence fitted by a third order polynomial

$$\begin{aligned} \Delta E(m_{raw}; x_h) = a &+ b(m_{raw} - \langle m_{raw} \rangle) \\ &+ c(m_{raw} - \langle m_{raw} \rangle)^2 \\ &+ d(m_{raw} - \langle m_{raw} \rangle)^3 \end{aligned} \tag{4.3}$$

The four parameters a, b, c, d in each x_h class are then plotted as a function of x_h and their dependence fitted with third and second-order polynomials. Thus one obtains a smooth correction function describing the mean dependence on m_{raw} and the hemisphere energy as determined from the Monte Carlo. Finally, a small bias correction is applied for the remaining mean energy residual as a function of the corrected energy.

In addition, for the case of hemispheres with large missing energy i.e. $x_h < 0.6$ a separate correction with different parameters a, b, c and d is derived.

The procedure described above leads to a single estimate of the B hadron energy. The resolution of the reconstructed B hadron energy is shown in figure 4.4. Fitting a double Gaussian[1] to the distribution leads to widths $\sigma_1 = 0.065$ and $\sigma_2 = 0.18$ for the inner and outer Gaussian.

[1] A double Gaussian is defined as sum of two independent Gaussians

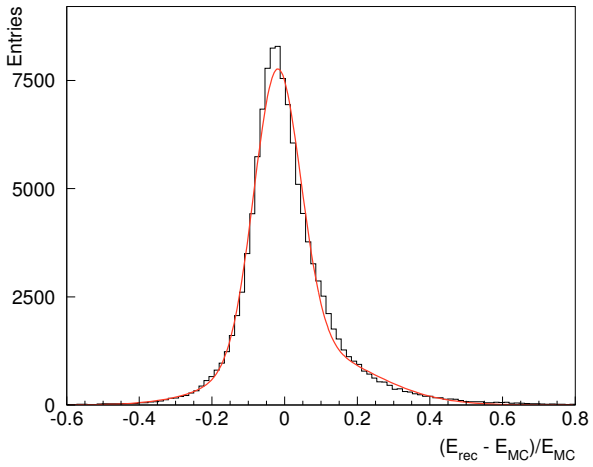

Figure 4.4: The resolution of the reconstructed B hadron energy. A double Gaussian is fitted to the distribution with widths $\sigma_1 = 0.065$ and $\sigma_2 = 0.18$ for the inner and outer Gaussian respectively.

4.2.4 1st Level Vertex Reconstruction

In each hemisphere an attempt is made to fit a secondary vertex to tracks with rapidity > 1.6 passing the following track selection criteria:

- impact parameter in the $r\phi$ plane $|\delta_{r-\phi}| < 4.0$ cm
- impact parameter in the z plane $|\delta_z| < 6.0$ cm
- $|\cos\theta| < 0.94$
- $\frac{\Delta E}{E} < 1.0$
- at least 1 $r-\phi$ track hit registered in the silicon vertex detector(VD).

To this class of tracks additional criteria (i.e. track is identified as kaon or lepton) are applied with the aim of selecting tracks for the vertex fitting stage that are likely to have originated from the decay chain of a weakly decaying B hadron state. Using the selected tracks a secondary vertex fit is performed in 3-dimensions (DAPLCON χ^2-fit routine from ELEPHANT [FKP96]) constrained to the direction of the B-candidate momentum vector (see section 4.2.3). The event primary vertex is used as a starting point and if the fit did not converge[2], the track making the largest χ^2 contribution is stripped away in an iterative procedure, and the fit is repeated.

[2]Here, non-convergence means the fit took more than 20 iterations. A further iteration is deemed necessary if the χ^2 is above 4 standard deviations during the first 10 iterations or above 3 standard deviations during the next 10 iterations.

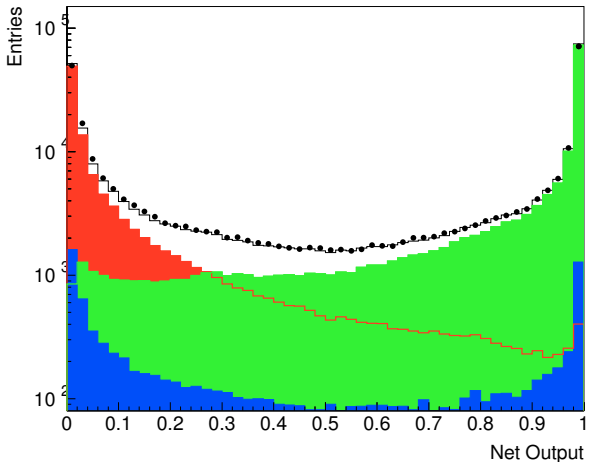

Figure 4.5: The TrackNet output distribution for Monte Carlo compared to data. The shaded distribution on the right illustrates tracks from the B decay chain, the left distribution tracks from the fragmentation process. The dark distribution is from non-b decays of the Z^0.

Tracks originating from the weak decay of the B hadron are separated from all other tracks with a dedicated neural network (so-called 'TrackNet'). The network calculates for every track the probability of the track originating from the weak B decay. For this purpose the information from the vertex fit, described in the previous section, is combined with the momentum, rapidity, helicity angle and other information of the track into the network. The network output of the TrackNet is illustrated in figure 4.5. One can clearly see the good separation power between b decay and fragmentation tracks.

Once a convergent fit has been obtained, the final stage of the secondary vertex fitting procedure involves an attempt to add tracks into the fit that failed the initial track selection criteria but nevertheless are consistent with originating from the vertex. The TrackNet output is the additional information used to add any remaining B hadron decay candidate tracks to the secondary vertex definition. The track of largest TrackNet in the hemisphere is added to the existing track list and retained if the resulting fit converges. This process continues iteratively for all tracks with TrackNet > 0.5.

4.2.5 Vertex Reconstruction

A precise proper time reconstruction is one of the key elements for a B_s oscillation analysis. Therefore a very good B hadron decay length reconstruction is necessary. In figure 4.6 the decay length resolution of the standard BSAURUS vertex algorithm of the previous section is shown. Clearly visible is the rather large bias in the direction of bigger decay length. The main rea-

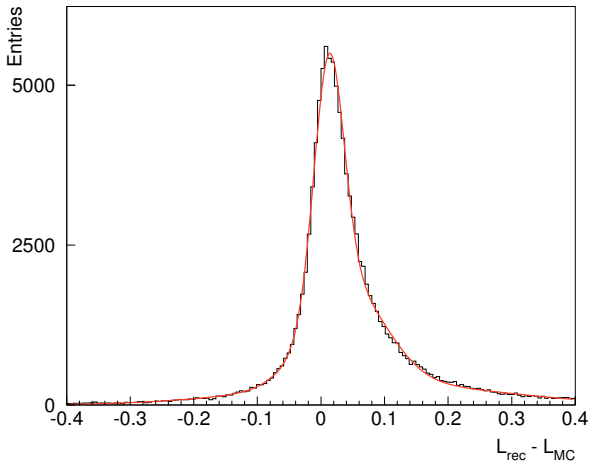

Figure 4.6: The decay length resolution distribution for the standard vertex algorithm of BSAURUS. The rather large forward bias is clearly visible. Three Gaussians were fitted to the distribution with:

$$\bar{l}_1 = 120\,\mu\text{m} \quad \sigma_1 = 250\,\mu\text{m}$$
$$\bar{l}_2 = 410\,\mu\text{m} \quad \sigma_2 = 600\,\mu\text{m}$$
$$\bar{l}_3 = 930\,\mu\text{m} \quad \sigma_3 = 1870\,\mu\text{m}$$

son for this behaviour is the cascade D decay[3] of the B hadron leading which will 'pull' the decay length longer if tracks from the D decay are mistakenly included in the vertex fit.

As this is not optimal for the B_s analysis, two dedicated algorithms for a better vertexing performance, described in the next section, were developed.

The BD-Net

As stated in the previous section the main problem of the standard vertex algorithm is the incorporation of tracks from the cascade D decay. In analogy to the TrackNet (see section 4.2.4) a neural network designed to discriminate between tracks originating from the weakly decaying B hadron and those from the subsequent cascade D meson decay was constructed.

The following input variables are used for the BD-Net:

- The angle between the track vector and the estimate of the B flight direction derived from the B hadron four-momentum vector (see section 4.2.3).
- The probability that the track originates from the fitted primary vertex (AABTAG algorithm).
- The probability that the track originates from the fitted secondary vertex (AABTAG algorithm).
- The momentum and angle of the track vector in the B rest frame.

[3] A B meson usually decays to a D meson: $B \to D \to K$. This is called a decay cascade

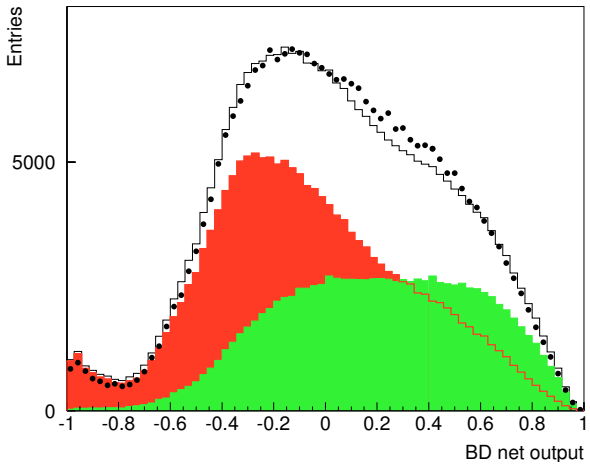

Figure 4.7: Output of the BD-Net for tracks in simulation (histogram) and data (points) for tracks with TrackNet > 0.5. The component histograms show the distribution for cascade D-decay tracks (light grey) and B decay tracks (dark grey).

- The TrackNet output.
- Kaon identification information (KaonNet see [AFM99] for details).
- Lepton identification information.

Some of the variables carry no direct discrimination power but are included as gauges of the quality of the other variables:

- A combined hemisphere variable containing information about the quality of the hemisphere, i.e. number of ambiguous vertex detector hits and number of tracks with no TPC information.
- A combined track quality variable containing i.e. number of vertex detector hits of the track.
- The hemisphere decay length significance L/σ_L.
- The hemisphere secondary vertex mass.
- The hemisphere rapidity gap between the track of highest rapidity below a TrackNet cut at 0.5 and the smallest rapidity above the cut at 0.5.

The output of the BD-Net is shown in figure 4.7 for simulation and data for tracks with TrackNet > 0.5. The two classes the network was trained on are shown in the histogram, namely tracks originating from cascade D-decays (light grey) and all other tracks which are mainly tracks from the B decay

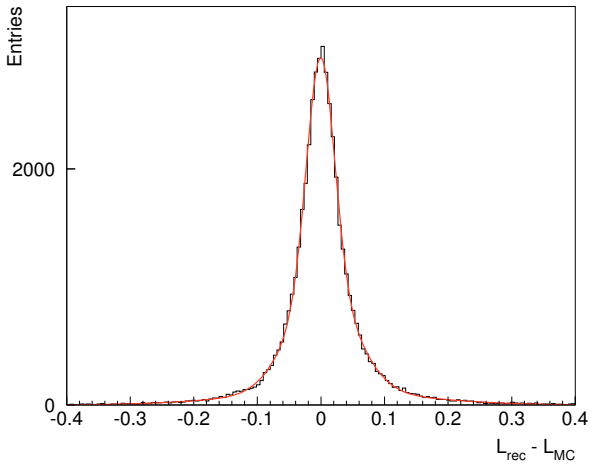

Figure 4.8: The decay length resolution distribution for the Strip-Down vertex algorithm. Three Gaussians were fitted to the distribution with

$$\bar{l}_1 = -5\,\mu\text{m} \quad \sigma_1 = 216\,\mu\text{m}$$
$$\bar{l}_2 = 21\,\mu\text{m} \quad \sigma_2 = 510\,\mu\text{m}$$
$$\bar{l}_3 = 110\,\mu\text{m} \quad \sigma_3 = 1426\,\mu\text{m}$$

(dark grey). The main reason for the small discrepancy between data and simulation are the different D branching ratios which are not known with high precision.

Strip-Down Vertex Fit

The basic idea of the Strip-Down Algorithm is the selection of candidate tracks most likely to come from the weakly decaying B. Therefore tracks are selected if they have TrackNet output bigger than 0.5 and BD-Net value less than -0.1. This BD-Net cut corresponds to 50 % efficiency in selecting a track from a weakly decaying B hadron with 75 % purity. Monte Carlo studies have shown that this BD-Net cut gives an optimal decay length resolution for the Strip-Down fit.

A secondary vertex χ^2-fit (DAPLCON routine) is made if there are 2 or more tracks selected. If the fit fails to converge within the algorithm criteria (same criteria as in the 1st level fit) and more than two tracks were originally selected, the track with highest χ^2 contribution is removed[4] and the fit repeated. This procedure is done iteratively until convergence is reached or two tracks are left. The fit is constrained by the direction estimated from the B hadron energy reconstruction and the starting point of the fit is the initial vertex position given by the standard algorithm.

The decay length resolution of the fit is shown in figure 4.8. One can see that the distribution is centred very nicely around 0. Also a fit of three

[4]This stripping procedure of tracks from the fit is responsible for the name of the algorithm

Gaussians indicates that the resolution is nearly bias free. The mean values are $\bar{l}_1 = -5\,\mu\text{m}, \bar{l}_2 = 21\,\mu\text{m}, \bar{l}_3 = 110\,\mu\text{m}$. The errors are $\sigma_1 = 216\,\mu\text{m}, \sigma_2 = 510\,\mu\text{m}, \sigma_3 = 1426\,\mu\text{m}$.

Indeed the nice performance of the algorithm does not come for free. In a lot of events less than two candidate tracks are selected or the fit with the selected tracks did not converge. This leads to an overall efficiency for this algorithm of around 40 % depending slightly on event selection criteria. Therefore a second algorithm with higher efficiency was developed. This is described in the next section.

Build-Up Vertex Fit

In the Build-Up method those two tracks with TrackNet bigger than 0.5 and smallest BD Net values are chosen to form a seed vertex. The two selected tracks have the highest probability to come from a weakly decaying B hadron. If the invariant mass of all remaining tracks with TrackNet$>$ 0.5 exceeds the D mass[5], that track with the lowest BD Net output is added to the seed vertex definition. This is done iteratively until the remaining mass drops below the D Meson mass. Finally all the candidate tracks are fitted to a vertex.

The approach with the D Meson mass is done to make sure that all tracks not belonging to the D vertex are taken into account for the B hadron vertex fit. As a result of the fit its decay length resolution is plotted in figure 4.9. It is clearly visible that this method leads to a bigger bias in forward direction than the Stripping method. A sum of three Gaussians is fitted to the distribution. The core Gaussian has a mean of $\bar{l}_1 = 20\,\mu\text{m}$ and an error $\sigma_1 = 238\,\mu\text{m}$, which represents the events where the fit worked as planned. The second Gaussian has a mean $\bar{l}_2 = 124\,\mu\text{m}$ and error $\sigma_2 = 589\,\mu\text{m}$. In these events the ordering of tracks within the BD-Net is not correct, lead to the incorporation of cascade D tracks in the fit and thus to a forward bias.

The advantage of the Build-Up fit method is the good overall efficiency of around 78.5 % compared to approximatly 40 % for the Strip-Down fit.

Comparing the different Vertex Fits

A direct comparison of the individual vertex algorithms is shown in figure 4.10. In each plot only the subset of events was used where both vertex algorithms converged. This inhibits that the effects of the advanced algorithms are only due to an improved event selection. One can clearly see in the left

[5] A D mass of 1.7 GeV/c is chosen.

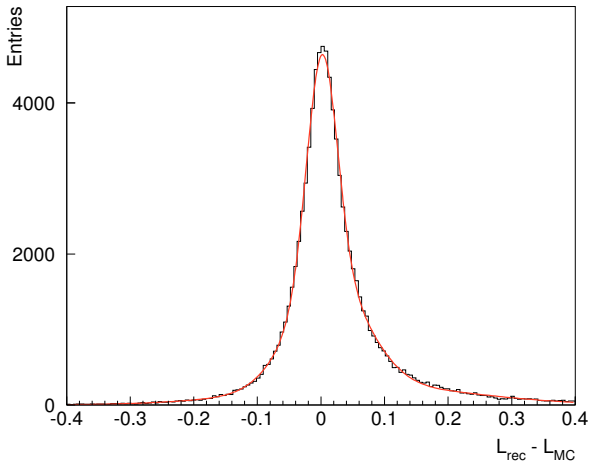

Figure 4.9: The decay length resolution distribution for the Build-Up vertex algorithm. Three Gaussians were fitted to the distribution with

$\bar{l}_1 = 20\,\mu\text{m} \quad \sigma_1 = 238\,\mu\text{m}$
$\bar{l}_2 = 124\,\mu\text{m} \quad \sigma_2 = 589\,\mu\text{m}$
$\bar{l}_3 = 770\,\mu\text{m} \quad \sigma_3 = 1657\,\mu\text{m}$

plot that the Strip-Down Fit removes the forward bias. Also visible is a slight improvement of the resolution.

The same effects, in a weaker way, are also visible in the right plots for the Build-Up Fit. Also the widths of the Build-Up distribution is nearly unchanged compared to the Standard Vertex.

As stated in the previous sections the overall efficiency for the Strip-Down Algorithm is 40 % and 78.5 % for the Build-Up method. The combined efficiency of requiring either one of the two is 81 %, which shows that around 20 % of the events are not accessable to either one of the algorithms. Monte Carlo studies have shown that those events often have badly reconstructed tracks and their resolution within the Standard Algorithm is bad. These events have a mean resolution of approximately 500 μm with a huge tail and are not usable in the analysis.

4.2.6 Flavour Tagging

The flavour tagging as one of the key elements is constructed in a special way. The BSAURUS approach is tagging the b quark charge (equivalent to the flavour) by first constructing a track probability and then combining them to give a probability at the hemisphere level.

Different networks were trained separately for tracks originating from fragmentation and decay. Also separately trained were the different B hadron types (B^+, B_d^0, B_s and B baryon) leading to 8 individual networks in total.

Combining the B species dependent flavour networks into a single flavour estimator, i.e. the production flavour, requires knowledge about the produced

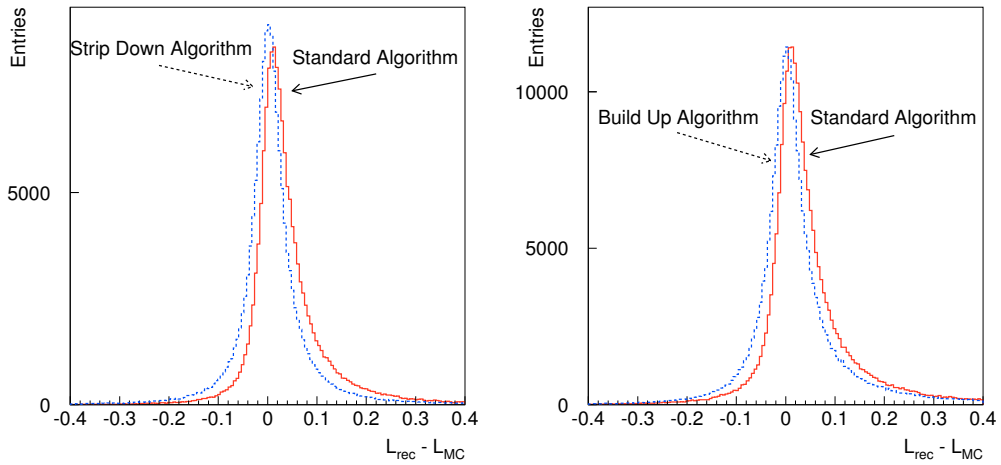

Figure 4.10: The left plot shows a comparison of the Strip-Down and the Standard Algorithm for the same set of events. It is clearly visible that nearly all forward bias is removed and the width of the distribution is also improved. The right plot shows the same for the Build-Up Algorithm. The effects are weaker, but an improvement of the forward bias is also visible. The width of the distribution is nearly unaffected.

B species. This is achieved by the B hadron identification network described in the next section.

B Hadron Identification Networks

The following section describes an attempt to decide in each hemisphere whether the hemisphere contains a B_s meson, B^+ or B^- mesons, a B_d^0 meson or a B baryon.

For this purpose a neural network consisting of 15 input nodes, described below, 17 hidden nodes and 4 output nodes was trained. Each output node delivers a probability for one of the four hypotheses it was trained on. The following input variables are used:

- based on the TrackNet output a probability P_B for each track coming from the B hadron or from the primary vertex is obtained. The weighted vertex charge, $\sum_i^{tracks} P_B(i) \cdot Q(i)$, distinguishes between charged and neutral B hadrons.

- the binomial error on the vertex charge defined as, $\sum_i^{tracks} \sqrt{P_B(i)(1-P_B(i))}$. This input gives a measure for the

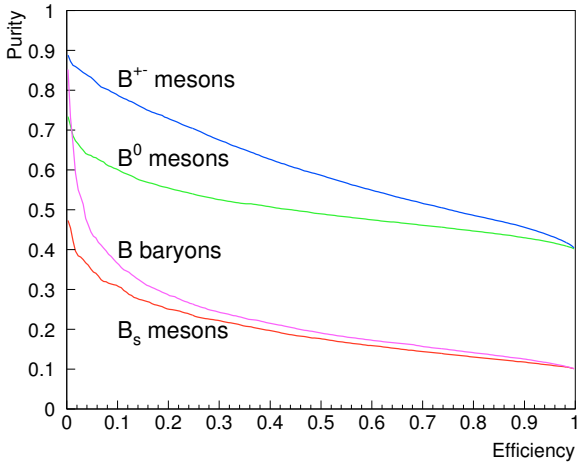

Figure 4.11: The purity vs. efficiency behaviour for the different B hadrons. The best performance is achieved for charged B hadrons. The weakest performance is given by the B_s network for which the B_d^0 background is nearly irreducible.

reliability of the vertex charge information.

- the number of charged 'pions' (tracks not identified as proton, kaon or lepton are called 'pion') in the hemisphere. This is most powerful for the case of B baryons and B_s mesons, which have a higher content of non-pion particles i.e. neutrons, protons and kaons in comparison to other B-species.

- the energy deposited in the hemisphere is sensitive to the presence of B baryons and B_s mesons due to the fact that associated neutrons and K_L^0 are often not reconstructed in the detector with the consequence that the total hemisphere energy tends to be smaller compared with the corresponding value obtained for B_d^0 or B^+ mesons.

- B_s mesons are normally produced with a Kaon as leading fragmentation particle with a further Kaon emerging from the weak decay (The same applies to the associated production of protons in jets with B baryons). Using this fact four input variables are constructed, one for each B hadron type, giving a measure of the likelihood for the presence of a leading fragmentation kaon/proton and a weakly decaying kaon/proton. For each track, the rapidity, the proton and kaon probability and the vertex position is taken into account in the calculation of the individual variable.

- the leading fragmentation Kaon can often be neutral in B_s meson production. This input uses information from neutral K^0's, reconstructed from decay tracks.

- a quality estimator for the whole hemisphere is calculated based on e.g. the number of reconstructed secondary interactions and tracks containing detector hits that could equally well fit to other tracks in the vertex detector. This provides the network with a measure of the quality of the input information.

- an additional approach, independent to the one mentioned before, is made to get a measure for the 'kaonness' of the leading fragmentation particle. The maximum of the Kaon net output for the three tracks with highest rapidity coming from the primary vertex is calculated.

- a charge correlation between the leading fragmentation particle charge and the secondary vertex charge provides a hint for the presence of B^+ mesons.

  In addition, input variables that give no inherent separation power of different B species were included to inform the network of the potential quality of the other input variables:

- the invariant mass of the reconstructed vertex.

- the energy of the B hadron to provide information on how hard the fragmentation was and therefore inform the network of how the available energy is expected to be shared between B hadron and fragmentation products.

The network was trained with an equal number of B hadrons of each species.

In figure 4.11 the efficiency vs. purity curves for the individual species are given. The best performance is achieved for charged B hadrons, while B_s mesons are hard to separate from the B_d^0 background.

Track Level Flavour Tagging

The track flavour networks are constructed to provide conditional probabilities for the tracks to have the same charge as the b quark in the B hadron. Therefore the network is trained with target value $+1(-1)$ if the track charge is correlated(anti-correlated) with the b quark charge. The following input variables are used:

- Particle identification: Kaon, Proton and Electron Net (see [AFM99] for details). Muon classification code (see [H^+,B^+91]).

- B-D Vertex separation (only used for decay flavour nets): BD Net output, BD Net output minus minimum BD Net value above TrackNet 0.5 and the track momentum in the B hadron rest-frame.
- Track level quality: Helicity angle in the B rest-frame, track quality flag, TrackNet output and track energy.
- Hemisphere level quality: Rapidity gap, secondary vertex mass, secondary vertex χ^2 fit probability, B energy and error on vertex charge measurement.

In total the track decay flavour network uses 19 input variables, while the track fragmentation network uses 16 variables (The BD information is not valid for fragmentation tracks).

The resulting track charge correlation conditional probabilities $P(Q_{same}|B_s)$ and $P(Q_{same}|B_d^0)$ optimised for B_s and B_d^0 mesons are plotted in figure 4.12. The plots illustrate the comparison between simulation and data.

Fragmentation and Decay Tags

Hemisphere level fragmentation and decay flavour tags are obtained from the track probabilities $P(Q_{same}|j)^k$, where $j = B^+, B_d^0, B_s$ or B baryon and $k =$ *fragmentation* or *decay*, by combining them into a likelihood ratio,

$$F(hem)_j^k = \sum_{tracks} ln\left(\frac{1+P(Q_{same}|j)^k}{1-P(Q_{same}|j)^k}\right) \cdot Q(track) \tag{4.4}$$

where Q(track) denotes the track charge. The tracks used in the likelihood sum are tracks with TrackNet > 0.5 for the decay flavour hypothesis and tracks with TrackNet < 0.5 for the fragmentation flavour hypothesis.

Figure 4.13 shows the hemisphere decay flavour networks optimised for B_s mesons (left plot) and B_d^0 mesons (right plot). The better performance of the B_d^0 net is clearly visible due to the fact that this network uses the information of the Kaon in the B_d^0 decay. This information can not be used in B_s decays as there are two different kaons with opposite charges.

4.2.7 Optimal Production Tag

The final step of flavour tagging is a general production flavour tag. Therefore all available information on the decay and fragmentation flavour is combined into a single network, in order to tag the B hadron quark flavour at production time.

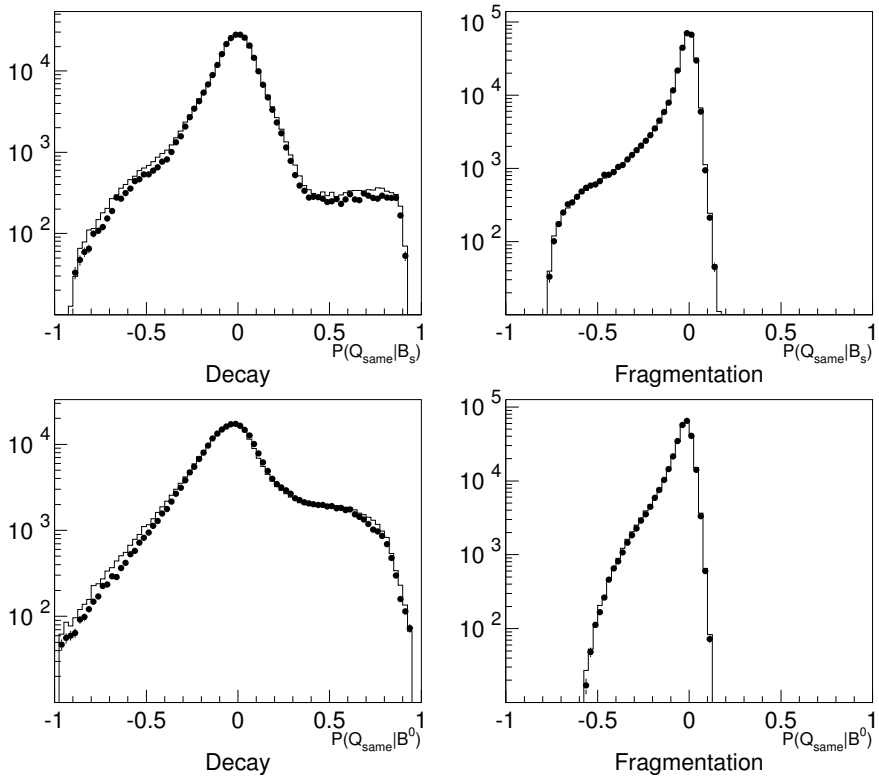

Figure 4.12: The conditional track-level decay and fragmentation probabilities optimised for B_s (top row) and B_d^0 (bottom row) mesons. The tracks plotted in the decay flavour plots on the left have TrackNet > 0.5, the fragmentation plots on the right TrackNet < 0.5.

The 9 input variables used for the network are:

- $F(hem)_{B_s}^{Frag} \cdot P_{BSpecies}(B_s)$.[6]
- $(F(hem)_{B^+}^{Decay} - F(hem)_{B^+}^{Frag}) \cdot P_{BSpecies}(B^+)$
- $(F(hem)_{B_d^0}^{Decay} \cdot (1 - 2\sin(\frac{1}{2}\Delta m_d \tau)^2) - F(hem)_{B_d^0}^{Frag}) \cdot P_{BSpecies}(B_d^0)$, τ is the reconstructed B lifetime calculated from the decay length of the 1st level vertex fit and the B hadron momentum estimate. This

[6] $P_{BSpecies}(B_s)$ denotes the B_s mesons probability given by the B hadron identification network.

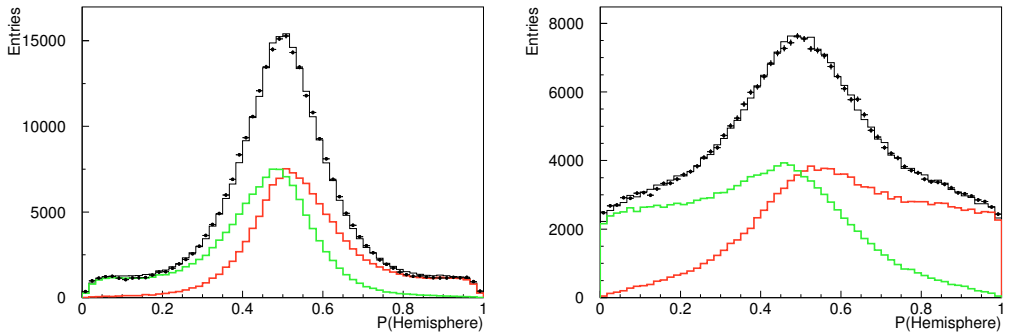

Figure 4.13: The left plot shows the B_s decay flavour net output, the right the B_d^0 flavour output, both for data and simulation. It is clearly visible that the performance of the B_d^0 net is much better; This can be understood from the fact that the charge of the Kaon in the B_d^0 decay is useful while it is not in the B_s decay.

construction attempts to take account of the B_d^0 oscillation frequency Δm_d.

- The jet charge defined as,

$$Q_J = \frac{\sum |P_L|_i^\kappa Q_i}{\sum |P_L|_i^\kappa} \tag{4.5}$$

where the sum is over all tracks and P_L is the longitudinal momentum component with respect to the thrust axis. Three different κ values were chosen, forming individual input variables ($\kappa = 0.3, 0.6, \infty$ ($\kappa = \infty$ selects the charge of the track with highest momentum in the hemisphere)).

- The TrackNet weighted vertex charge and its significance.

The result of the network is plotted in figure 4.14 for data and simulation. Drawn in light and dark grey are the distributions for b and $\bar{b}$ quarks. The tagging purity is approximately 73 % at 100 % efficiency.

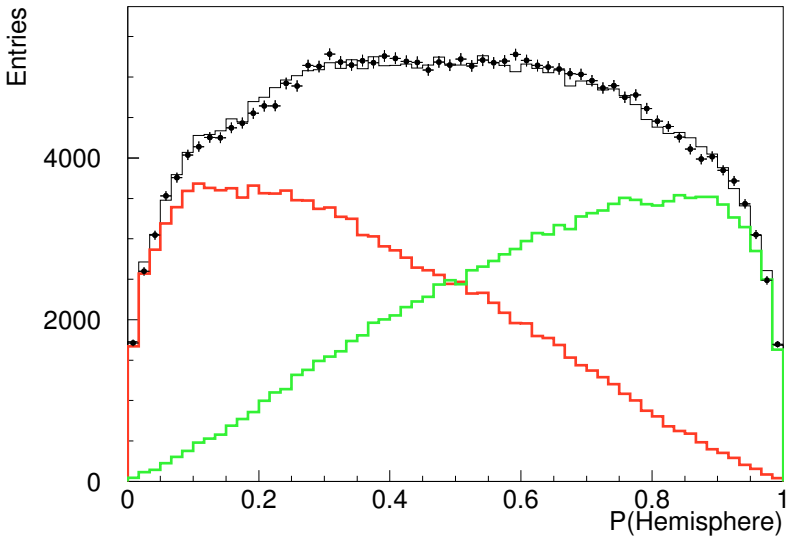

Figure 4.14: The plot illustrates the optimal production flavour net distribution for b and $\bar{b}$ quarks. The tagging power is around 73 % purity at 100 % efficiency.

Chapter 5

Analysis

This main analysis chapter can be divided into five parts: All the ingredients are put together in the first three parts, i.e. event selection, classification of decay length, momentum and flavour tags. The fitting method ('Method of Maximum Likelihood') is explained in part four and part five describes extracting the results.

The event selection of the data sample is described in the first part of the chapter. The fundamental requirements are a selection of multihadronic events, a modest cut on the b-tagging, which ensures a $b\bar{b}$ sample purity of approximately 90 % and a successful run of BSAURUS which ensures a successful fitted secondary vertex for every event hemisphere. Additional requirements are made for each hemisphere in order to improve the quality of the sample by cutting on the expected vertex fit error. The removal of hemispheres containing a lepton with high transverse momentum reduces the statistical overlap with other DELPHI analyses [KPP$^+$01, Lip, B$^+$]. In total 408000 hemispheres are selected in the data sample from 1992 to 2000, which is roughly 10 times more than the DELPHI analysis using semi-leptonic B decays [KPP$^+$01].

Optimal performance of the analysis is achieved if the events are classified with respect to their decay length and B meson momentum resolution. Events with a good resolution get a higher statistical weight during the fitting routine. The classification algorithm for the decay length reconstruction is based on the expected decay length error of the different vertex fitting algorithms. The charged multiplicity of the hemisphere is used for the classification for the case of the B meson momentum. In the second part of the chapter an exact description of the classification algorithms is provided. For the mathematical treatment of the resolution functions in the fitting routine it is necessary to transform the resolution distributions in analytically accessable functions, and this procedure is also described in this section.

Although the tagging algorithms for production and decay are already discussed in the previous BSAURUS chapter 4, a critical point of the analysis is the calibration of both algorithms for the data sample. The chosen fitting method requires that every input quantity can be interpreted as probability. These probabilities for the decay and the production tag are derived from the simulation and due to imperfect modelling, are usually found to be overly optimistic for the data. Therefore a dedicated calibration algorithm was developed using the fraction of events tagged as mixed from the data sample compared to the expectation from the simulation [Klu01]. This calibration is described starting at section 5.3.

All results are extracted in terms of a Maximum Likelihood Fit (see [Cow98] for an introduction). The detailed formulation of the fit is provided in the fourth part of the chapter. In the last part the result for the B_d^0 meson mixing frequency Δm_d is provided together with a discussion of the systematic errors. The much faster oscillation of the B_s meson requires another approach: The so-called 'amplitude method' [HGA97], which is basically a Fourier transformation of the data sample, is introduced and some tests on simulated events are described. Finally the result for the B_s case together with a discussion of the systematic errors is provided.

5.1 Data Sample

An inclusive oscillation analysis is only possible on a well selected, high quality multihadronic sample. In addition, the b quark purity should be better than 95 % suppressing c-quark and uds-quark background which is hard to parametrise in a likelihood function.

A description of the event and hemisphere selection is provided in the following two sections.

5.1.1 Event Selection

The first step of the event selection is the separation of multihadronic events ($e^+e^- \to Z \to q\bar{q}$) from various backgrounds. The background sources are widely spread, e.g. two photon and Bhabba events, leptonic events, beam-gas interaction events and cosmic rays. Multihadronic events are selected similar to the BSAURUS selection in section 4.2.1 by requiring at least 5 charged tracks and 12 % of the beam energy in the event. The selected number of

Selection	92/93	94/95	96/00	$q\bar{q}$ MC	bb MC
Multihadron	1381635	2022756	347039	1542366	1246369
N_{JET}	1381635	2022754	347038	1542365	1246369
$\cos_{THR}(\Theta)$	1251488	1821920	312410	1381482	1121791
$\cos_{JETS}(\alpha)$	1186471	1709362	288968	1302557	1053918
B tagging	115456	223625	37139	169233	590638

Table 5.1: The number of events for data and simulation. The rows show the number of events after the multihadron selection (first row), number of jets smaller than 5 (second row), cosine of the thrust axis smaller than 0.75 (third row), the cosine angle of the B candidate jets (fourth row) and the combined event b tagging cut (fifth row).

events for the individual years[1] in data and simulation[2] are listed in row one of table 5.1.

In the second step the number of jets is reconstructed via the LUCLUS [Sjo94] algorithm. The chosen transverse momentum cutoff value is $d_{join} = 5\,GeV$. All events with 5 or more reconstructed jets are rejected. This cut is chosen very loose as there is a more effective angular cut selecting 2 jet events described in the next section. The selection result of this N_{JET} cut is listed in row two of table 5.1.

As described in section 2.2.1 the vertex detector has a limited polar acceptance, making all events in very forward direction unusable because as a successful vertex reconstruction is essential. Events with a cosine of the thrust ($\cos_{THR}(\Theta)$) greater than 0.75 are rejected. Table 5.1 lists the remaining events of this cut in row 3.

As outlined in section 4.2.2 of chapter 4 BSAURUS selects in each candidate hemisphere the jet which is most likely from the b quark. Nice 2 jet events are selected if the cosine of the angle between the two jets ($\cos_{JETS}(\alpha)$) is smaller than -0.9. The result of this selection is illustrated in row 4 of table 5.1.

The final event selection cut is based on the combined event b tagging of section 4.1. The cut provides a sample b purity of 90 %. The last line of table 5.1 lists the finally selected events.

[1]The years 1996-2000 were no real Z^0 runs like the period from 1992-1995. But for calibration purposes a Z^0 phase was done before every high energy run of the year. During the years a reasonable amount of data was collected and is commonly used in various B analyses.

[2]The 1994 Monte Carlo sample was chosen for comparison.

Selection	92/93	94/95	96/00	$q\bar{q}$ MC	bb MC
BSAURUS	216191	424016	70864	319891	1121156
Vertex	114496	292899	51937	228873	811634
Lepton	107622	277907	49303	216132	765681
Prod. & Decay	100098	262329	46809	203070	722520

Table 5.2: The first row gives the number of hemispheres with successful BSAURUS vertex runs. In the second the Strip-Down or BuildUp vertex selection is made, in row three the high p_T lepton and in row four, the final selection step, the production and decay tag cut is made.

5.1.2 Hemisphere Selection

After the event selection the point of view switches to the hemisphere level. Each individual hemisphere is a possible source for a oscillating B meson. As a lot of the used analysis tools are based on the BSAURUS package the minimal requirement for each candidate hemisphere is a successful BSAURUS vertex reconstruction, as described in section 4.2.4. The number of selected hemispheres is listed in the first row of table 5.2, which is built similar to table 5.1.

Based on these candidate hemispheres a dedicated B oscillation selection algorithm was developed including the following steps:

- At least one of the two optimised vertex algorithms, either the Strip Down or the Build Up, have to be finished successfully. In addition the fit error of at least one successful vertex fit should not exceed $600\mu m$. Otherwise the hemisphere is rejected. As the decay length resolution is most critical to any oscillation analysis it makes no sense using events with bad decay length reconstruction.

- if a lepton candidate is identified in the hemisphere with a transverse momentum greater than 1.2 with respect to the jet axis the lepton is assigned to, the hemisphere is discarded. This cut removes events already used in a DELPHI B oscillation analysis [Lip, A+96a] optimised for high p_T leptons. This cut ensures a reduced statistical overlap of the analyses.

- The production flavour net (see 4.2.7) of the opposite hemisphere must have a reasonable result between zero and one. Otherwise the event is rejected.

- The B (either s or d) decay flavour likelihood (see 4.2.6) has a reasonable value not equal 0. In this case at least one track fulfils the necessary selection criteria.

One can see in table 5.2 the individual selection steps described in the previous section.

The hemisphere selection, especially the strong vertex selection criteria leads to an improvement of the b purity to 96.5%. The remaining background splits into 3% c quarks and 0.5% light quarks[3].

5.2 Proper Time Reconstruction

The proper time τ is calculated using:

$$\tau = \frac{ml}{cp} \tag{5.1}$$

were m is the mean B hadron mass, c the speed of light, p the estimated B hadron momentum and l the reconstructed decay length. The expected error on the proper time is estimated using:

$$\sigma_\tau = \sqrt{\left(\frac{m\delta l}{cp}\right)^2 + \left(\frac{ml\delta p}{cp^2}\right)^2} \tag{5.2}$$

It is clearly visible that the expected error splits into two components, which behave differently. The decay length error δl gives a constant contribution to σ_τ, while the contribution of the momentum error δp to σ_τ increases linearly with τ.

This error behaviour motivates the separated treatment of the decay length and momentum resolution contribution in the likelihood fit, including the convolution procedure.

In the next section the parameterisation of the decay length and its error is described. Afterwards the the classification algorithm for the momentum reconstruction is discussed.

5.2.1 Decay Length Resolution Classes

Five different decay length classes with increasing resolution were defined, based on simulation studies. Each event was classified using the algorithm described below. If an event failed to fit in a certain class it was tested for the next.

[3]The term "light quarks" combines u, d and s quarks

1. The Strip Down fit worked and the expected fit error is below 200 μm.
2. The Build Up fit worked and the expected fit error is below 230 μm.
3. The Strip Down fit worked and the expected fit error is below 300 μm.
4. The Build Up fit worked and the expected fit error is below 380 μm.
5. Either the Strip Down or the Build Up fit worked and the expected fit error is below 600 μm.

The decay length resolution of the individual classes is plotted in figure 5.1. A double Gaussian is fitted to every distribution. The inner and outer Gaussian σ_1 and σ_2 is shown in every plot. Also shown for every class is the class ratio, which is the fractional contribution of each class to the total sample.

The fitting routine for both vertex algorithms had the constraint of reconstructing positive decay length only. This complicates the parameterisation of the resolution functions for the likelihood fit. Based on simulation studies seven different decay length regions based on the truth decay length l_{true} were defined and the resolution function was plotted. The following borders were defined, starting at region 1: $l_{true} < 50\,\mu\text{m}$, $< 100\,\mu\text{m}$, $< 200\,\mu\text{m}$, $< 500\,\mu\text{m}$, $< 1000\,\mu\text{m}$. The seventh region contained all events with greater decay length. In figure 5.2 the resolutions for the first six regions are plotted starting at the top left with region 1 and ending on the bottom right with region 6. The plots are done for the best resolution class 1. The distribution gets more and more symmetric towards the higher region classes five and six. An asymmetric double Gaussian with six parameters $P_{2..7}$ is fitted to every class:

$$f(x,P_i) = P_1 \times \begin{cases} P_2 e^{-\frac{1}{2}(\frac{x}{P_3})^2} + (1-P_2)e^{-\frac{1}{2}(\frac{x}{P_4})^2} & : \quad x \geq 0 \\ P_5 e^{-\frac{1}{2}(\frac{x}{P_6})^2} + (1-P_5)e^{-\frac{1}{2}(\frac{x}{P_7})^2} & : \quad x < 0 \end{cases} \tag{5.3}$$

The parameter P_1 is used for normalisation. Starting from a set of parameters P_i^k for every region $k \in \{1..6\}$ a linear interpolation function was calculated:

$$P_i(l_{true}) = \begin{cases} \frac{P_i^1 - P_i^0}{l_1} \cdot l_{true} & : \quad l_{true} \leq l_1 \\ \frac{P_i^k - P_i^{k-1}}{l_k - l_{k-1}} \cdot l_{true} & : \quad l_{k-1} < l_{true} < l_k \\ P_i^6 & : \quad l_{true} \geq l_k \end{cases} \tag{5.4}$$

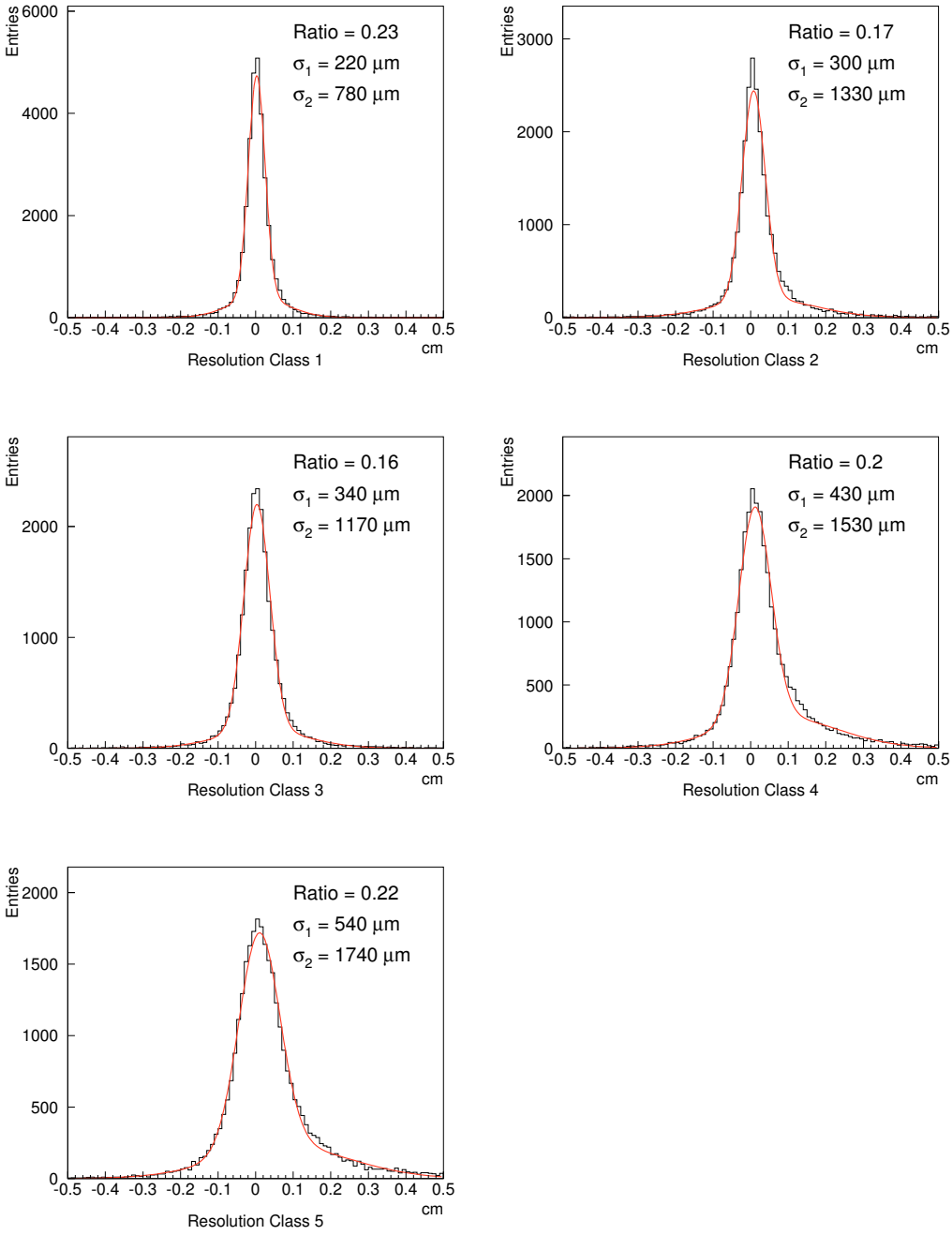

Figure 5.1: The decay length resolution of the five different classes are plotted. A double Gaussian is fitted to every distribution with inner Gaussian σ_1 and outer Gaussian σ_2. The ratio gives the contribution of every classes with respect to the total.

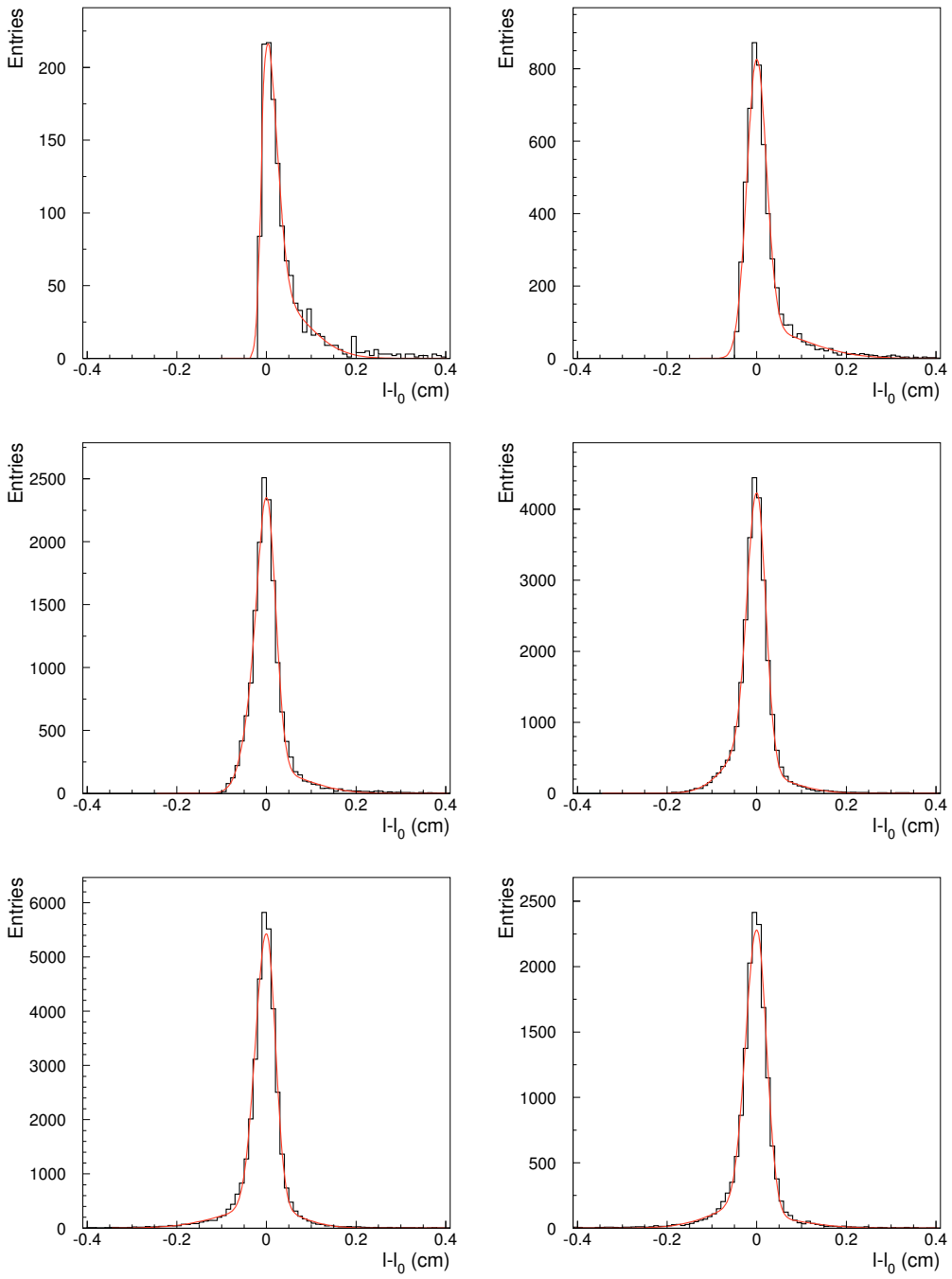

Figure 5.2: The six plots illustrate the different shapes of the decay length resolution for several regions of l_{true} for the resolution class 1. As both vertex algorithms reconstruct only positive decay length the resolution function for events with $l_{true} < 50\ \mu$m is largely forward biased (top left plot). For bigger true decay length the distribution gets more and more symmetric (bottom right plot).

The different region borders define the sampling points l_k. P_i^0 for $l_{true} = 0$ was defined by hand, as $P_3^0 = P_4^0 = 0$, $P_2^0 = 0.5$, $P_5^0 = P_5^1$, $P_6^0 = P_6^1$ and $P_7^0 = P_7^1$.

The final form of the resolution functions for the different resolution classes t used in the likelihood fit program is,

$$R^t(l_m - l_{tr}, l_{tr}) = \frac{1}{G} \begin{cases} P_2^t(l_{tr}) e^{-\frac{1}{2}(\frac{l_m - l_{tr}}{P_3^t(l_{tr})})^2} & \\ +(1 - P_2^t(l_{tr})) e^{-\frac{1}{2}(\frac{l_m - l_{tr}}{P_4^t(l_{tr})})^2} & : \quad l_m \geq l_{tr} \\ P_5^t(l_{tr}) e^{-\frac{1}{2}(\frac{l_m - l_{tr}}{P_6^t(l_{tr})})^2} & \\ +(1 - P_5^t(l_{tr})) e^{-\frac{1}{2}(\frac{l_m - l_{tr}}{P_7^t(l_{tr})})^2} & : \quad l_m < l_{tr} \end{cases} \tag{5.5}$$

were l_m denotes the measured and l_{tr} the truth decay length. The factor G is used for preservation the normalisation of the resolution function.

In figures 5.3, 5.4 and 5.5 a comparison of the resolution function based on the simulation and the parameterisation of equation 5.5 is shown in different bins of l_{true}.

5.2.2 B Hadron Momentum Classes

Independent from the decay length classification routine described in the previous section a classification of the reconstructed B hadron momentum is done. The B hadron momentum was reconstructed with the algorithm described in section 4.2.3. Studies of simulated events have shown a large correlation between the number of charged tracks and the resolution obtained in the energy reconstruction. For this reason, the number of charged tracks in the hemisphere was chosen to define different resolution classes. In total 16 different classes were defined, starting with 2 charged tracks per hemisphere in class one and ending with 17 and more tracks in class 16.

This behaviour can be understood from the fragmentation behaviour of B hadrons. In the case of a very hard fragmentation most of the energy goes into the B and almost no fragmentation tracks are produced. A weakly fragmentating event or an event with an accompanying gluon jet leads to many non B tracks and therefore to a bigger multiplicity.

In figure 5.6 the relative momentum resolution for four different samples is shown. For illustration purposes only collections of the 16 original classes are made with 4 classes in every sample. One can clearly see the decreasing resolution for increasing multiplicity (From top left to bottom right).

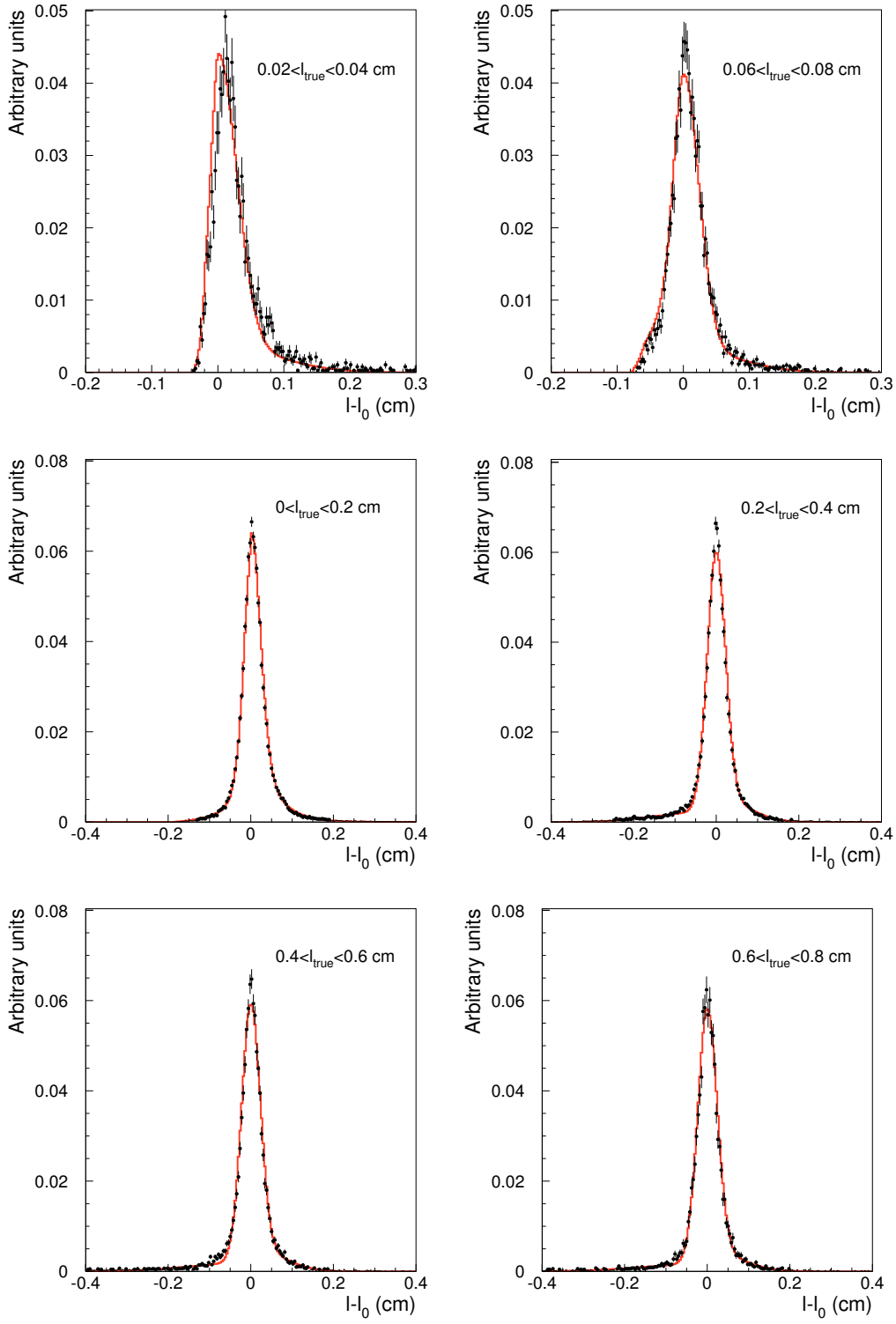

Figure 5.3: The plots illustrate the decay length resolution distribution for class 1 events, compared to the parameterisation function $R^1(l_m - l_{tr}, l_{tr})$. The two top plots are made for very small values of $l_{true} < 400\,\mu$m and $< 800\,\mu$m. The other four plots are made in steps of $0.2\,cm$ starting at $0\,cm$.

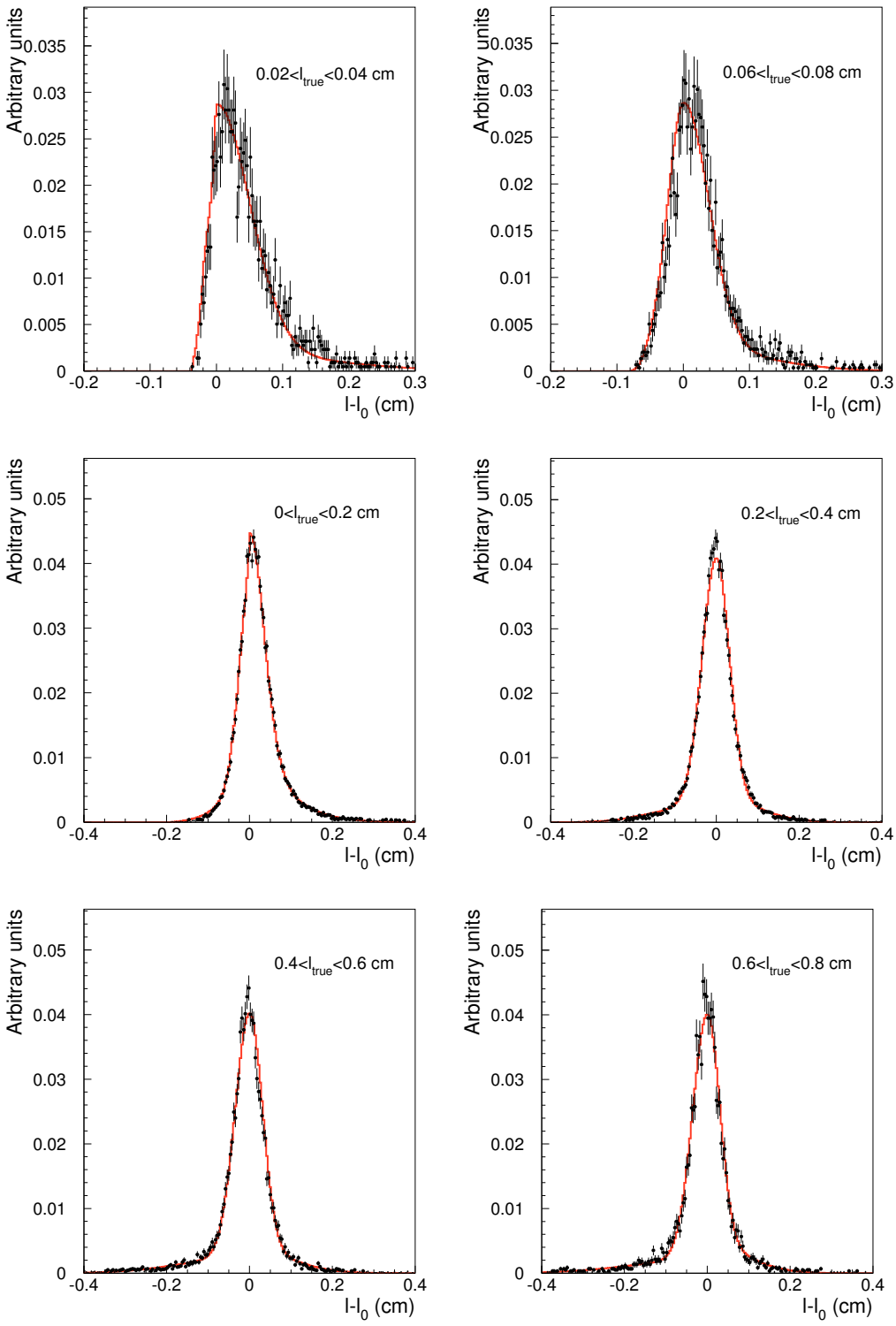

Figure 5.4: The plots illustrate the decay length resolution distribution for class 3 events, compared to the parameterisation function $R^3(l_m - l_{tr}, l_{tr})$. The two top plots are made for very small values of $l_{true} < 400\,\mu$m and $< 800\,\mu$m. The other four plots are made in steps of $0.2\,cm$ starting at $0\,cm$.

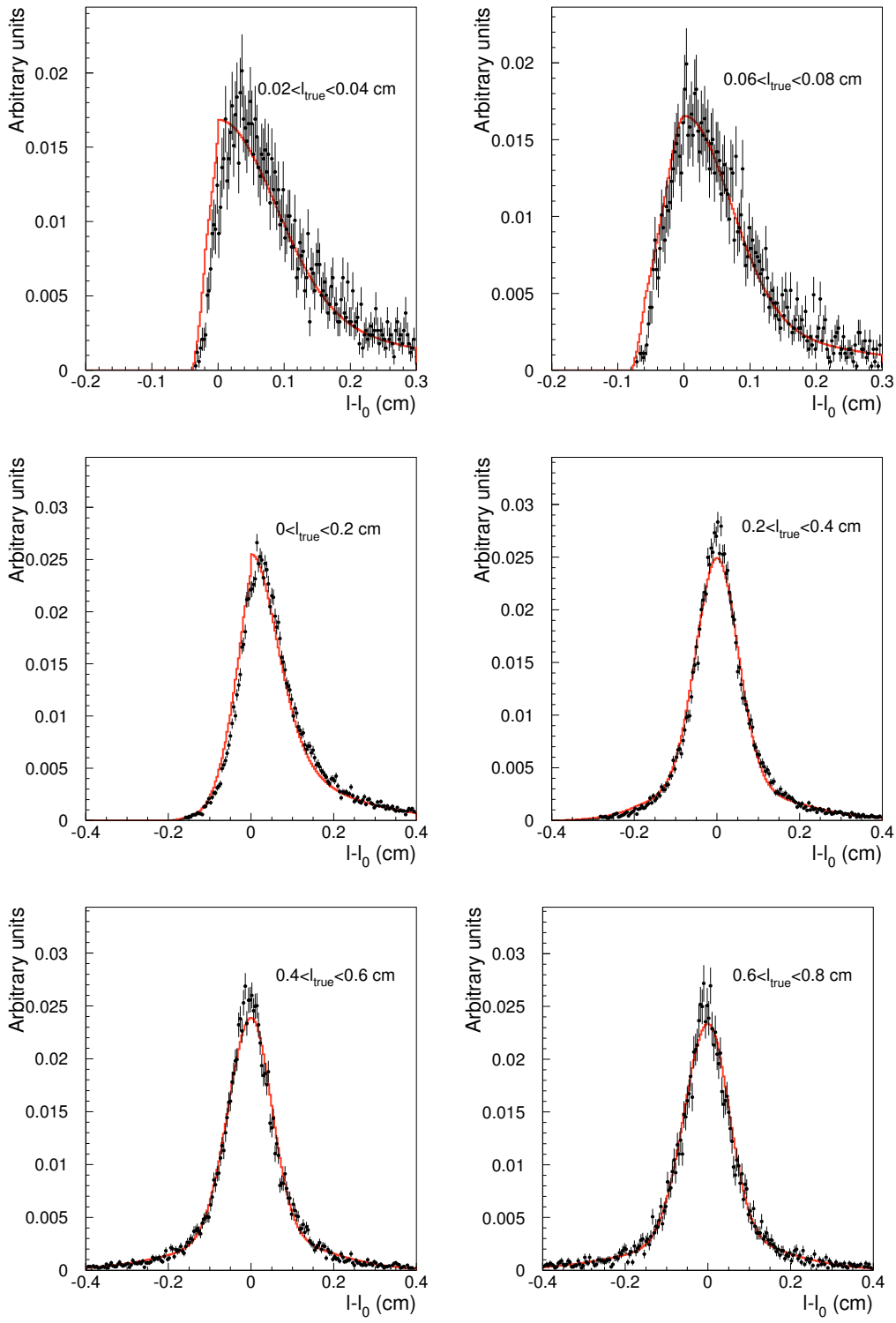

Figure 5.5: The plots illustrate the decay length resolution distribution for class 5 events, compared to the parameterisation function $R^5(l_m - l_{tr}, l_{tr})$. The two top plots are made for very small values of $l_{true} < 400\,\mu$m and $< 800\,\mu$m. The other four plots are made in steps of $0.2\,cm$ starting at $0\,cm$.

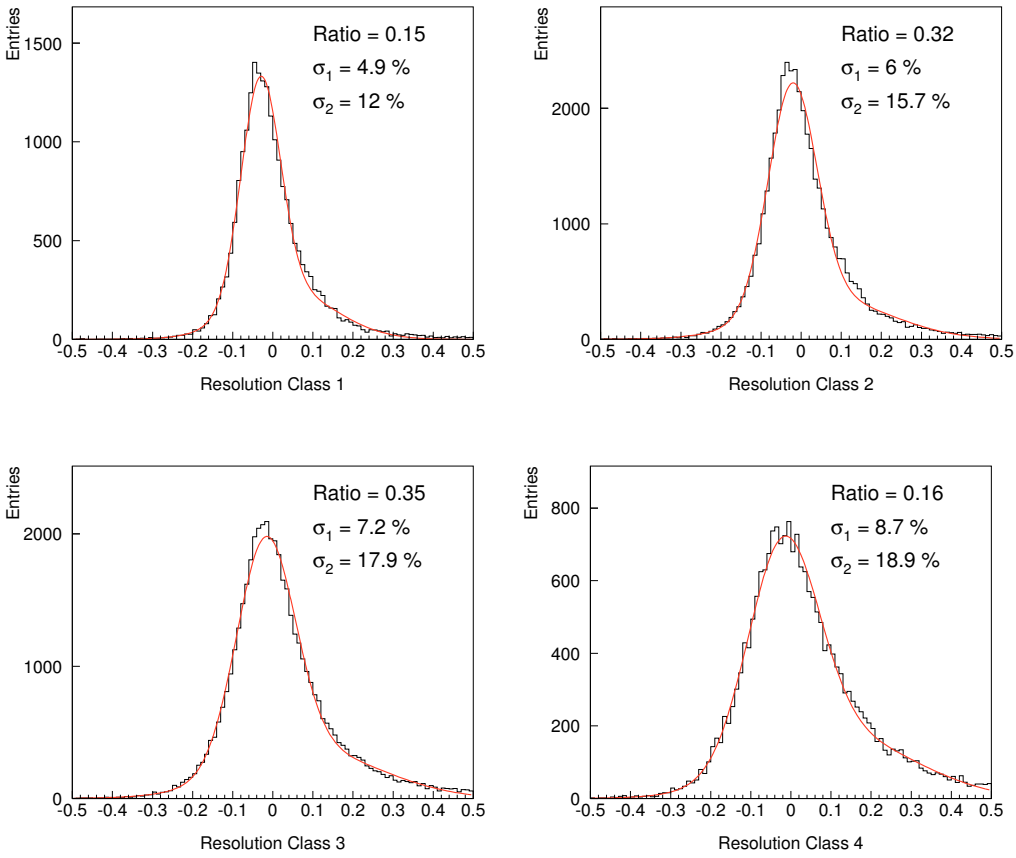

Figure 5.6: The 16 momentum classes are combined into 4 samples with 4 classes in every bunch. This is done for illustration purposes only. In the final Likelihood Fit all 16 classes are used. The relative B Hadron momentum resolution for the four different samples from the top left to the bottom right ($N_{charged} \leq 5$, $5 < N_{charged} \leq 9$,$9 < N_{charged} \leq 13$ and $13 < N_{charged}$).

The energy resolution function was parametrised using a double Gaussian as a function of the true momentum p and the reconstructed momentum p_{rec} in the following way:

$$\begin{aligned}\mathcal{R}^s((p_{rec}-p)/p) = \quad & T_1 \quad \frac{1}{\sqrt{2\pi T_3}} e^{-\frac{1}{2}(\frac{T_2-(p_{rec}-p)/p}{T_3})^2} \\ & +(1-T_1) \quad \frac{1}{\sqrt{2\pi T_5}} e^{-\frac{1}{2}(\frac{T_4-(p_{rec}-p)/p}{T_5})^2}\end{aligned} \tag{5.6}$$

The index s denotes the momentum class of the event. The different parameters T_i are derived from the simulation for every class s separately.

5.3 Tagging Procedure

The production and decay flavour tagging was done by the neural networks described in section 4.2.6. For the production tag only the opposite side flavour information was used. Several attempts were made including the fragmentation flavour information of the candidate side into the production tag, but the correlation to the decay tag were too large and uncontrollable on the data sample.

In an event by event B oscillation analysis the tagging purities ϵ_b for different B hadron species, c quarks or light quarks are incorporated. In the theoretical case of a perfectly trained neural network, the net output should be identical to its purity[4]. This statement is only valid for a perfectly trained network and for the hypothesis the network is trained on, i.e. the production flavour network is trained on $b\bar{b}$ events and thus larger deviations are possible for $c\bar{c}$- and light-quark events. Compensating deviations from the optimal behaviour is most easily done by modification of the tagging probability P. For this purpose a slope α is introduced and the new probability P_{new} is defined as:

$$\begin{aligned} R &= \frac{P_{old}}{1-P_{old}} \\ R' &= R \cdot \alpha \\ P_{new} &= \frac{R'}{1+R'} \end{aligned} \tag{5.7}$$

A slope α of 1 means that the probability remains unchanged, smaller values of α lead the probability closer to 0.5 which is equivalent to a random tag.

[4]In the case of a decision between two possibilities only "purity" can be replaced by "tagging probability"

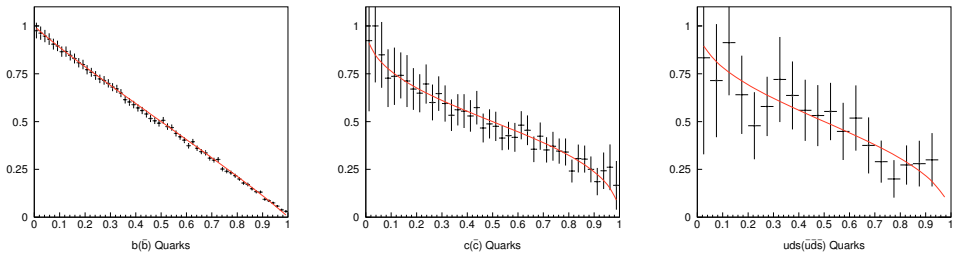

Figure 5.7: The production tag slopes for b, c and light quarks. The fitted slope values α are are 0.941 for b, 0.564 for c and 0.586 for light quarks.

The parameter α is obtained for the production tag by fitting equation 5.7 with free α to the purity distributions of b,c and light quarks. The result of the fit is plotted in figure 5.7. (The fitted slopes α are 0.941 for b, 0.564 for c and 0.586 for light quarks.)

A similar correction procedure with more correction steps is necessary for the decay tag, because of the the B species and decay length dependence of the tag, which has to be modelled correctly for the later likelihood fit. The first correction step is done for the individual B species and the c and light quarks. All decay tag slope values are listed in table 5.3. The individual fits of the B_d^0 decay tag slopes for the different B hadron species are visible in figure 5.8. The weakest deviation from slope 1 of the decay tag can be found for the hypotheses the decay tag is trained on, as expected for a well trained network.

Detailed studies of the tagging behaviour indicated another correction which is necessary for the decay tag. As expected the tagging power of a sample of hemispheres with very short decay length is weaker than another

Particle	B_d^0 decay tag slope	B_s decay tag slope
B_s	0.415	0.941
B_d^0	1.149	1.012
B^+	1.149	1.076
Λ_b	0.636	0.870
c quarks	0.970	0.809
light quarks	0.017	0.127

Table 5.3: The decay tag slopes for the B_d^0 and B_s tag. On can see that the decay tags nearly have no power for light quarks.

with bigger decay length. This can be easily understood by the pollution of the decay tag with tracks from the fragmentation, where the separation between the primary and secondary vertices is not as good as in events with larger decay length. This decay length dependence of the tagging is already modelled in the networks (see figure 5.9 left plot), but the real behaviour obtained from a 'like-sign ratio'[5] calculation indicates a weaker drop (see figure 5.9 right plot). This difference was compensated by a decay length dependent term:

$$\alpha' = \alpha(1 + \beta e^{-6 \cdot l}) \tag{5.8}$$

with a small correction factor β for every B hadron type, dependent on the decay length l. The power factor -6 was obtained from simulation studies.

Finally a very small correction of about $\pm 5\%$ to the slopes α, dependent on the 5 different decay length resolution classes was made.

The overall decay flavour distributions for the different B hadrons look pretty much the same, but they are weaker for charm and light quarks (e.g. there are no light quarks with decay probability > 0.8 or < 0.2). This flavour dependent tagging acceptance has been taken into account and parametrised using results from the Monte Carlo simulation.

5.4 Acceptance and Background Functions

Theoretically expected for all B decays is an exponential function with lifetime τ_B. As it is more complicated separating primary from secondary vertices in b events with very short lifetime, or even recognising events with very short decay length as B events, the b tagging cut (see section 4.1) and the first level vertex reconstruction reduces the efficiency for those events. The deviation from the exponential decay is taken from simulation and an acceptance is parametrised accordingly. Note that for B_s and B_d^0 oscillations the fraction of like sign events is relevant and the acceptance drops out to first order.

The proper time distribution for background events from charm and light quarks was taken from the simulation as there is no nice analytic description of it and the low statistic make them quite difficult to deal with. The two distributions are plotted in figure 5.10.

[5] The fraction of events with a positive correlation between the decay tag ($> 0.5 \equiv b$ or $< 0.5 \equiv \bar{b}$) and the simulated truth flavour divided by all events defines a so-called 'like-sign ratio'

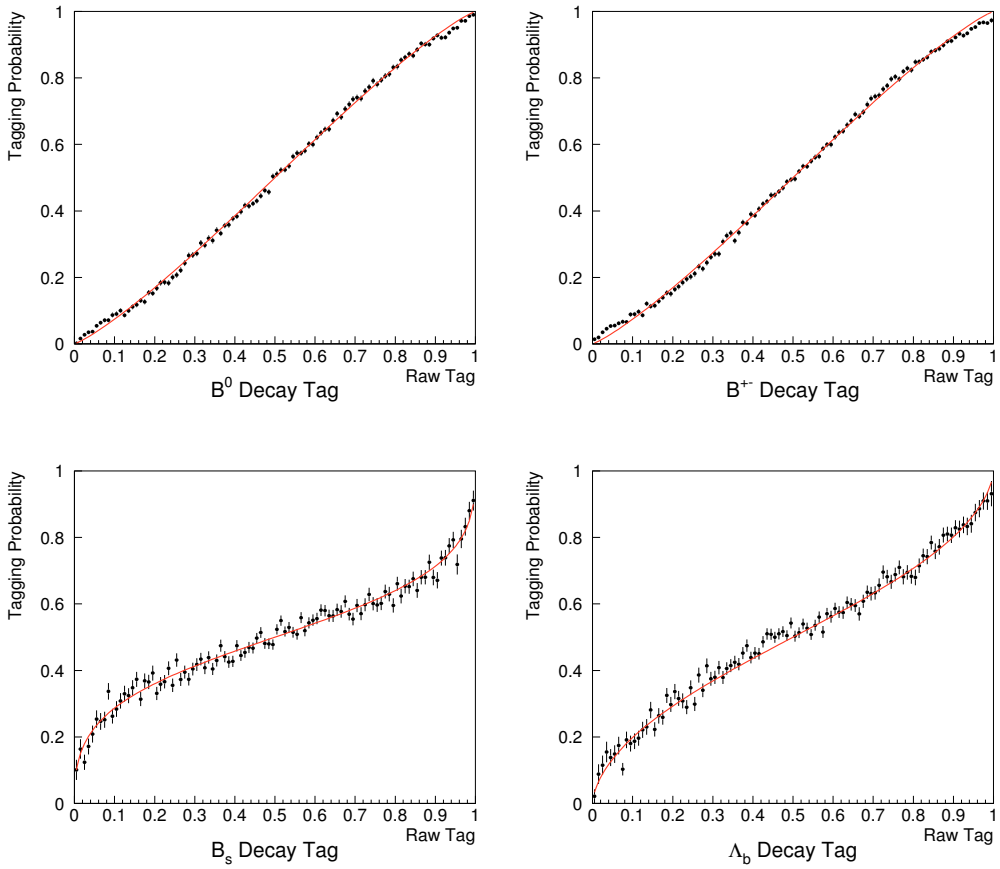

Figure 5.8: The tagging purity curves for the B_d^0 decay tag of the different b hadron species. For the B_d^0 mesons the curve is nearly linear, while there are bigger deviations for the other types.

5.5 The Likelihood Function

As described in section 5.2 the measurement of the decay length l and momentum p is only possible with finite resolution, parametrised by their resolution functions. The reconstructed proper time τ_{rec} is calculated using equation 5.1. The probability function $\mathcal{P}(t_{true})$ for a B hadron can be written as function of true l_{tr} and p_{tr}[6],

$$\mathcal{P}(t_{true}) = \mathcal{P}(l_{tr}, p_{tr}) = \frac{e^{-l_{tr}m/(\tau_b c p_{tr})}}{\tau_b} \tag{5.9}$$

[6] $tr \equiv true$ and $rec \equiv reconstructed$

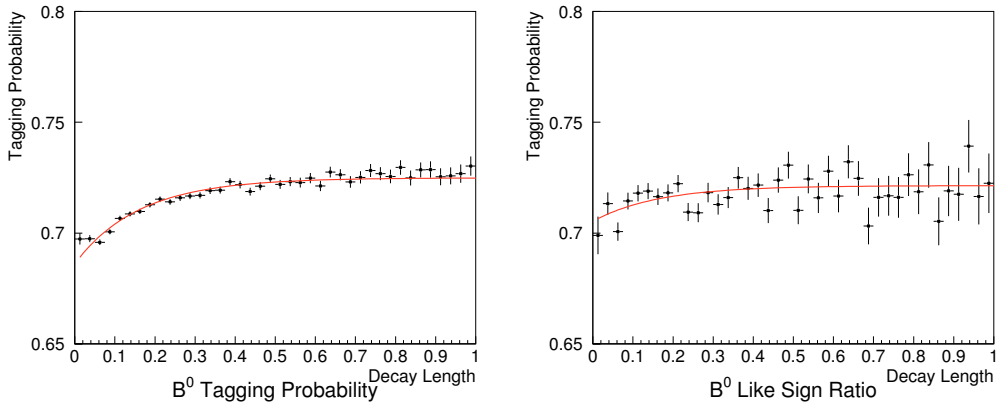

Figure 5.9: On can clearly see the stronger drop of the B_d^0 decay tagging purity for simulated B^+ mesons in the left plot for small decay length compared to the right one. This difference indicates that the forecast of the decay tag is slightly too pessimistic. The difference between the distributions has been taken into account and parametrised.

were τ_b denotes the B hadron lifetime. The probability $\mathcal{P}_b^{t,s}(t_{rec})$ for a B hadron to be observed at reconstructed proper time t_{rec} is an convolution over an exponential B decay probability $\mathcal{P}(t_{true})$, an acceptance function $A(l,p)$, the true B hadron momentum distribution $\mathcal{F}(p)$ taken from simulated events and the resolution functions:

$$\begin{aligned}\mathcal{P}_b^{t,s}(t_{rec}) &= \int_{l=0}^{\infty}\int_{p=0}^{\infty} A(l_{tr},p_{tr})\mathcal{F}(p_{tr})\mathcal{R}^t(l_{rec}-l_{tr},l_{tr}) \\ &\cdot\ \mathcal{R}^s((p_{rec}-p_{tr})/p_{tr})\mathcal{P}(l_{tr},p_{tr})dl_{tr}dp_{tr} \qquad (5.10)\end{aligned}$$

The indices t and s denote the decay length and momentum class the event is assigned to. The probabilities for light and charm quark events, $\mathcal{P}_l(t_{rec})$ and $\mathcal{P}_c(t_{rec})$ are taken directly from the simulation as described in the previous section.

Based on the production and decay tagging probabilities from section 5.3, the combined mixing probability for a given event is defined as

$$P_{comb} = P_{prod}P_{decay} + (1-P_{prod})(1-P_{decay}) \qquad (5.11)$$

were P_{prod} and P_{decay} are the slope corrected probabilities, dependent on the

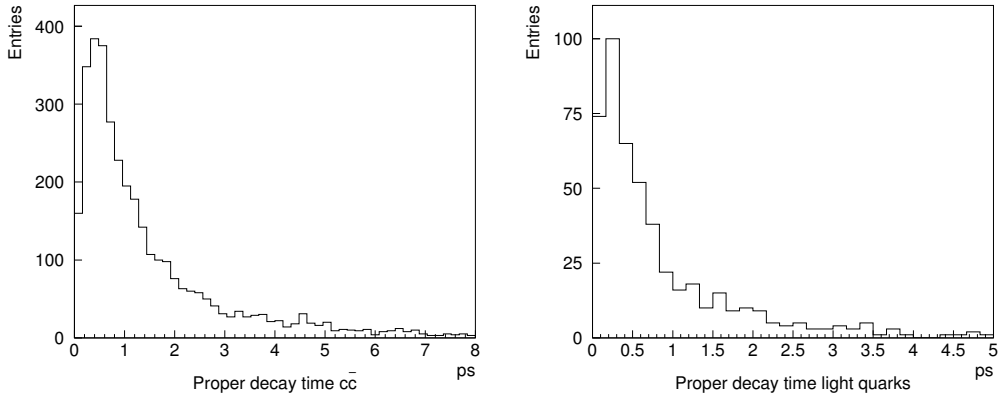

Figure 5.10: The proper time distribution for charm and (left plot) and light quarks (right plot). Both distributions have a long tail to bigger decay times. A separation within the individual decay length resolution classes was not possible due to the reduced statistic.

quark type, the B hadron type, the decay length and the decay length class[7]. If P_{comb} is larger than 50% the events are called 'like-sign'. 'Unlike-sign' events are those for which P_{comb} is less than 50%.

The total probability for a like-sign event is:

$$\begin{aligned}\mathcal{P}^{like}(t_{rec}) &= h_b \sum_q f_q\, \epsilon_{B_q}\, \mathcal{P}^{mix}_{B_q}(t_{rec}) + \sum_q f_q\, (1-\epsilon_{B_q})\, \mathcal{P}^{unmix}_{B_q}(t_{rec}) \\ &+ h_c\, \epsilon_c\, \mathcal{P}_c(t_{rec}) + h_l\, (1-\epsilon_l)\, \mathcal{P}_l(t_{rec}) \end{aligned} \tag{5.12}$$

and correspondingly for the unlike-sign events:

$$\begin{aligned}\mathcal{P}^{unlike}(t_{rec}) &= h_b \sum_q f_q\, (1-\epsilon_{B_q})\, \mathcal{P}^{mix}_{B_q}(t_{rec}) + \sum_q f_q\, \epsilon_{B_q}\, \mathcal{P}^{unmix}_{B_q}(t_{rec}) \\ &+ h_c\, (1-\epsilon_c)\, \mathcal{P}_c(t_{rec}) + h_l\, \epsilon_l\, \mathcal{P}_l(t_{rec}) \end{aligned} \tag{5.13}$$

The tagging purities ϵ_{B_q},ϵ_c and ϵ_l[8] were calculated from the combined tagging probability P_{comb} using,

$$\epsilon_{B_q,c,l} = 0.5 + |P_{comb} - 0.5| \tag{5.14}$$

[7] $P_{comb}(B_q, q, l_{rec}) \equiv P_{comb}$; B_q denotes the B hadron type, q the quark type (b, c, light) and l_{rec} the reconstructed decay length.

[8] A tagging probability of $P_{comb} = 0.1$ has the same statistical power as $P_{comb} = 0.9$. Therefore only the distance ('tagging purity') from $P_{comb} = 0.5$ is necessary in the fit.

The fraction of b, c and light quark events is represented by $h_{b,c,l}$, the B hadron fractions by f_q. Both fraction types are a function of the decay length resolution class as it is e.g. less likely to find a light quark in resolution class 1 than in class 5. The numbers were taken from the simulation, which was corrected to match the latest PDG values [GAA+00]. The c quark contribution is ϵ_c in the like-sign probability and $(1-\epsilon_c)$ in the unlike-sign term, which is opposite to all other contributions. The reason for this can be found in the production tag, which has opposite sign for c quarks compared to b quarks.

For the mixed(unmixed) B_q mesons the following expression is used:

$$\begin{aligned}\mathcal{P}^{mix(unmix)}(l_{tr}, p_{tr}) &= \mathcal{P}(B_q \to \bar{B}_q(B_q)) = \\ &\frac{1}{2\tau_q}\; e^{-l_{tr} m/(\tau_b c p_{tr})}[1 \mp \cos(\Delta m_q\, l_{tr} m/(c\, p_{tr}))] \quad (5.15)\end{aligned}$$

The probabilities $\mathcal{P}_{B_q}^{like(unlike)}(t_{rec})$ for the reconstructed proper time is calculated using the convolution formula 5.10. The unmixed probability $\mathcal{P}_{B_q}^{unlike}(t_{rec})$ for B^+ mesons and B baryons is calculated using equation 5.10 directly.

In the search for B_d^0 and B_s oscillations a likelihood fit was performed with the likelihood constructed from the elements described above:

$$\begin{aligned}\mathcal{L} = &- \sum_{like-sign} \ln(\mathcal{P}^{like}(t_{rec}, P_{prod}, P_{decay})) \\ &- \sum_{unlike-sign} \ln(\mathcal{P}^{unlike}(t_{rec}, P_{prod}, P_{decay})) \quad (5.16)\end{aligned}$$

Based on this likelihood sum two different approaches are made for the extraction of B_d^0 and B_s oscillations results. The B_d^0 oscillation measurement is described in the next section, followed by the B_s part which is based on the amplitude method [HGA97, BA99].

5.6 B_d^0 Oscillation Measurement

The mass difference between the two B_d^0 states is determined by fitting the fraction of like-sign events as a function of proper time, using the likelihood function described in the previous section.

The values for the fractions f_q are taken from the LEP Heavy Flavour working group [OG01], background fractions h_q are taken from the simulation and the B lifetimes τ_q are fixed to the values of the PDG [GAA+00]. The c and light quark background distributions and the acceptance function

Typ	Class 1	Class 2	Class 3	Class 4	Class 5
B_s fraction	0.076	0.104	0.084	0.104	0.100
B_d^0 fraction	0.407	0.370	0.419	0.396	0.412
B^+ fraction	0.456	0.443	0.429	0.405	0.396
B hadron fraction	0.061	0.083	0.068	0.095	0.092
Charm background	0.014	0.027	0.020	0.045	0.044
Light quark background	0.005	0.005	0.004	0.005	0.005

Table 5.4: The different B hadron and background fractions for the individual decay length classes. The B hadron fractions for the total sample without any selection cuts was $B_s = 0.097$, $B_d^0 = B^+ = 0.4$ and B baryon $= 0.103$.

$A(l,p)$ are parametrised using the simulation. The tagging purities ϵ_c and ϵ_l are taken from the simulation. The unknown B_s mixing frequency Δm_s is set to a very high value of $\Delta m_s = 18\,\mathrm{ps}^{-1}$, which is the most likely value within the Standard Model derived from other CKM matrix element measurements [C$^+$01]. The selection of the decay length classes 1 to 5 shifts the B hadron fractions slightly. The amount of such shifts is determined from the simulation. In Table 5.4 the B hadron fractions and, the charm and light quark background is listed for the 5 decay length classes.

Studies of data events have shown that the tagging power of the production tag and the decay tag is too optimistic (e.g. a expected B decay tag purity of 0.7 in the simulation is only 0.65 in the data). This effect can be adjusted by introducing slope corrections γ for the combined mixing probability P_{comb}. Equation 5.7 was used to recalculate the probability P_{comb} with a certain value γ. Equation 5.14 is used for recalculation of the mixing purities ϵ_b

The fit has three free parameters, Δm_d, the B_d^0 mass difference and two correction slopes γ for variation of the tagging purities ϵ_b of the B hadrons.

Parameter	92/93	94/95	96/00
γ_1	0.960 ± 0.09	0.703 ± 0.028	0.913 ± 0.10
γ_2	0.750 ± 0.08	0.789 ± 0.039	0.534 ± 0.08
Δm_d	0.497 ± 0.028	0.506 ± 0.016	0.522 ± 0.048

Table 5.5: The individual fit results for the different years. The different correction factors of the years can be explained by the different Kaon identification properties, which vary within the years and are important for the B_d^0 decay tag.

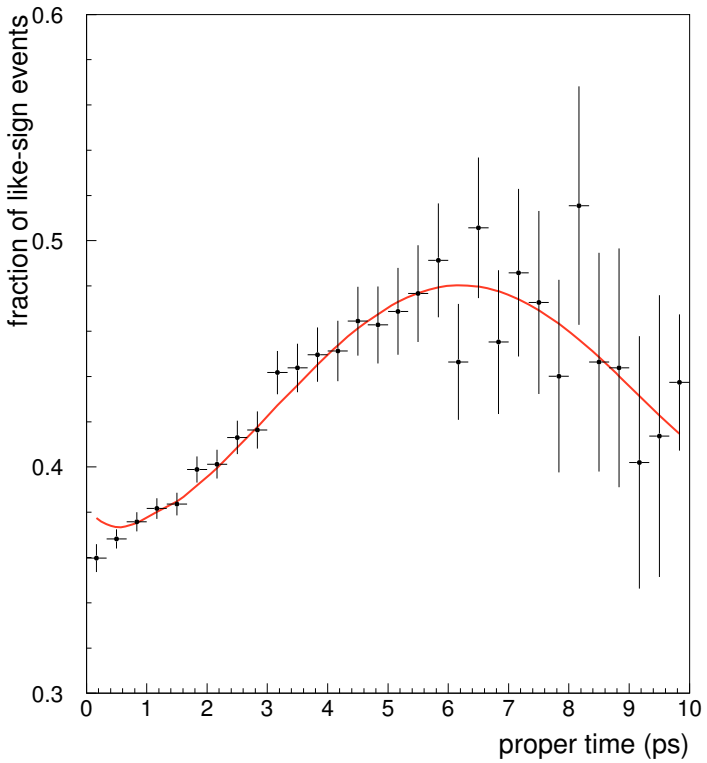

Figure 5.11: The fraction of like-sign events as a function of proper time using 1992-2000 data. The data is shown with error bars, the solid lines correspond to the fit result of $\Delta m_d = 0.501 \pm 0.013\,\mathrm{ps}^{-1}$.

Data studies indicated that charged B^+ mesons behave slightly differently for the B_d^0 decay tag, as expected from the simulation. Therefore the tagging purity for charged B^+ mesons ϵ_{B^+} was modelled as a function of the correction slope γ_1. The other B hadrons were varied as a function of γ_2. Putting the tagging purities in the fit made the fit very stable and independent of the knowledge of the exact tagging behaviour.

A check of the fit is done for 500k simulated events resulting in $\Delta m_d = 0.467 \pm 0.012$, $\gamma_1 = 1.002 \pm 0.02$ and $\gamma_2 = 0.997 \pm 0.02$ as fit result. This is to be compared with the true value in the Delphi simulation of $\Delta m_d = 0.475\,\mathrm{ps}^{-1}$

The data fit was done separately for the different years, as the tagging properties vary year-to-year. The individual results including the correction factors γ are listed in table 5.5 Using the different year-dependent correction factors the whole data set is re-fitted with a final result of $\Delta m_d = 0.501 \pm$

$0.013\,\mathrm{ps}^{-1}$.

The result of the fit is shown in figure 5.11. The plot shows the weighted like-sign fraction versus the reconstructed proper time for the 1992-2000 data compared to the fit result. There is a small deviation for small proper time which can be explained by different c and light background behaviour, or a contribution from the B_s oscillation, which is not included in the plot. A discussion of the systematic error is provided in the next section.

5.6.1 Discussion of Systematic Errors

The systematic errors of the B_d^0 oscillation measurement are mostly based on the uncertainties of the physical input quantities, i.e. B hadron lifetimes and B hadron production fractions. Effects of the tagging are compensated by the additional free fit parameters γ described in the previous section. As the likelihood fit measures the ratio of like-sign to unlike-sign events all acceptance function effects drop out to first order.

A breakdown of the systematic errors affecting the measurement is shown in table 5.6. The B_s, B baryon and the B_d^0, B^+ fractions are changed in the sample and correspondingly the other B fractions are recalculated. As a constrained it is assumed that B_d^0 mesons and B^+ mesons have exactly the same production fraction. For the production fraction errors the latest results of the LEP Heavy Flavour Steering Group are taken [OG01]. Finally the full correlation matrix of the fraction measurements is used and a total error on Δm_d extracted. The width of the Gaussian of the decay length and momentum resolution parameterisation is changed by a relative 10%. A scaling factor of 1 % is applied in order to describe the uncertainty for the absolute scale of the proper time reconstruction. Also changed are the B lifetimes, different Δm_s values of $15\mathrm{ps}^{-1}$ and $20\mathrm{ps}^{-1}$ for the B_s mixing and the charm and light quark background fraction $h_{c,l}$. A more detailed discussion about the range of the variation of the different quantities can be found in section 5.7.6, where the systematic effects of the B_s oscillation analysis are described. The total systematic error amounts to $0.0183\mathrm{ps}^{-1}$.

The final result thus is:

$$\Delta m_d = 0.501 \pm 0.013\ (\mathrm{stat}) \pm 0.018\ (\mathrm{sys})\ \mathrm{ps}^{-1} \tag{5.17}$$

The total error is therefore $0.022\,\mathrm{ps}^{-1}$.

Error Source	Values	Systematic Error on $\Delta m_d\,(\mathrm{ps}^{-1})$
B_s lifetime	$\pm 0.05\,\mathrm{ps}$	0.0001
B_d^0 lifetime	$\pm 0.015\,\mathrm{ps}$	0.0008
B^+ lifetime	$\pm 0.01\,\mathrm{ps}$	0.0006
B baryon lifetime	$\pm 0.05\,\mathrm{ps}$	0.0007
B_s fraction	± 0.011	
B_d^0 and B^+ fraction	± 0.011	
B baryon fraction	± 0.018	
Total B fraction contribution		0.0145
Decay length resolution	$\pm 10\%$ width scale	0.0002
Momentum resolution	$\pm 10\%$ width scale	0.0001
Proper time scale factor	$\pm 1\%$	0.0103
Light quark and c background	$\pm 20\%$ $h_{c,l}$ fraction	0.0009
Δm_s dependence	$\Delta m_s = 15,\ 20\,\mathrm{ps}^{-1}$	0.0012
Total systematic error		0.0181

Table 5.6: The systematic errors affecting the Δm_d measurement. Only a total systematic error for the B fractions is given as there are combined using the full correlation matrix information [OG01]. Tagging effects do not contribute as there are left free in the fit.

5.7 B_s Oscillation Analysis

The oscillation frequency Δm_s for B_s mesons is still not measured as the oscillations are still not resolved in contrast to Δm_d of the B_d^0 mesons. Two main reasons can be found for this condition. The B_s production fraction of $\approx 10\%$ leads to a signal to background ratio of 1/9 which is 6 times lower than in the B_d^0 case with a signal to background ratio of 2/3. In addition as measurements during the years have shown the B_s oscillation frequency Δm_s is much higher compared to the B_d^0 case, making the analysis very sensitive to the achieved proper time resolution.

In the following sections the B_s amplitude fit method and the effects of decay length and momentum resolution are described. Various fit results for different Δm_s frequencies are shown, to provide an easier understanding of the final result. Finally a breakdown of the systematic effects is given.

5.7.1 The Amplitude Method

In order to combine the information provided by the different analyses in the absence of a measurement of Δm_s a technique known as the *amplitude method* is used [HGA97, BA99]. The fit to the reconstructed proper time distribution of events tagged as like- and unlike-sign is performed with a fixed frequency ω and the following expressions for the mixed and unmixed B_s probability:

$$\mathcal{P}_{B_s}^{unmix}(t_{tr}) = \frac{1}{2\tau_{B_s}} e^{-\frac{t_{tr}}{\tau_{B_s}}} [1 + A\cos(\omega t_{tr})] \tag{5.18}$$

and similarly:

$$\mathcal{P}_{B_s}^{mix}(t_{tr}) = \frac{1}{2\tau_{B_s}} e^{-\frac{t_{tr}}{\tau_{B_s}}} [1 - A\cos(\omega t_{tr})] \tag{5.19}$$

The parameter A is left as the free fit parameter. A scan over the frequencies ω is performed and at each value the amplitude $A(\omega)$ and its error $\sigma_A(\omega)$ is measured. Averaging values from different analysis is straightforward, as the error $\sigma_A(\omega)$ provides the relative weight of the individual analyses. The expected value of the amplitude is unity if $\omega = \Delta m_s$ and compatible with zero in all other cases.

5.7.2 Resolution Dependence of the B_s Signal

As mentioned in the introduction of this section, the B_s mixing frequency is quite high and the sensitivity of an analysis is mostly provided by its proper time resolution. The proper time is calculated using equation 5.1 with expected error as given in equation 5.2.

The influence on the likelihood for different decay length resolution is quite different compared to the momentum resolution. A simulation sample of approximately 30k B_s events was chosen with $\Delta m_s = 6\,\mathrm{ps}^{-1}$. Using the convolution formula 5.10 the expected signal for an amplitude $A = 1$ for different resolutions is calculated. The calculation was performed for Gaussian decay length resolutions of 200 μm, 350 μm and 500 μm, either with relative momentum resolution of 5 % or 10 %, also parametrised by a Gaussian. The result of the calculation is illustrated in the like-sign ratio plots of figure 5.12. Weakening the decay length resolution reduces the absolute size of the oscillations, while the momentum resolution is represented in the drop of the oscillations with proper time.

The same calculations with the same parameters were done for the individual decay length classes 1-5 of the real resolution parameterisation. The results are illustrated in figure 5.13. The bottom right plot of figure 5.13

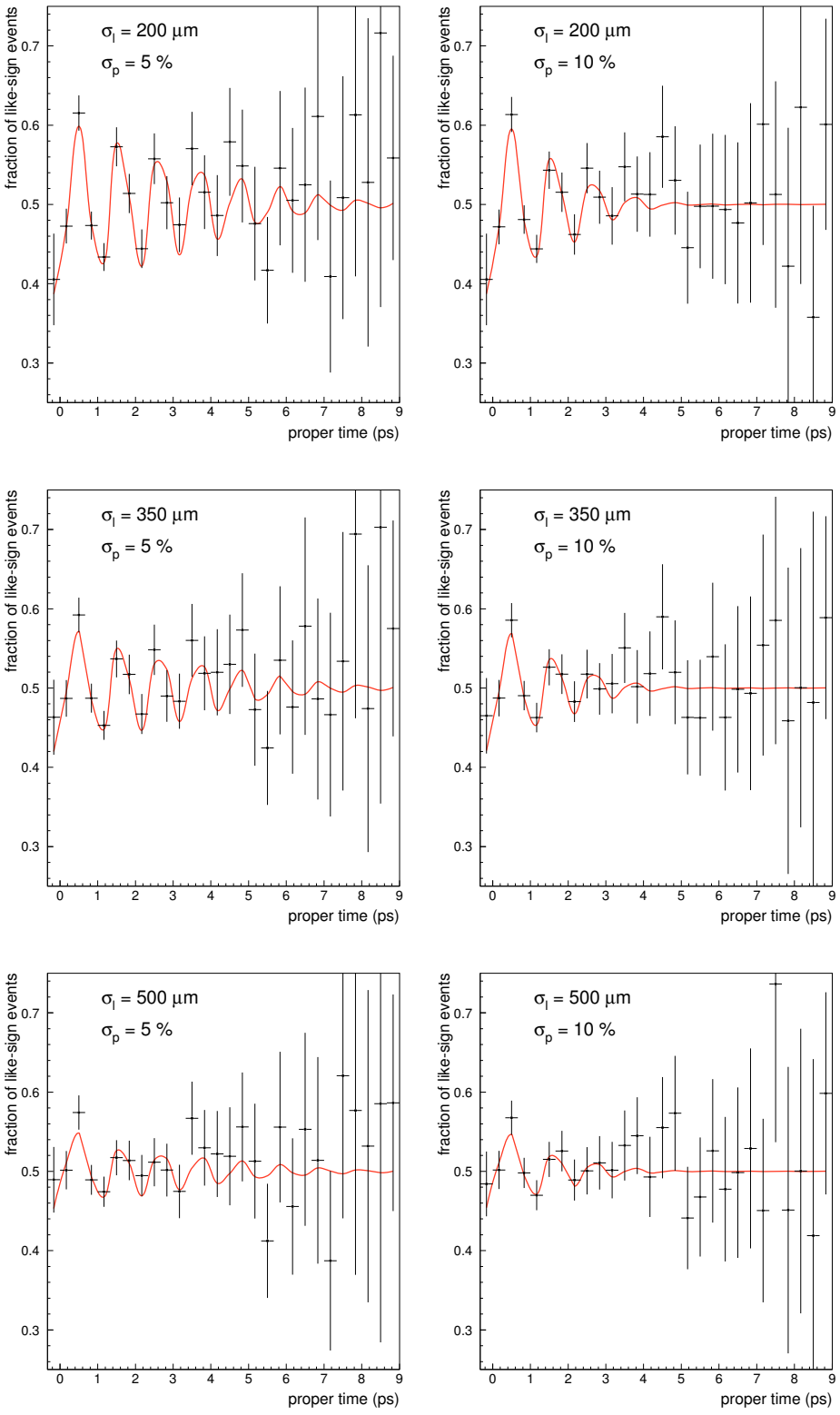

Figure 5.12: Expected like-sign ratio plots with amplitude $A = 1$ and $\Delta m_s = 6\,\mathrm{ps}^{-1}$ for 30k B_s events. The effect of different decay length resolution σ_l and momentum resolution σ_p is shown.

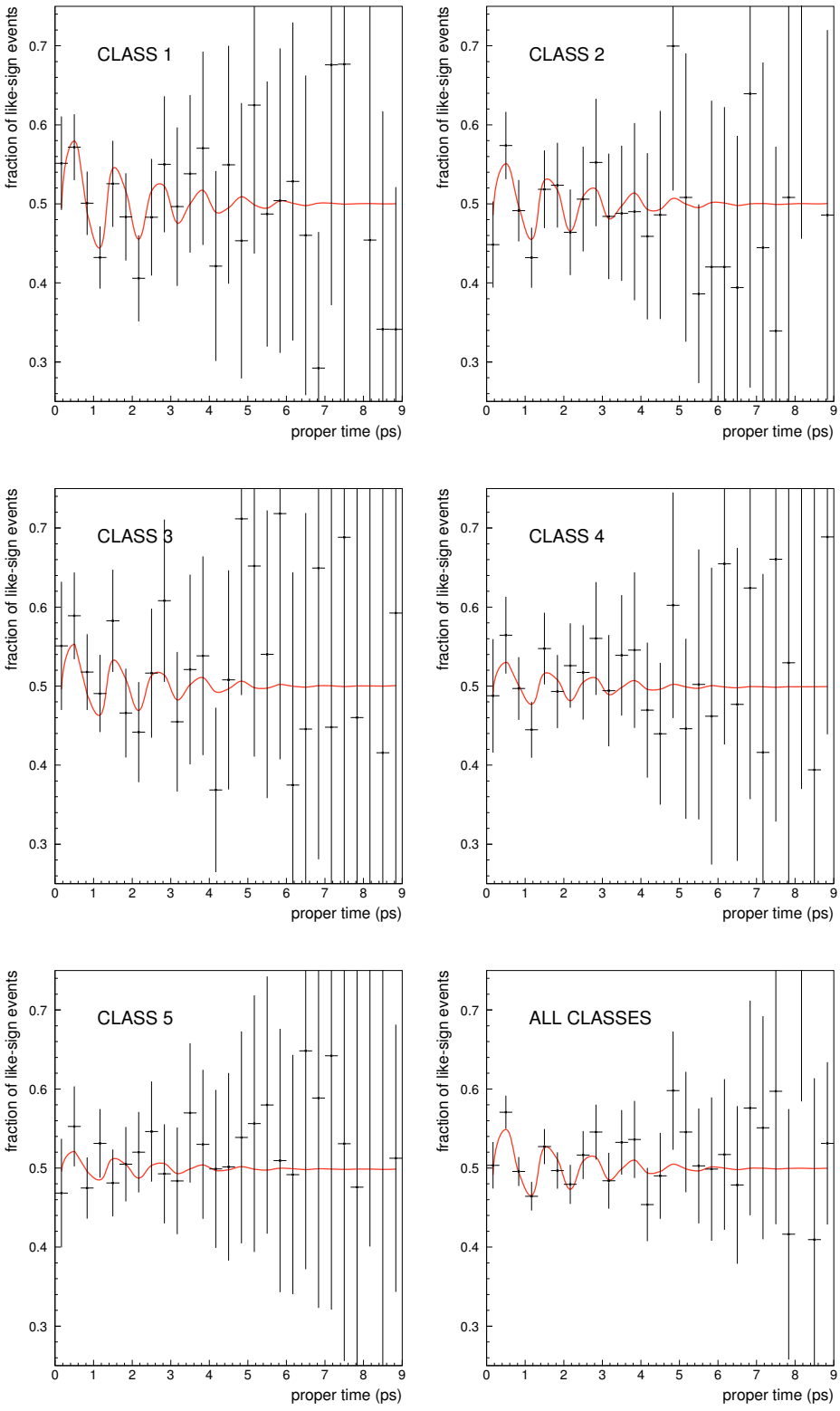

Figure 5.13: Expected like-sign ratio plots with amplitude $A = 1$ and $\Delta m_s = 6\,\mathrm{ps}^{-1}$ for 30k B_s events. The effect of the different decay length resolution classes one to five is shown (see section 5.2.1). In the bottom right the classes are mixed adjusted to the expected mixture.

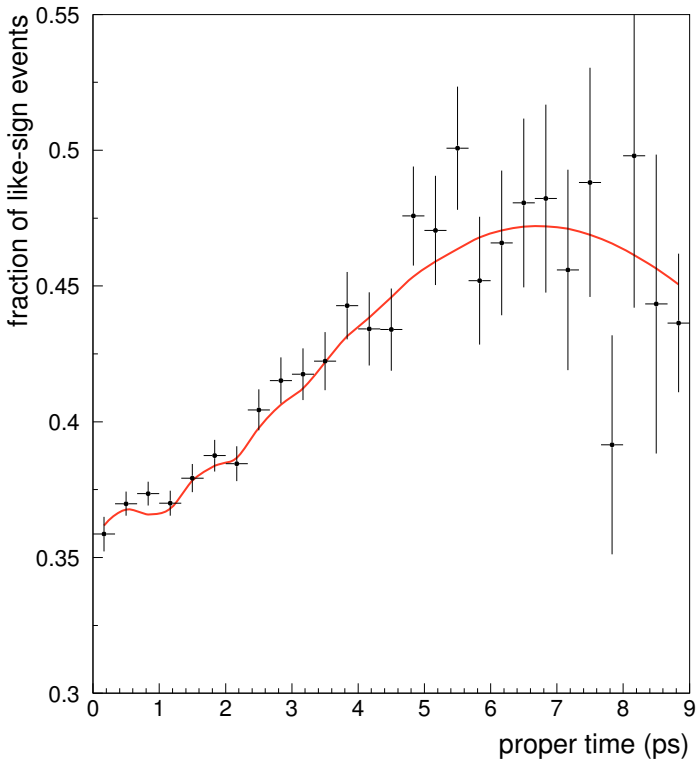

Figure 5.14: Expected like-sign ratio plot with amplitude $A = 1$ and $\Delta m_s = 6\,\mathrm{ps}^{-1}$ using 300k simulated $b\bar{b}$ events.

shows the expected signal for a natural mixture of the classes one to five as provided by the simulation.

Although the bottom right plot in figure 5.13 looks quite promising for a pure B_s sample, the significance drops drastically if one uses a naturally mixed $b\bar{b}$ sample with a B_s fraction of approximately 10 %. The expected like-sign ratio using 300k $b\bar{b}$ events with $\Delta m_s = 6\,\mathrm{ps}^{-1}$ and amplitude $A = 1$ is shown in figure 5.14.

5.7.3 Fitting different B_s Signals

A check of the likelihood fit and the amplitude method was performed by fitting 300k $b\bar{b}$ events with different true oscillation frequencies Δm_s. In the available simulated event sample from 1994/95 Δm_s is set to $11.5\,\mathrm{ps}^{-1}$. The simulation of other oscillation frequencies was provided by a routine which

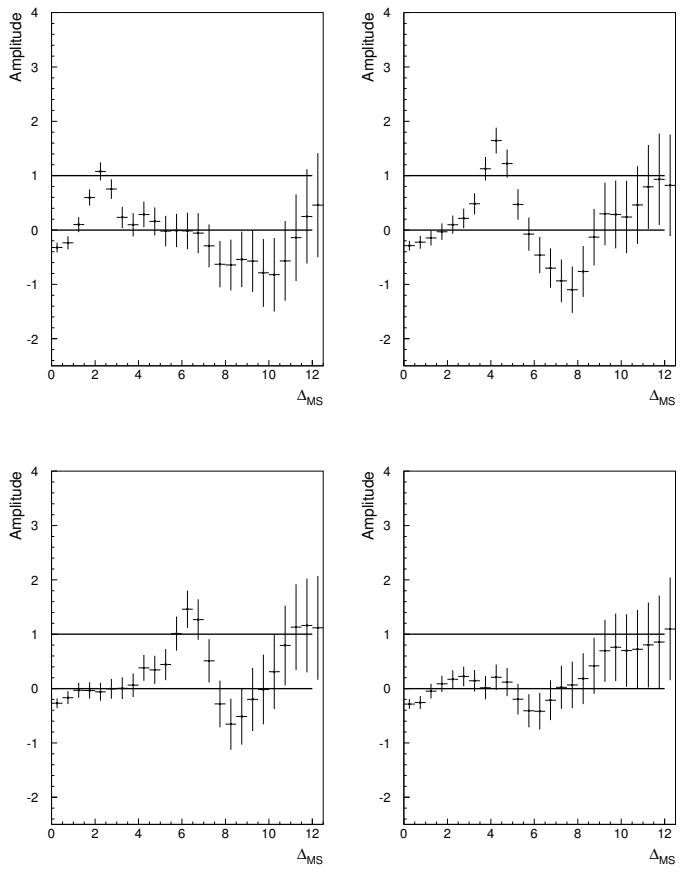

Figure 5.15: Result of an amplitude fit to 300k $b\bar{b}$ events with simulated $\Delta m_s = 2.25$, 4.25, 6.25, 8.25, ps^{-1} (Top left to bottom right).

calculated for every B_s event the probability $\mathcal{P}_{change}$ to change the sign of the decay flavour. Based on a random number R between zero and one, the decay probability P_{decay} was transformed into $1 - P_{decay}$ if $R < \mathcal{P}_{change}$ is fulfilled. $\mathcal{P}_{change}$ was calculated in the following way:

$$\mathcal{P}^{unmix}_{tar} = \frac{1}{2}(\cos(\omega_{tar} t_{tr}) + 1) \quad , \mathcal{P}^{mix}_{tar} = 1 - \mathcal{P}^{unmix}_{tar} \tag{5.20}$$

$$\mathcal{P}^{unmix}_{sim} = \frac{1}{2}(\cos(\omega_{sim} t_{tr}) + 1) \quad , \mathcal{P}^{mix}_{sim} = 1 - \mathcal{P}^{unmix}_{sim}) \tag{5.21}$$

$$P_{change} = \begin{cases} (\mathcal{P}_{sim}^{unmix} - \mathcal{P}_{tar}^{unmix})/\mathcal{P}_{sim}^{unmix} & : \quad \mathcal{P}_{tar}^{mix} > \mathcal{P}_{sim}^{mix} \wedge P_{mix} = 0 \\ (\mathcal{P}_{sim}^{mix} - \mathcal{P}_{tar}^{mix})/\mathcal{P}_{sim}^{mix} & : \quad \mathcal{P}_{tar}^{mix} \leq \mathcal{P}_{sim}^{mix} \wedge P_{mix} = 1 \end{cases} \tag{5.22}$$

ω_{sim} denotes the original Δm_s from the simulation, ω_{tar} the desired frequency, t_{tr} the true proper decay time and P_{mix} equals one if the simulated event is mixed otherwise it is zero. In all other cases P_{change} is set to zero.

The amplitude fit result for signal frequencies $\Delta m_s = 2.25, 4.25, 6.25, 8.25, \text{ps}^{-1}$ is shown in the four plots of figure 5.15 from top left to bottom right. The fit behaves as expected and a clear signal is visible for the first three frequencies. In the case of $\Delta m_s = 8.25, \text{ps}^{-1}$ the error of the amplitude at this point is already ≈ 0.6 reaching the sensitivity of the analysis.

The like-sign fit distributions for the different Δm_s values is shown in figure 5.16. The amplitude A was set to the value obtained by the fit at the corresponding frequency $\omega = \Delta m_s$. The like-sign ratio distributions confirm the results from the amplitude fit, as a nice oscillation is clearly visible in the first three cases and no real statement can be made in the fourth case.

5.7.4 Modelling the Data

As already stated in section 5.3, modelling the mixing purity $\mathcal{P}_{comb}$ is very important for any B_s oscillation analysis. In the B_d^0 case this was achieved by introducing two additional free slope fit parameters calibrating the measurement. This is not possible for the amplitude method.

The problem can be solved in a different way. The B_s decay slopes are listed in the right column of table 5.3. The slopes for B_s and B_d^0 are equal within 5%, also the purity vs. efficiency behaviour for both mesons is the same. This can be understood from the fact that only the separation of the primary B decay vertex from the cascade D vertex counts for the flavour determination. The B_s decaying into Φ and K^+K^- inhibits a usage of the Kaon identification for the flavour determination. Finally it can be seen that a calibration of the B_d^0 part leads within 5% to a calibration of the B_s part if the same slope correction is applied. Studies of simulation and data have shown that, equal to the B_d^0 oscillation case, the charged B^+ mesons need a separate correction slope. The individual correction slopes for the different data samples are listed in table 5.7.

The like-sign ratio plot for the full data sample of 408k events from the years 1992 to 2000 is plotted in figure 5.17. The fit expectation is plotted for amplitude $A = 0$. All physical input variables are taken either from the LEP

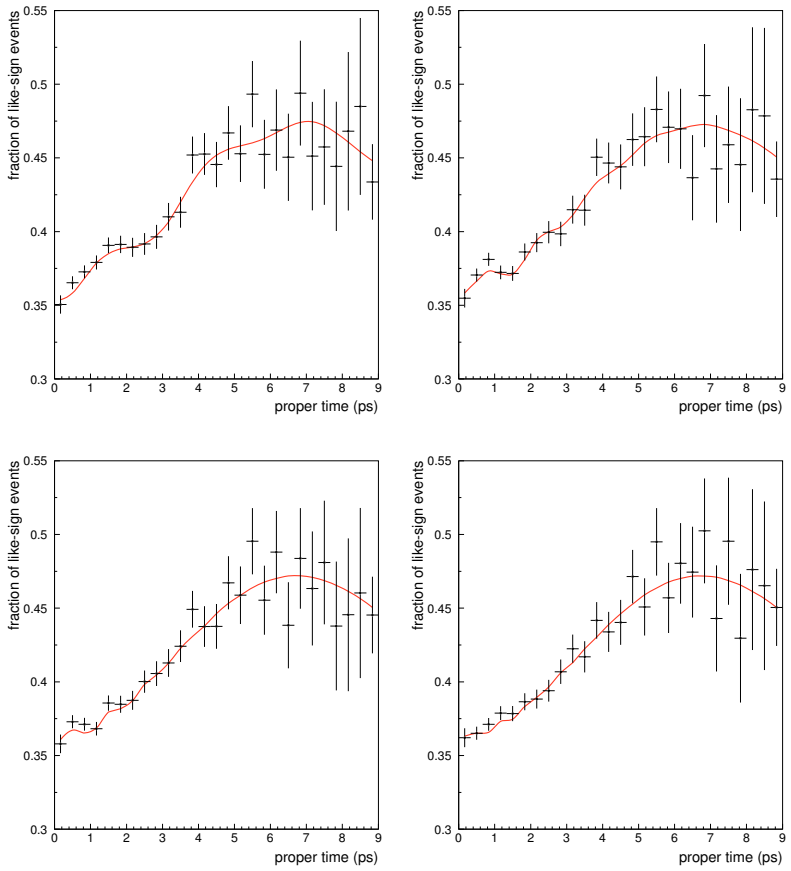

Figure 5.16: Taking the result of the amplitude fit, the like-sign ratio plots for the different Δm_s = 2.25, 4.25, 6.25, 8.25, ps^{-1} values are shown. The amplitude A is set to the value fitted at the signal Δm_s in each plot.

Heavy Flavour working group [OG01] or from the latest PDG book [GAA+00] in the same way as in the B_d^0 oscillation measurement section.

5.7.5 The Amplitude Fit

The amplitude fit described in the previous sections is performed on the full data sample from 1992 to 2000 with 408k events in total. The fit is done for $\omega_{B_s} = 0.25\,\text{ps}^{-1}$ in steps of $0.5\,\text{ps}^{-1}$ up to $20.25\,\text{ps}^{-1}$. Using the amplitude and the error it is possible to obtain the 95 % CL exclusion region. This region corresponds to $A + 1.645\sigma_A$. This curve is shown in figure 5.18 in light grey. From the data one can conclude that the time dependence of the $B_s - \bar{B}_s$ oscillations is not seen and the frequency Δm_s is not measured. A limit

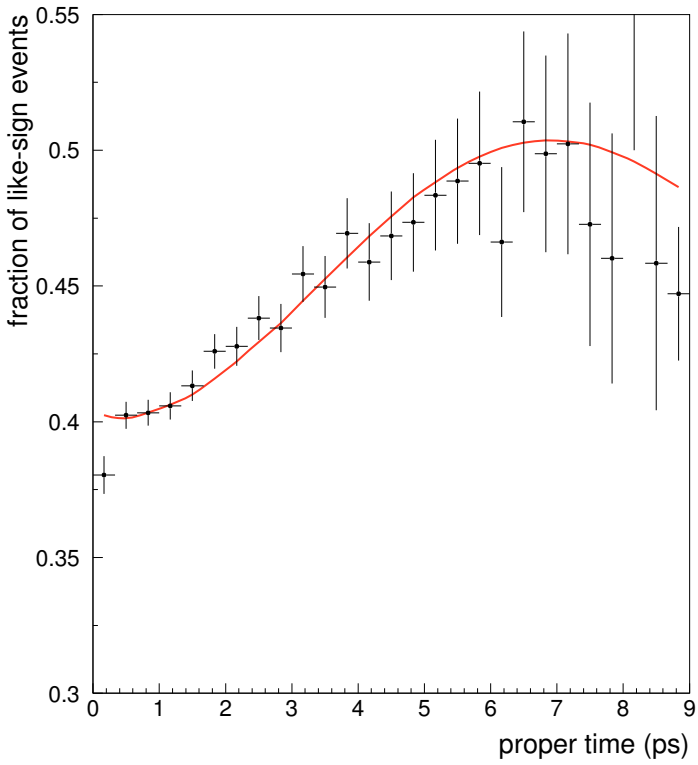

Figure 5.17: The like-sign ratio for the full data sample 1992 to 2000. The unknown B_s oscillation contribution is neglected. Therefore the plot shows the good description of the data by the likelihood parametrisation.

on the mass difference of the two physical states ($\equiv$ frequency) neglecting systematic errors can be put:

$$\Delta m_s > 4.9\,\mathrm{ps}^{-1} \quad \text{at } 95\ \%\ \mathrm{CL} \tag{5.23}$$

Using the amplitude error σ_A one can extract the sensitivity or expected limit on Δm_s at 95 % CL[9]:

$$\text{Sensitivity} = 7.8\,\mathrm{ps}^{-1} \tag{5.24}$$

The results including systematic errors are provided in the next section.

[9]Defined as: $1.645\sigma_A = 1$

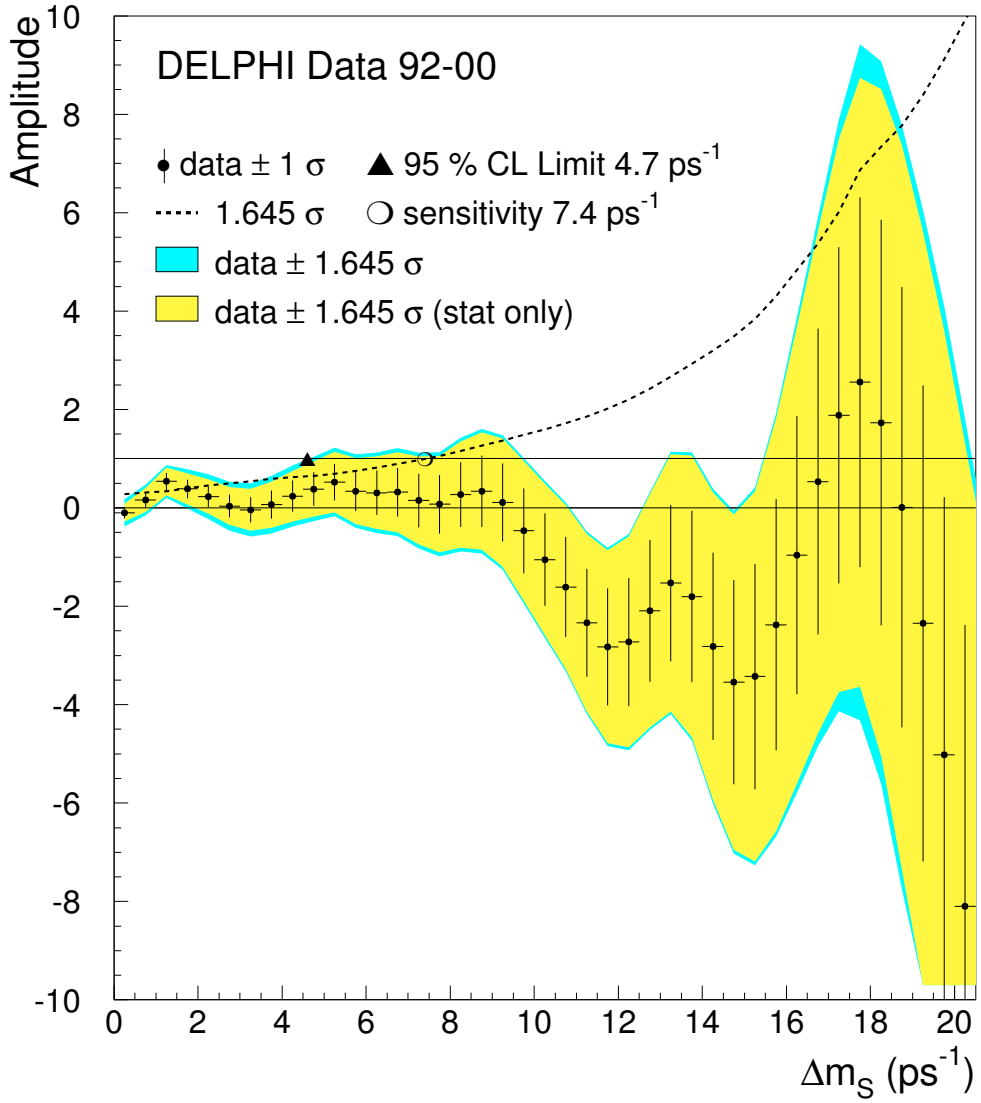

Figure 5.18: The final amplitude fit for the full data set 1992 to 2000 with 408k events in total. The dotted line is the sensitivity curve $1.645\sigma_{tot}$. The light grey area corresponds to $A \pm 1.645\sigma_{stat}$, while the dark area includes systematic errors, $A \pm 1.645\sigma_{tot}$. A 95 % CL limit can be extracted with $\Delta m_s > 4.7\,\mathrm{ps}^{-1}$ and an achieved sensitivity of $7.4\,\mathrm{ps}^{-1}$

Sample	Neutral meson slope	Charged meson slope
Data 1992/93	0.90 ± 0.05	0.68 ± 0.06
Data 1994/95	0.92 ± 0.03	0.57 ± 0.05
Data 1996/00	0.90 ± 0.08	0.60 ± 0.11

Table 5.7: The data correction slopes for the different years. In the second column the corrections for the neutral B hadrons are listed, while in the third column the corrections for the charged B^+ mesons are listed.

5.7.6 Discussion of Systematic Errors

The systematic errors were studied by changing one parameter at a time, i.e. f_{B_s} and redoing the full amplitude fit. The systematic error is evaluated in the following way [HGA97]:

$$\sigma_A^{syst} = A_1 - A_0 + (1 - A_0)\frac{\sigma_{A_1}^{stat} - \sigma_{A_0}^{stat}}{\sigma_{A_1}^{stat}}, \tag{5.25}$$

where $A_0(A_1)$ and $\sigma_{A_0}(\sigma_{A_1})$ denote the fitted amplitude and the error before (after) changing the parameter.

The following parameters have been varied:

- The B_s production fraction f_{B_s} within its error of $\pm 1.5\%$. The other fractions are adjusted accordingly. The number is slightly bigger than the proposed number of the LEP working group [OG01] of $\pm 1.1\%$. The bigger error takes into account the effect of shifting the relative fractions of the B hadrons in the individual decay length class. This is one of the strongest contributions, as the relative uncertainty of 15% directly leads to a systematic error of this amount.

- The tagging power is varied by $\pm 10\%$ by changing the slope correction factor. The 10 % relative change is derived from the statistical uncertainties listed in table 5.7.

- The decay length resolution is varied by $\pm 10\%$, by changing the width of the Gaussians in the decay length parameterisation. This effect gives the biggest contribution to the systematic error for very high Δm_s. The chosen variation range is derived from the fine tuning of the track impact parameter, which provides a handle for the agreement of the track covariance matrices between simulation and data done for the Delphi b-tagging [BM95]. Sensitive to this agreement is the decay length resolution and the expected decay length error, which is plotted

in figure 5.19. One can see that the good agreement between data and simulation is valid for the different vertex algorithms and years. Therefore a variation of 10 % is quite reasonable and conservative.

- The momentum resolution is varied by ±10%, by changing the width of the Gaussians in the parameterisation. The contribution to the systematic error is quite weak over the full Δm_s range. A detailed study on momentum resolution effects was done by a separate analysis, which measures the B hadron fragmentation function inclusively [Ker02]. Based on this analysis the variation of the momentum resolution is chosen.

- The charm and light quark background is changed by ±20%. Their contribution to the systematic error is very weak, except for very small values of Δm_s. An analysis, which is built on a double tag measurement method on the data, derived the relative deviation of the b-tagging between data and simulation [Sch02]. Based on this analysis the variation of the charm-quark and light-quark background is estimated.

Other source for systematic errors, like B hadron lifetimes and acceptance function, were also studied. It was found that their systematic error contribution is negligible as these terms drop out completely in first order due to the fact, that a ratio fit is made.

The numerical stability of the convolution was also tested, as already discussed in section 3.2.6. A breakdown of the individual systematic error contributions can be found in table 5.8.

As final result on the mass difference of the two physical states,

$$\Delta m_s > 4.7\,\mathrm{ps}^{-1} \quad \text{at } 95\ \%\ \mathrm{CL} \tag{5.26}$$

is extracted. The sensitivity or expected limit on Δm_s at 95 % CL is:

$$\text{Sensitivity} = 7.4\,\mathrm{ps}^{-1} \tag{5.27}$$

The result of the fit including the systematic errors is plotted in figure 5.18 in dark grey. The systematic error is at most 25 % of the statistical error for very small Δm_s and approximately 10 % for higher Δm_s.

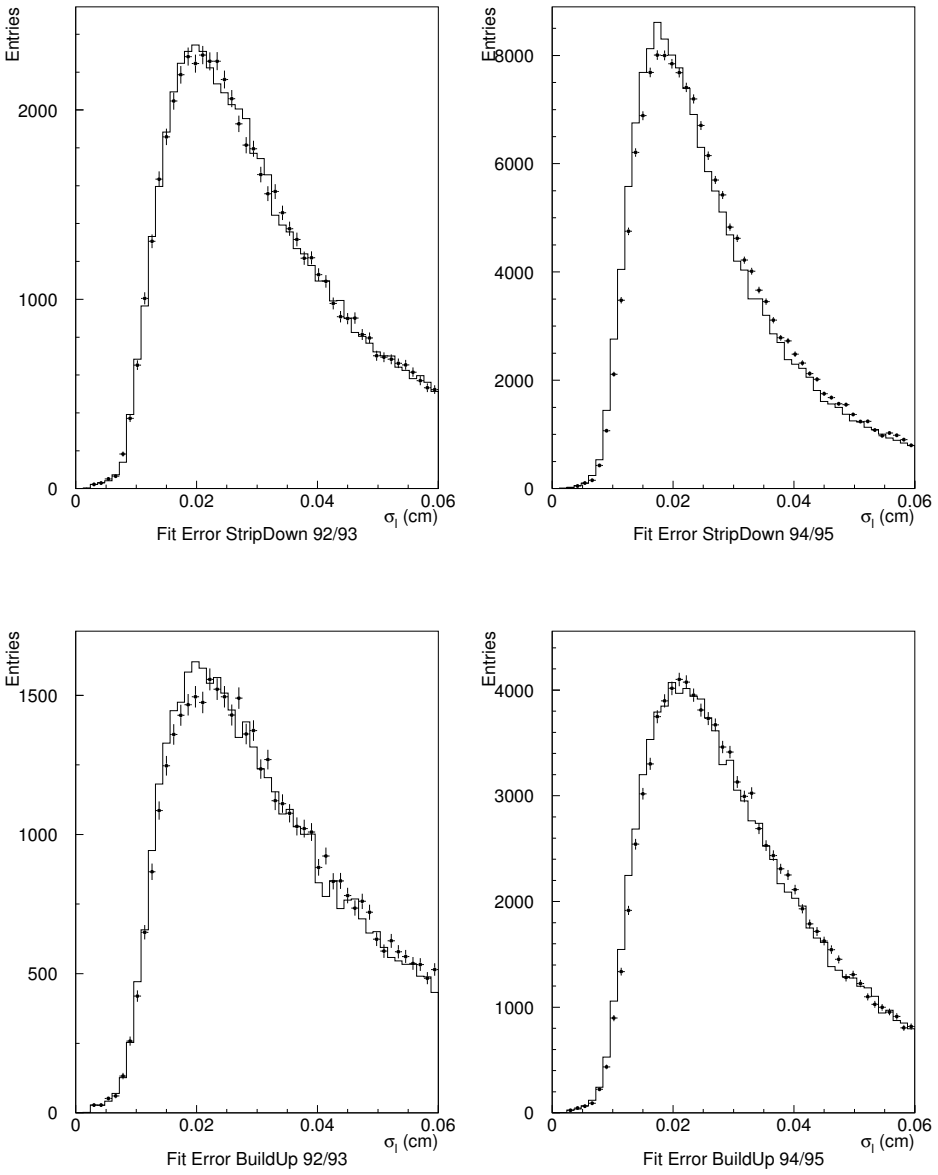

Figure 5.19: The expected decay length error resulting from the vertex fitting procedures. A good agreement is visible for both vertex algorithms. Although the distributions are quite different for the different years, the agreement between data and simulation is good. Therefore the assignment of a 10 % variation to the systematic error from the decay length resolution is reasonable and quite conservative.

Δm_s	$0.25\,\mathrm{ps}^{-1}$	$6.25\,\mathrm{ps}^{-1}$	$12.25\,\mathrm{ps}^{-1}$	$18.25\,\mathrm{ps}^{-1}$
Amplitude	-0.099	0.307	-2.730	1.733
σ_{stat}	± 0.111	± 0.450	± 1.303	± 4.122
Momentum Resolution	$^{+0.006}_{-0.007}$	$^{+0.016}_{-0.013}$	$^{+0.062}_{-0.064}$	$^{+0.141}_{-0.255}$
Decay Length Resolution	$^{+0.024}_{-0.022}$	$^{+0.063}_{-0.061}$	$^{+0.136}_{-0.106}$	$^{+1.471}_{-0.710}$
B_s fraction	$^{+0.026}_{-0.020}$	$^{+0.198}_{-0.155}$	$^{+0.243}_{-0.193}$	$^{+0.755}_{-0.544}$
Tagging Purity	$^{+0.110}_{-0.106}$	$^{+0.035}_{-0.033}$	$^{+0.094}_{-0.090}$	$^{+0.418}_{-0.400}$
Charm and Light Background	$^{+0.021}_{-0.021}$	$^{+0.010}_{-0.010}$	$^{+0.017}_{-0.017}$	$^{+0.087}_{-0.088}$
σ_{total}	0.169	0.497	1.337	4.463

Table 5.8: Systematic error contribution to the full 1993 to 2000 data set amplitude fit. The total error was calculated using only the positive systematic error part, as these contributions are relevant for the limit and the sensitivity of the analysis.

Chapter 6

Conclusion

An inclusive neutral B meson oscillation analysis was performed using the full DELPHI data set taken on the Z^0 from 1992 to 2000.

After cuts in data quality, momentum and vertex reconstruction and after removal of all high p_T lepton events, which are treated separately in a different analysis [KPP+01], a sample of 408k events was available. Dedicated production and decay tag algorithms, momentum and decay length reconstruction algorithms were developed, partly based on the BSAURUS package [ABF+01].

A likelihood fit was performed to extract the B_d^0 oscillation frequency. Systematic uncertainties were mostly compensated by the additional free fit parameters describing the tagging power of charged and neutral B mesons in the fit. The mass difference of the two physical B_d^0 states was measured to be:

$$\Delta m_d = 0.501 \pm 0.013\,(\text{stat}) \pm 0.018\,(\text{sys})\ \text{ps}^{-1}$$

The total error was therefore $0.022\,\text{ps}^{-1}$

Within the amplitude method [HGA97] a likelihood fit searching for B_s oscillations was performed on the same data set. The time dependence of the B_s oscillations was not resolved but a limit on the mass difference of the two physical B_s states was set at:

$$\Delta m_s > 4.7\,\text{ps}^{-1} \quad \text{at } 95\ \%\ \text{CL}.$$

Using the assigned error on the oscillation amplitude the sensitivity, or the expected limit on Δm_s at 95% CL was found to be:

$$\text{Sensitivity} = 7.4\,\text{ps}^{-1}$$

List of Figures

List of Tables

Bibliography

[A+74] J. J. Aubert et al., *EXPERIMENTAL OBSERVATION OF A HEAVY PARTICLE J*, Phys. Rev. Lett. **33** (1974), 1404–1406.

[A+87] H. Albrecht et al., *OBSERVATION OF B0 - anti-B0 MIXING*, Phys. Lett. **B192** (1987), 245.

[A+91] P. Aarnio et al., *The DELPHI detector at LEP*, Nucl. Instrum. Meth. **A303** (1991), 233–276.

[A+92] W. Adam et al., *Design and performance of the DELPHI data acquisition system*, IEEE Trans. Nucl. Sci. **39** (1992), 166.

[A+96a] P. Abreu et al., *Measurement of the B(d)0 oscillation frequency using kaons, leptons and jet charge*, Z. Phys. **C72** (1996), 17–30.

[A+96b] ———, *Performance of the DELPHI detector*, Nucl. Instrum. Meth. **A378** (1996), 57–100.

[A+96c] M. Acciarri et al., Phys. Lett. (1996), no. B383, 487–498.

[A+00] T. Affolder et al., *A measurement of sin(2beta) from B –¿ J/psi K0(S) with the CDF detector*, Phys. Rev. **D61** (2000), 072005.

[A+01a] K. Abe et al., *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001), 091802.

[A+01b] B. Aubert et al., *Observation of CP violation in the B0 meson system*, Phys. Rev. Lett. **87** (2001), 091801.

[ABF+01] T. Allmendinger, G. Barker, M. Feindt, C. Haag, and M. Moch, *BSAURUS: A package for inclusive B reconstruction in DELPHI.*

[Abr93] 1993, Proceedings of the 5th International Symposium on Heavy Flavour Physics.

[AFM99] Zoltan Albrecht, Michael Feindt, and Markus Moch, *MACRIB: High efficiency - high purity hadron identification for DELPHI.*

[AGIS83] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *PARTON FRAGMENTATION AND STRING DYNAMICS*, Phys. Rept. **97** (1983), 31.

[B$^+$] A. W. Borgland et al., *Study of $B_s^0 - \overline{B_s^0}$ oscillations using inclusive leptons with large P_t*, 29th International Conference on High-EnergyPhysics, Vancouver, Canada, 23-29 Jul 1998.

[B$^+$81] C. Bebek et al., *EVIDENCE FOR NEW FLAVOR PRODUCTION AT THE UPSILON (4S)*, Phys. Rev. Lett. **46** (1981), 84.

[B$^+$91] J. Buytaert et al., *The Forward muon detector of the DELPHI experiment at LEP*, Nucl. Instrum. Meth. **A310** (1991), 596–606.

[B$^+$93] D. Buskulic et al., *Observation of the time dependence of B(d)0 - anti-B(d)0 mixing*, Phys. Lett. **B313** (1993), 498–508.

[B$^+$95] V. Bocci et al., *Basic concepts and architectural details of the DELPHI trigger system*, IEEE Trans. Nucl. Sci. **42** (1995), 837.

[B.96] Murray B., *Measurement of the beam position in DELPHI*, DELPHI-PHYS **590** (1996).

[B$^+$01] H. N. Brown et al., *Precise measurement of the positive muon anomalous magnetic moment*, Phys. Rev. Lett. **86** (2001), 2227–2231.

[BA99] G. Boix and D. Abbaneo, *The B/s oscillation amplitude analysis*, JHEP **08** (1999), 004.

[BBBD92] E. Bagan, Patricia Ball, V. M. Braun, and H. G. Dosch, *QCD sum rules in the effective heavy quark theory*, Phys. Lett. **B278** (1992), 457–464.

[BJW90] Andrzej J. Buras, Matthias Jamin, and Peter H. Weisz, *LEADING AND NEXT-TO-LEADING QCD CORRECTIONS TO ϵ PARAMETER AND $B0 - \bar{B}0$ MIXING IN THE PRESENCE OF A HEAVY TOP QUARK*, Nucl. Phys. **B347** (1990), 491–536.

[Blo] Craig A. Blocker, *Measurement of sin(2beta) from j/psi k(s) decays*, To be published in the proceedings of 3rd Workshop on Physics and Detectors for DAPHNE (DAPHNE 99), Frascati, Italy, 16-19 Nov 1999.

[BM95] G. Borisov and C. Mariotti, *Fine tuning of track impact parameter resolution of the DELPHI detector*, DELPHI Note (1995), no. 142, PHYS 567.

[BM96] ———, *Fine tuning of track impact parameter resolution of the DELPHI detector*, Nucl. Instrum. Meth. **A372** (1996), 181–187.

[Bor] G. V. Borisov, *Lifetime tag of events with B hadrons with the DELPHI detector*, IFVE-94-98.

[Bur01] Andrzej J. Buras, *Flavor dynamics: CP violation and rare decays*.

[C+01] M. Ciuchini et al., *2000 CKM-triangle analysis: A critical review with updated experimental inputs and theoretical parameters*, JHEP **07** (2001), 013.

[Cab63] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963), 531–532.

[CCFT64] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *EVIDENCE FOR THE 2 PI DECAY OF THE K(2)0 MESON*, Phys. Rev. Lett. **13** (1964), 138–140.

[CKKW01] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001), 063.

[CM99] Andrzej Czarnecki and William J. Marciano, *Lepton anomalous magnetic moments: A theory update*, Nucl. Phys. Proc. Suppl. **76** (1999), 245–252.

[Cow98] G. Cowan, *Statistical Data Analysis*, Oxford Science Publications, 1998.

[Dav97] C. T. H. Davies, *NRQCD and HQET*, Nucl. Phys. Proc. Suppl. **53** (1997), 392–394.

[del89] *DELPHI Data Analysis Program (DELANA) User's Guide*, DELPHI Note (1989), no. 44, PROG 137.

[Fey49] R. P. Feynman, *Space-time approach to quantum electrodynamics*, Phys. Rev. **76** (1949), 769–789.

[FGML73] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler, *ADVANTAGES OF THE COLOR OCTET GLUON PICTURE*, Phys. Lett. **B47** (1973), 365–368.

[Fis36] R. Fisher, *The use of multiple measurement in taxonomic problems*, Annals of Eugenics **7** (1936), 179–188.

[FKP96] M. Feindt, C. Kreuter, and O. Podobrin, *ELEPHANT Reference Manual*, DELPHI Note **PROG 217** (1996), no. 82.

[For93] R. Forty, CERN-PPE/93-165 (1993).

[G+96] Alexander G. et al., Z. Phys. (1996), no. C72, 377–388.

[GAA+00] D.E. Groom, M. Aguilar-Benitez, C. Amsler, R.M. Barnett, P.R. Burchat, C.D. Carone, C. Caso, G. Conforto, O. Dahl, M. Doser, S. Eidelman, J.L. Feng, L. Gibbons, M. Goodman, C. Grab, A. Gurtu, K. Hagiwara, K.G. Hayes, J.J. Hernández, K. Hikasa, K. Honscheid, C. Kolda, M.L. Mangano, A.V. Manohar, A. Masoni, K. Mönig, H. Murayama, K. Nakamura, S. Navas, K.A. Olive, L. Pape, A. Piepke, M. Roos, M. Tanabashi, N.A. Törnqvist, T.G. Trippe, P. Vogel, C.G. Wohl, R.L. Workman, W.-M. Yao, B. Armstrong, J.L. Casas Serradilla, B.B. Filimonov, P.S. Gee, S.B. Lugovsky, F. Nicholson, K.S. Babu, D. Besson, O. Biebel, P. Bloch, R.N. Cahn, A. Cattai, R.S. Chivukula, R.D. Cousins, T. Damour, K. Desler, R.J. Donahue, D.A. Edwards, J. Erler, V.V. Ezhela, A. Fassò, W. Fetscher, D. Froidevaux, M. Fukugita, T.K. Gaisser, L. Garren, S. Geer, H.-J. Gerber, F.J. Gilman, H.E. Haber, C. Hagmann, I. Hinchliffe, C.J. Hogan, G. Höhler, P. Igo-Kemenes, J.D. Jackson, K.F. Johnson, D. Karlen, B. Kayser, S.R. Klein, K. Kleinknecht, I.G. Knowles, E.W. Kolb, P. Kreitz, R. Landua, P. Langacker, L. Littenberg, D.M. Manley, J. March-Russell, T. Nakada, H.R. Quinn, G. Raffelt, B. Renk, L. Rolandi, M.T. Ronan, L.J. Rosenberg, H.F.W. Sadrozinski, A.I. Sanda, M. Schmitt, O. Schneider, D. Scott, W.G. Seligman, M.H. Shaevitz, T. Sjöstrand, G.F. Smoot, S. Spanier, H. Spieler, M. Srednicki, A. Stahl, T. Stanev, M. Suzuki, N.P. Tkachenko, M.S. Turner, G. Valencia, K. van

Bibber, R. Voss, D. Ward, L. Wolfenstein, and J. Womersley, *Review of Particle Physics*, The European Physical Journal **C15** (2000), 1+.

[GIM70] S. L. Glashow, J. Iliopoulos, and L. Maiani, *WEAK INTERACTIONS WITH LEPTON - HADRON SYMMETRY*, Phys. Rev. **D2** (1970), 1285–1292.

[Gla] A. Glazov, *Measurement of direct CP violation in the neutral kaon system*, Prepared for 28th SLAC Summer Institute on Particle Physics: Neutrinos from the Lab, the Sun, and the Cosmos (SSI 2000), Stanford, California, 14-25 Aug 2000.

[Gla61] S. L. Glashow, *PARTIAL SYMMETRIES OF WEAK INTERACTIONS*, Nucl. Phys. **22** (1961), 579–588.

[GM87] Thomas D. Gottschalk and Duncan A. Morris, *A NEW MODEL FOR HADRONIZATION AND e+ e- ANNIHILATION*, Nucl. Phys. **B288** (1987), 729.

[GMP55] M. Gell-Mann and A. Pais, Phys. Rev. **97** (1955), 1387.

[Gro01] LEP Higgs Working Group, *Search for the standard model Higgs boson at LEP.*

[H^+] H. Herr et al., *The Results of the combined beam test of the DELPHI hadron calorimeter, the forward electromagnetic calorimeter and barrel muon chambers (pi+, e+ runs)*, HU-SEFT-1990-07.

[H^+77] S. W. Herb et al., *OBSERVATION OF A DIMUON RESONANCE AT 9.5-GeV IN 400-GeV PROTON - NUCLEUS COLLISIONS*, Phys. Rev. Lett. **39** (1977), 252–255.

[H.92] Niederreiter H., *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM (1992), Philadephia.

[Hal60] J.H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals.*, Numer. Math. **2** (1960), 84–90.

[Haw01] R. Hawkings, *LEP measurements of |V(cb)| and |V(ub)|*, Nucl. Instrum. Meth. **A462** (2001), 126–131.

[Haz] M. Hazumi, *Measurement of the CP violation parameter sin(2phi(1)) in B0/d decays.*

[HGA97] Moser H.-G. and Roussaire A., Nucl. Instrum. Meth. **A384** (1997), 491.

[Hig64] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964), 132–133.

[HKP91] J. Hertz, A. Krogh, and R. Palmer, *Introduction to the theory of neural computation*, Studies in the sciences of complexity, Addison-Wesley Publishing.

[HQ] ed. Harrison, P. F. and ed. Quinn, H. R., *The BaBar physics book: Physics at an asymmetric B factory*, Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[J.00] Bossert J., *Quasi-Zufallszahlen und ihre Anwendung in der Hochenergiephysik*, IEKP-KA **2001-11** (2000).

[Jam94] F. James, *RANLUX: A FORTRAN implementation of the high quality pseudorandum number generator of Luscher*, Comp. Phys. Commun. **79** (1994), 111–114.

[JJR97] F. James, Hoogland J., and Kleiss R., *Multidimensional sampling for simulation and integration : measure discrepancies, and quasi-random numbers*, Comp. Phys. Comm. **99** (1997), 180–220.

[Kan] G. Kane, *MODERN ELEMENTARY PARTICLE PHYSICS*, REDWOOD CITY, USA: ADDISON-WESLEY (1987) 344 P. (THE ADVANCED BOOK PROGRAM).

[Ker02] U. Kerzel, *First measurement of the b− quark fragmentation function $f(z)$ in Z^0 decays with the DELPHI detector at LEP I*, Master's thesis, Universität Karlsruhe, 2002, To be published.

[Klu01] P. Kluit, *Calibration of the tagging algorithms for an oscillation analysis*, Private Communications (2001).

[KM73] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973), 652.

[KPP+01] P. Kluit, F. Parodi, Roudeau P., Stocchi A., and A. Villa, *Search for $B_s - \bar{B}_s$ oscillations in DELPHI using high p_t leptons*, DELPHI CONF **055** (2001), no. 483.

[L+56] E. E. Lande et al., Phys. Rev. **103** (1956), 1901.

[Lee69] Benjamin W. Lee, *Renormalization of the sigma model*, Nucl. Phys. **B9** (1969), 649–672.

[Lip] I. Lippi, *B/d0 and B/s0 oscillation with DELPHI*, Given at International Europhysics Conference on High-Energy Physics (HEP 97), Jerusalem, Israel, 19-26 Aug 1997.

[LL00] Laurent Lellouch and C. J. David Lin, *Standard model matrix elements for neutral B meson mixing and associated decay constants.*

[Lon92] Leif Lonnblad, *ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model*, Comput. Phys. Commun. **71** (1992), 15.

[Mar] G. Martinelli, *B physics, the HQET and CP violation*, Given at 2nd International Conference on Phenomenology of Unification: From Present to Future, Rome, Italy, 21-24 Apr 1997.

[Mor89] Duncan A. Morris, *HEAVY QUARK FRAGMENTATION FUNCTIONS IN A SIMPLE STRING MODEL*, Nucl. Phys. **B313** (1989), 634.

[Mos93] H.G. Moser, CERN-PPE/93-164 (1993).

[MR81] J. L. McClelland and D. E. Rumelhart, *An interactive activation model of context effects in letter perceptions*, Psychological Review **88** (1981), 375–407.

[MT85] G. Marchesini and Luca Trentadue, *SIMULATION OF PARTON CASCADE AND PERTURBATIVE QCD*, Phys. Lett. **B164** (1985), 150.

[MW84] G. Marchesini and B. R. Webber, *SIMULATION OF QCD JETS INCLUDING SOFT GLUON INTERFERENCE*, Nucl. Phys. **B238** (1984), 1.

[OG01] LEP B Oscillation and Lifetime Working Group, *http://lepbosc.web.cern.ch/lepbosc.*

[P+00] Abreu P. et al., *Study of $B_s - \bar{B}_s$ oscillations and B_s lifetimes using hadronic decays of B_s mesons*, Eur. Phys. J. **C18** (2000), no. 229.

[P+02] ———, *Search for $B_s - \bar{B}_s$ oscillations and a measurement of $B_d^0 - \bar{B}_d^0$ oscillations using events with a soft lepton or an inclusively reconstructed vertex.*, Eur. Phys. J. (2002), To be published.

[Per] D. h. Perkins, *INTRODUCTION TO HIGH-ENERGY PHYSICS*, Reading, Usa: Addison-wesley (1982) 437p.

[PFM85] Billoir P., Fruhwirth, and Regler M., *Track element merging strategy and vertex fitting in complex modular detectors*, Nucl. Instrum. Meth. **A241** (1985), 115–131.

[Pot01] B. Potter, *Combining QCD matrix elements at next-to-leading order with parton showers*, Phys. Rev. **D63** (2001), 114017.

[Pre01] Soren Prell, *Measurements of CP violation, mixing and lifetimes of B mesons with the BaBar detector.*

[PSSZ83] C. Peterson, D. Schlatter, I. Schmitt, and Peter M. Zerwas, *SCALING VIOLATIONS IN INCLUSIVE e+ e- ANNIHILATION SPECTRA*, Phys. Rev. **D27** (1983), 105.

[Q+89] S. Quinton et al., *AN OVERVIEW OF THE FIRST AND SECOND LEVEL TRIGGER OF DELPHI*, IEEE Trans. Nucl. Sci. **36** (1989), 390.

[RHW86] D. E. Rummelhart, G. E. Hinton, and R. J. Williams, *Learning Representation by Back-Propagating Errors*, Nature **323** (1986).

[Ros58] F. Rosenblatt, *The Percepton: A probabilistic Model for Information Storage and Organization in the Brain.*, Psychological Rev. **65** (1958).

[Sac94] Y. Sacquin, *Description of the DELPHI DST Content*, DELPHI Note **161** (1994), no. PROG 210.

[Sal] A. Salam, *WEAK AND ELECTROMAGNETIC INTERACTIONS*, Originally printed in *Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377.

[Sch02] T. Scheidle, *Calibration of the* DELPHI *b-tagging using a double-tag method*, Master's thesis, Universität Karlsruhe, 2002, To be published.

[Sjo94] Torbjorn Sjostrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, Comput. Phys. Commun. **82** (1994), 74–90.

[Sto94] S. Stone (ed.), *B DECAYS*, World Scientific, 1994.

[Tat01] G. Tatishvili, *A new measurement of direct CP-violation by experiment NA48 at CERN*, Nucl. Phys. Proc. Suppl. **96** (2001), 310–315.

[tH71] G. 't Hooft, *RENORMALIZABLE LAGRANGIANS FOR MASSIVE YANG-MILLS FIELDS*, Nucl. Phys. **B35** (1971), 167–188.

[TKPA01] Allmendinger T., P. Kluit, F. Parodi, and Stocchi A., *Search for $B_s - \bar{B}_s$ oscillations in inclusive samples*, DELPHI CONF **054** (2001), no. 482.

[UKJS98] J. Urban, F. Krauss, U. Jentschura, and G. Soff, *Next-to-leading order QCD corrections for the B0 anti-B0 mixing with an extended Higgs sector*, Nucl. Phys. **B523** (1998), 40–58.

[Vai95] Arkady I. Vainshtein, *Review of Selected Topics in HQET.*

[Wea01] James Weatherall, *CP violation in the B0 meson system with BaBar.*

[Web84] B. R. Webber, *A QCD MODEL FOR JET FRAGMENTATION INCLUDING SOFT GLUON INTERFERENCE*, Nucl. Phys. **B238** (1984), 492.

[Wei67] S. Weinberg, *A MODEL OF LEPTONS*, Phys. Rev. Lett. **19** (1967), 1264–1266.

[Wol83] Lincoln Wolfenstein, *PARAMETRIZATION OF THE KOBAYASHI-MASKAWA MATRIX*, Phys. Rev. Lett. **51** (1983), 1945.

[Zha01] Xin Zhao, *Measurement of —V(cb)— and charmless hadronic B decays at CLEO.*

Acknowledgements

So many people have helped me over the last few years that it would be impossible to name them all. Those I do not mention by name have not been forgotten.

I would like to thank my supervisor Michael Feindt, for his advice and support and for his huge insight teaching me more on particle physics than anybody else. Thanks to Günther Quast for the pedagogical improvement and co-supervision of the thesis. Thanks also to the Institute of Experimental Nuclear Physics in Karlsruhe and the DELPHI Experiment for giving me the opportunity to undertake this research and the "Land Baden Württemberg" and the "Graduiertenkolleg für Teilchen- und Astroteilchenphysik" for providing funding in the form of a graduate scholarship.

Among those who have helped me in Karlsruhe, I would like to thank Zoltan and Markus for many useful discussions and for their careful reading of previous drafts. I am also indebted to Gary for providing a running version of Bsaurus all the time, reading drafts of this thesis and for his "refusal" to speak German, which lead to an improvement of my English on the other side. Thanks also to the Delphi B Mixing and Lifetime group at CERN, for their advice and interest, especially Peter for kicking the last errors out of the Likelihood Fit during my stay at NIKHEF.

A special thank belongs to Patrick for his vast knowledge on Linux, he was kind enough to share a pice of it together with the root password - it helped a lot. Thanks also to Jens and Marcel for the discussions on computers and for being throughly entertaining office mates.

Thanks to Fritz, Gerrit, Christoph, Oli and Benno for reminding me that there are other things than particle physics in the world and for the excellent parties we had especially at Fritz's summer residence.

Most important of all, I thank my parents and Sabine for their love and support, for encouraging me when things were difficult and for sharing my excitement when things were going well. This thesis is dedicated to them.