Study of B_s Oscillations with the ALEPH detector at LEP[1,2]

Gaëlle Boix

Universitat Autònoma de Barcelona
Departament de Física
Institut de Física d'Altes Energies
Edifici Cn E-08193 Bellaterra (Barcelona)

and

CERN EP-Division ALE
CH-1211 Geneva 23

June 2001

[1]PhD. Dissertation.

[2]Work supported by the Commission of the European Communities, contract ERBFMBICT982894.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model [1, 2], since its first formulation, has had an enormous phenomenological success, with most of its predictions confirmed by experimental results. However, several questions are still unanswered. For instance, it is not understood why there are three and only three fermion families, the pattern of masses and mixings cannot be explained, the origin of the Standard Model flavour structure is still mysterious, etc. These questions, among others, are presently studied both by theorists and experimentalists in particle physics. In this thesis the quark mixing part of the Standard Model, with the CKM matrix [3, 4], is addressed.

The CKM matrix elements are interesting to be studied on their own, as fundamental parameters of the Lagrangian of the Standard Model. One of the major goals of these studies is to understand the phenomenon of CP violation. The violation of CP is one of the three necessary ingredients to generate the observed excess of matter over antimatter in the universe [5]. However, the amount of CP violation which can be generated by the Standard Model appears to be orders of magnitude below what cosmological observations require [4, 6]. This argument gives strong motivation to search for new physics in that context.

Many experiments are currently studying physical observables in which CP violation would be visible, both with kaons and B mesons. However, insights into CP violation may also be provided by CP conserving experiments. One of the major measurements in this respect is the mixing of neutral B mesons, $\mathrm{B_d}$ and $\mathrm{B_s}$, and consequently the determination of the $|V_{\mathrm{td}}|$ and $|V_{\mathrm{ts}}|$ CKM matrix elements. These two elements are, up to now, only accessible through flavour changing neutral current processes like B mixing. Their accurate measurement is important to test the unitarity of the CKM matrix. Moreover, physics beyond the Standard Model could be revealed with the measurement of the B mixing parameters [7].

The mass difference between $\mathrm{B_d}$ and $\overline{\mathrm{B}}_{\mathrm{d}}$ (Δm_{d}) is now measured with few percent relative accuracy [8]. As of today, no precise measurement of Δm_{s} is available. The ALEPH experiment at LEP has greatly contributed to the overall present knowledge of $\mathrm{B_s}$ oscillations. Together with the other LEP experiments, ALEPH has collected the data sample in which the largest number of $\mathrm{B_s}$ decays can be reconstructed. The detector performance on charged and neutral particle momentum measurement, secondary vertex reconstruction and particle identification are adequate for $\mathrm{B_s}$ oscillation studies.

When I started my work on B_s oscillations, a first hint for an oscillation signal seemed to be appearing from the world combination, but no clear method existed to estimate the probability that this hint originate from a statistical fluctuation. My first contribution was a detailed study [9] of the amplitude method. Analytical expressions for the expected shape of the amplitude were derived and a procedure making use of toy experiments was proposed as a tool to estimate the significance of a possible oscillation signal.

At that time, the single most sensitive B_s oscillations analysis was the ALEPH study with an inclusive semileptonic event sample [10]. The ALEPH data taken at energies close to the Z mass had just been reprocessed and some of the published analyses were being revisited to profit from the improved reconstruction. A critical study of Ref. [10] showed that room for improvement existed beyond the gain from the reprocessing, and encouraged me to perform a new B_s oscillation analysis with an inclusive semileptonic event sample in ALEPH. The main item for which improvement was understood to be possible was the reconstruction of the B meson decay length. However, all the other relevant ingredients were studied in detailed and significantly improved. As a result, the sensitivity of the analysis was sizeably increased. The precision on the measured amplitude at high frequency was improved by more than a factor of two. A hint for a Δm_s measurement, with a statistical significance of ~ 1.5 standard deviations, was obtained with the analysis in Ref. [10]. The new analysis confirmed the hint with a significance increased to slightly more than two standard deviations.

This thesis dissertation is organized as follows. The theoretical framework for B_s oscillations, the CKM matrix, and their connection with CP violation are revised in Chapter 2. The experimental method for B mixing and oscillation studies, together with a detailed discussion on the amplitude method are the subject of Chapter 3. In Chapter 4, the LEP collider and the ALEPH detector are briefly described. Those analysis tools used for the data analysis but not developed by myself are presented in Chapter 5. The core of this thesis is developed in Chapter 6: it contains the description of the inclusive semileptonic event sample selection for the B_s oscillation study. Two complementary selections of event samples enriched in B_s mesons are presented in Chapter 7. The results on B_s oscillations obtained with the three event samples described in this thesis are provided in Chapter 8. The ALEPH combination and the world combination are also presented. Finally, the conclusions are given in Chapter 9.

Chapter 2

Flavour Physics and B-Mixing in the Standard Model

Particle physics has a well established theoretical framework which has been successful for many years: the Standard Model [1, 2]. This model explains and includes basically all the experimental observations made so far, in particular high precision tests, at the per-mil level, which have been performed in the past decade. However, many parameters in the model are not predicted by the theory, and have to be measured experimentally. The particles (or fields) of the Standard Model are divided into gauge fields associated with the gauge groups, and matter fields. All matter fields are fermions of two types distinguished by the kind of interactions they are sensitive to; these particles are quarks and leptons. They appear in three families (or generations) each containing two quarks and two leptons. The families all have the same interactions with the gauge fields; only their masses (and couplings to the Higgs field) are different. It is common to use the term *flavour* to denote the different particles and therefore to say u-flavour, or b-flavour.

The flavour sector of the model studies fermion masses and mixings, and is, at present, one of the least well tested. Several of its relevant parameters, in particular quark mixing, are only known to an accuracy of $\mathcal{O}(30\%)$, while other parameters of the model, in the electroweak sector, are measured with a precision of one percent or even better, and allow a fine test of the Standard Model. The so-called flavour parameters are 13: six quark masses, four parameters related with quark mixing, and in the lepton sector, assuming massless neutrinos, three charged lepton masses. If neutrinos are proved to be massive, three new masses and a set of mixing parameters should be added.

Some of the remaining open questions in the Standard Model are related with the flavour parameters, hence the interest of improving their experimental knowledge. The flavour parameters exhibit a "family hierarchy", *i.e.*, their relative magnitude depends on the family. No theoretical explanation exists for this hierarchy, the phenomenon is called *flavour puzzle*. The strength of flavour changing neutral current processes (FCNC) depends on the flavour parameters. The phenomenon of CP violation is also closely related to flavour physics and it is one of the least tested aspects of the Standard Model. Most of the theoretically proposed extensions of the Standard Model predict new sources for CP violation, and some for FCNC as well. These reasons make important the study of flavour physics inside the Standard

Model, and maybe beyond.

Particle-antiparticle mixing happens in neutral meson systems which are not auto-conjugated (the π^0 meson, for instance, being its own antiparticle, does not mix). Two B mesons have this property: B_d and B_s. From the experimental point of view, the two neutral B mesons have very distinct mixing properties. The B_d system has already been deeply studied. The focus of this thesis is devoted to the B_s system whose experimental knowledge is far behind.

The flavour structure of the Standard Model is presented in the first Section of this Chapter. Stress is put on the aspects still to be understood or measured, and in particular on the Cabbibo-Kobayashi-Maskawa (CKM) matrix [3, 4] which describes the quark mixing. The present state-of-the-art is given with a brief review of methods already used to measure some of the matrix elements. A very brief introduction to CP violation in the Standard Model, which arises from the CKM matrix structure with three families, follows.

The phenomenology of neutral mesons mixing is introduced in Section 2.2. Although the experimental focus of the present thesis is on B_s oscillations, both B_d and B_s are included in the theoretical discussion. Stress is put on the differences between them and their importance from the experimental point of view. A discussion on how B_s mixing is related to some of the elements of the CKM matrix and CP violation is given, as well as the present state-of-the-art.

Finally, some of the possible extensions to the Standard Model are very briefly presented with their implications on neutral B-meson mixing.

2.1 Cabbibo-Kobayashi-Maskawa (CKM) Matrix

In the Standard Model, only the charged current sector, with exchange of a W-boson, can accommodate family and flavour-changing transitions. The interaction Lagrangian density responsible for these transitions in the quark sector, is written

$$\mathcal{L}_{CC} = \frac{g}{2\sqrt{2}} \sum_{i,j} [\bar{u}_i \gamma^\mu (1-\gamma^5) V_{ij} d_j] W_\mu^+ + h.c. \tag{2.1}$$

where V_{ij} are terms of the Cabbibo-Kobayashi-Maskawa (CKM) matrix, and the indices i and j run over the three quark families.

The CKM matrix V_{CKM} is a rotation from the quark mass eigenstates, d, s, and b, to a set of different states d′, s′, and b′, the weak eigenstates. Therefore, if no other quark families exist, the 3×3 CKM matrix should be unitary. A common representation is

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}. \tag{2.2}$$

To a first order approximation, the CKM matrix is simply the identity matrix, so that the dominant transitions are inside a single family: u → d, c → s, t → b. Family-changing transitions between quarks through a $W^\pm$-gauge boson are made possible by the fact that none of the off-diagonal elements is exactly zero. These off-diagonal elements of the matrix can therefore be measured through the experimental study of the family-changing transitions. These transitions occur in some mesons decays, in FCNC processes, as well as in the CP violation inside the Standard Model and constitute the main interest of flavour physics.

2.1.1 Estimation of CKM elements

The values of the CKM matrix elements are, like the fermion masses, fundamental input parameters of the Standard Model. Experimental measurements are needed in order to give insights into their values.

An intense experimental effort is being made to measure the CKM elements. They can be determined basically in three ways, with direct measurements of Standard Model tree-level processes, with indirect measurements of Standard Model loop processes, or applying unitarity, *i.e.*, using some of the relations from $V_{\rm CKM}V_{\rm CKM}^{\dagger} = 1$.

The present knowledge of the matrix elements comes from the following sources:

- $|\boldsymbol{V_{\rm ud}}|$
  Three different methods exist to measure $|V_{\rm ud}|$. *Nuclear superallowed Fermi beta decays:* the method consists of comparing the rate of nuclear beta decays of spin 0^+ and isospin 1 states to that of the muon. This rate is proportional to $|V_{\rm ud}|^2$. The measurement is performed with several nuclei and the average is taken. This method is, at present, the most precise, but it is limited by the uncertainty on the theoretical corrections, especially on nucleus dependent corrections. The result obtained with this method is $|V_{\rm ud}| = 0.9740 \pm 0.0005$ [11]. *Neutron decay:* $|V_{\rm ud}|$ is in this case extracted from the neutron lifetime measurement. The limiting factor is the experimental uncertainty on g_A/g_V. One of the available values is $|V_{\rm ud}| = 0.9790 \pm 0.0017$ [12]. *Pion beta decay:* it is the cleanest method from the theoretical point of view. However, the branching ratio for the process $\pi^+ \to \pi^0 {\rm e}^+ \nu_{\rm e}$ is only of the order of 10^{-8}, and therefore it is limited from the experimental side. The present value obtained is $|V_{\rm ud}| = 0.9679 \pm 0.0161$ [11]. The theoretical uncertainty estimates of the above results are controversial and so is the global average value of $|V_{\rm ud}|$. The present value quoted in the PDG (Particle Data Group) is: $|V_{\rm ud}| = 0.9735 \pm 0.0008$ [13].

- $|\boldsymbol{V_{\rm us}}|$
  Mainly two ways are exploited to determine $|V_{\rm us}|$. K_{e3} *decays:* these decays are: $\rm K^+ \to \pi^0 e^+ \nu_e$ and $\rm K_L^0 \to \pi^- e^+ \nu_e$. Chiral perturbation theory is used to extract the value of $|V_{\rm us}|$ from the measured K_{e3} decay rates. The present estimated value reads $|V_{\rm us}| = 0.2196 \pm 0.0023$, Ref. [14]. *Semileptonic hyperon decays:* in this case the $|V_{\rm us}|$ extraction is dominated by large theoretical uncertainties. The best value quoted in the PDG [13] with this method is $|V_{\rm us}| = 0.2176 \pm 0.0026$, but the value obtained from K_{e3} decays is recommended for use.

- $|\boldsymbol{V_{\rm cd}}|$
  The value of $|V_{\rm cd}|$ can be obtained from single charm production in deep inelastic neutrino (antineutrino)-nucleon scattering, supplemented by measurements of semileptonic branching fractions of charmed mesons. The present average is: $|V_{\rm cd}| = 0.224 \pm 0.016$ [13].

- $|\boldsymbol{V_{\rm cs}}|$
  The value of $|V_{\rm cs}|$ used to be obtained in a similar manner as for $|V_{\rm us}|$, from D_{e3} decays, by comparing the data with the theoretical decay width. This method uses model-dependent form-factors which introduce an important uncertainty in the final

result, which reads $|V_{cs}| = 1.04 \pm 0.16$ [13]. Recent measurements from the ALEPH, DELPHI, and OPAL Collaborations [15] of the inclusive charm production rate in W-boson decays, $R_c^W = \Gamma(W \to cX)/\Gamma(W \to \text{hadrons})$, allow a more precise value to be extracted, $|V_{cs}| = 0.95 \pm 0.08$ [16], which is compatible with the one previously quoted. No average of these two numbers is provided by the PDG.

- **$|V_{cb}|$**
The heavy quark effective theory (HQET) [17] provides a nearly model-independent treatment of B semileptonic decays to charmed mesons, assuming that both the b and c quarks are heavy enough for the theory to apply. The matrix element $|V_{cb}|$ can be determined from B semileptonic decays studies in the HQET framework [18]. There are two experimental methods to obtain $|V_{cb}|$: the *inclusive* method, which uses the semileptonic decay width of b-quark decays, and the *exclusive* method, where $|V_{cb}|$ is extracted by studying the exclusive $\overline{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decay. The latest result from LEP with the inclusive method, $|V_{cb}| = 0.04070 \pm 0.00050_{(exp)} \pm 0.00204_{(th)}$, is largely dominated by the theoretical uncertainties but it is more precise than the exclusive one, $|V_{cb}| = 0.0398 \pm 0.0018_{(exp)} \pm 0.0022_{(th)}$ [19]. The LEP average, $|V_{cb}| = 0.0404 \pm 0.0018$, can be compared to the average from similar measurements performed at CLEO: $|V_{cb}| = 0.0404 \pm 0.0034$ [13]. No clear procedure to average LEP and CLEO results exist and therefore no average is provided by the PDG.

- **$|V_{ub}|$**
Until now, three different methods have been used to measure $|V_{ub}|$. *Endpoint:* at CLEO, with data at the $\Upsilon(4S)$ resonance, the decay of $b \to u\ell\bar{\nu}$ is measured from the lepton energy spectrum above the endpoint of the $b \to c\ell\bar{\nu}$ spectrum. *Exclusive:* decays such as $B \to \pi\ell\bar{\nu}_\ell$ and $B \to \rho\ell\bar{\nu}_\ell$, sensitive to $|V_{ub}|$ are accessible at CLEO. These two methods are, however, limited by theoretical model dependences. *Inclusive:* at LEP the charmless semileptonic B branching fraction can be determined by discriminating $b \to u$ from $b \to c$ decays. The value of $|V_{ub}|$ is then extracted from the measured branching fraction. The latest LEP average available gives $|V_{ub}| = (4.13^{+0.42}_{-0.47\,(\text{stat+exp})}\,{}^{+0.43}_{-0.48\,(b\to c\ \text{syst})}\,{}^{+0.24}_{-0.25\,(b\to u\ \text{syst})} \pm 0.20_{(\text{HQET})}) \times 10^{-3}$ [20]. The CLEO Collaboration has produced an average of their results which can be compared to that of LEP and reads, $|V_{ub}| = (3.25 \pm 0.14_{(\text{stat})}\,{}^{+0.21}_{-0.29\,(\text{syst})} \pm 0.55_{(\text{theo})}) \times 10^{-3}$ [21]. No average of these results is provided by the PDG.

- **$|V_{tb}|$**
No measurement of the $|V_{tb}|$ matrix element has been performed without the assumption of unitarity of the CKM matrix and three quark generations in the Standard Model until now. The ratio $BR(t \to Wb)/BR(t \to Wq)$, where q is a down-type quark, can be measured with the $t\bar{t}$ data collected at the Tevatron. This measurement, with the above assumptions, allows a value of $|V_{tb}|$ to be extracted: $|V_{tb}| = 0.99 \pm 0.15$ [22]. A more accurate, and less model-dependent, estimate is expected using the single top production cross-section with data from the second phase of the Tevatron accelerator at Fermilab. It is also possible to extract the value of $|V_{tb}|$ from electroweak loop corrections without assuming unitarity, the value obtained is $|V_{tb}| = 0.77^{+0.18}_{-0.24}$ [23].

- **$|V_{ts}|$ and $|V_{td}|$**
These two matrix elements are not measured directly at present. Their value can be

estimated from top quark loop correction to some processes (*e.g.* b → sγ), they are also closely related to the B meson mixing, as explained later in Section 2.2.1.

Some comments

If the CKM matrix is unitary, its elements in the first row (Eq. 2.2) should verify $|V_{\rm ud}|^2 + |V_{\rm us}|^2 + |V_{\rm ub}|^2 = 1$. If the values of the matrix elements quoted above are taken, a discrepancy of more than one and a half standard deviations from unitarity is obtained: $|V_{\rm ud}|^2 + |V_{\rm us}|^2 + |V_{\rm ub}|^2 = 0.9959 \pm 0.0025$; the discrepancy is larger, 0.9955 ± 0.0020, if the second value of $|V_{\rm us}|$ is taken instead of the first. Although not felt as a real problem yet [11, 13], it gives strong motivation for pursuing on the $|V_{\rm ud}|$ and $|V_{\rm us}|$ estimates, both from the experimental and theoretical sides. The element $|V_{\rm ub}|$ has the biggest relative uncertainty of the three, however, it is intrinsically too small to make a difference on the unitarity test. A definite deviation from unitarity would be proof of the existence of physics beyond the Standard Model.

Within the Standard Model, all branching fractions of the W-boson depend, at tree level, on the six matrix elements $|V_{\rm qq'}|$ not involving the top quark. The semileptonic branching ratio of the W-boson is inversely proportional to the squared sum of these six matrix elements,

$$1/BR(\mathrm{W} \to \ell\overline{\nu_\ell}) \;=\; 3\{1 + [1 + \alpha_s(M_{\rm W}^2)/\pi] \sum_{i=\mathrm{u,c},\, j=\mathrm{d,s,b}} |V_{ij}|^2\} \tag{2.3}$$

where α_s is the strong coupling constant. Taking the experimental values quoted above for $|V_{\rm ud}|$, $|V_{\rm us}|$, $|V_{\rm ub}|$, $|V_{\rm cd}|$, and $|V_{\rm cb}|$, an estimate for $|V_{\rm cs}|$ can be extracted: $|V_{\rm cs}| = 0.989 \pm 0.016$, as in Ref. [16], which is much more precise than the direct measurement obtained so far.

Only two, $|V_{\rm ud}|$ and $|V_{\rm us}|$, out of nine CKM matrix elements, are measured with an uncertainty smaller than 10%. Being input parameters to the Standard Model, it is important for the global understanding of the theory that the precision on their measurements increases. The aim is to be able to measure all the CKM elements independently and with high precision, and with them test the overall consistency.

2.1.2 CKM matrix and CP Violation

In the framework of the Standard Model, not all nine elements of the CKM matrix are independent. In general, an $N \times N$ unitary matrix has N^2 independent real parameters, $N(N-1)/2$ of which are Euler-type angles and $N(N+1)/2$ are complex phases. In the CKM matrix not all of these parameters are physically meaningful, because, for N quark generations, $2N-1$ phases can be absorbed by the freedom to select the phases of the quark fields. The number of physical parameters is therefore $N^2 - (2N-1) = (N-1)^2$, with $(N-1)(N-2)/2$ complex phases. In the case of $N = 3$ families, as in the Standard Model, only four independent parameters are left, three angles and a single complex phase.

The unitarity condition, which is expressed by the orthonormality of the rows and columns of the matrix,

$$\sum_i |V_{i\alpha}|^2 = 1, \qquad \sum_i V_{i\alpha} V_{i\beta}^* = 0 \;, \tag{2.4}$$

could, in principle, be used to determine the whole matrix once four elements are known. The constraints are such that even a small relative uncertainty in one of the large (and easiest to measure) quantities, completely dominates the smaller elements, rendering such a determination impossible. For this reason, a complete determination of the matrix requires the measurement of some of the smallest elements. These elements are the ones which are measured in processes involving B physics.

The complex phase appearing in the CKM matrix introduces CP violation in the Standard Model (it is also necessary that all quarks of the same charge have different masses [24]). The matrix elements can be parameterized as a function of four independent parameters in many manners. One of the most commonly used is due to Wolfenstein [25], with a generalization to include higher order corrections first presented in Ref. [26]

$$\begin{pmatrix} V_{\rm ud} & V_{\rm us} & V_{\rm ub} \\ V_{\rm cd} & V_{\rm cs} & V_{\rm cb} \\ V_{\rm td} & V_{\rm ts} & V_{\rm tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \quad (2.5)$$

In this parametrization all the matrix elements are expressed as a function of the four real parameters λ, A, ρ, and η. In order to include high order corrections, up to $\mathcal{O}(\lambda^6)$, new variables, $\bar{\rho}$ and $\bar{\eta}$, are defined as $\bar{\rho} = \rho\,(1-\lambda^2/2)$ and $\bar{\eta} = \eta\,(1-\lambda^2/2)$. The expression of some of the CKM matrix elements is modified by the inclusion of $\mathcal{O}(\lambda^6)$ corrections as:

$$\begin{array}{ll} V_{\rm ud} = 1-\frac{\lambda^2}{2} & V_{\rm cd} = -\lambda + \frac{1}{2}A^2\lambda^5[1-2(\rho+i\eta)] \\ V_{\rm cs} = 1-\frac{\lambda^2}{2} - \frac{1}{8}\lambda^4(1+4A^2) & V_{\rm td} = A\lambda^3(1-\bar{\rho}-i\bar{\eta}) \\ V_{\rm ts} = -A\lambda^2 + \frac{1}{2}A(1-2\rho)\lambda^4 - i\eta A\lambda^4 & V_{\rm tb} = 1-\frac{1}{2}A^2\lambda^4\,. \end{array}$$

In the parametrization in Eq. 2.5, $\eta \neq 0$ is needed to incorporate CP violation in the Standard Model, and $\lambda \sim 0.22$ is the sine of the Cabbibo angle.

Recalling the unitarity condition expressed in Eq. 2.4, and applying it to first row and first column, the following equation holds

$$V_{\rm ud}V_{\rm ub}^* + V_{\rm cd}V_{\rm cb}^* + V_{\rm td}V_{\rm tb}^* = 0\,. \quad (2.6)$$

Since V_{ij} are complex numbers, this relation can be represented as a triangle in the plane of complex numbers: it is the *Unitarity Triangle*. It is customary to rescale the triangle dividing all the terms by $|V_{\rm cd}V_{\rm cb}^*|$, which, to a very good accuracy, $\mathcal{O}(\lambda^7)$, is real and equal to $A\lambda^3$. Therefore the triangle can be drawn as in Fig. 2.1.

The values of the triangle sides can be written as

$$\begin{array}{rcccccl} \overline{\rm CA} & \equiv & \dfrac{|V_{\rm ud}V_{\rm ub}^*|}{|V_{\rm cd}V_{\rm cb}^*|} & = & \sqrt{\bar{\rho}^2+\bar{\eta}^2} & = & \left(1-\dfrac{\lambda^2}{2}\right)\dfrac{1}{\lambda}\left|\dfrac{V_{\rm ub}}{V_{\rm cb}}\right|, \quad \text{and} \\ \overline{\rm BA} & \equiv & \dfrac{|V_{\rm td}V_{\rm tb}^*|}{|V_{\rm cd}V_{\rm cb}^*|} & = & \sqrt{(1-\bar{\rho})^2+\bar{\eta}^2} & = & \dfrac{1}{\lambda}\left|\dfrac{V_{\rm td}}{V_{\rm cb}}\right|. \end{array} \quad (2.7)$$

A precise measurement of the sides of the triangle, through the measurement of the CKM matrix elements, verifying that they really form a triangle would constitute a very important test of the Standard Model and of the understanding of the CP violation mechanism. Furthermore, the angles α, β, and γ of the triangle are also experimentally accessible. Any significant

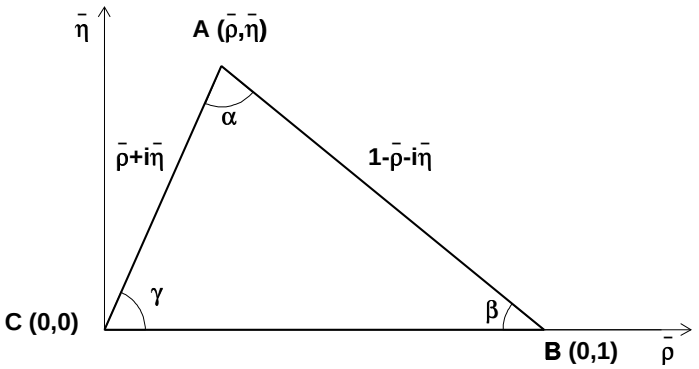

Figure 2.1: Unitarity triangle in the complex $(\bar{\rho}, \bar{\eta})$ plane.

inconsistency between all these measurements would indicate the presence of physics beyond the Standard Model. Completing the picture shown above is today one of the most active domains in particle physics.

2.2 Phenomenology of B mixing and oscillations

Particle anti-particle mixing was first observed and studied in the neutral kaon system [27] in the fifties. It is a phenomenon which has always been of fundamental importance for testing the Standard Model and suggesting extensions to it. Being a FCNC process it involves heavy quark loops and consequently it is a perfect testing ground for heavy flavour physics: from the calculation of the $\mathrm{K}^0_\mathrm{L} - \mathrm{K}^0_\mathrm{S}$ mass difference, the value of the charm quark mass was estimated before its discovery [28], for instance.

Two neutral B-mesons exhibit the phenomenon of mixing; these are $\mathrm{B_d} = (\bar{\mathrm{b}}\mathrm{d})$ and $\mathrm{B_s} = (\bar{\mathrm{b}}\mathrm{s})$. Hereafter they are generically denoted $\mathrm{B_q}$. A $\mathrm{B_q}$ meson can be produced in either of the two flavour-eigenstates $|\mathrm{B_q}\rangle = |(\bar{\mathrm{b}}\mathrm{q})\rangle$ or $|\overline{\mathrm{B}}_\mathrm{q}\rangle = |(\mathrm{b}\bar{\mathrm{q}})\rangle$. Due to electroweak flavour-changing interactions, this initial state evolves into a time-dependent quantum superposition of the two flavour states as, $a(t)|\mathrm{B_q}\rangle + b(t)|\overline{\mathrm{B}}_\mathrm{q}\rangle$, which satisfies the equation

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathcal{H}_\mathrm{eff}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \qquad \text{with} \qquad \mathcal{H}_\mathrm{eff} = M - i\,\frac{\Gamma}{2}\,, \tag{2.8}$$

where $\mathcal{H}_\mathrm{eff}$ is the effective Hamiltonian. The operators M and Γ describe the dispersive and absorptive parts of the $\mathrm{B_q}-\overline{\mathrm{B}}_\mathrm{q}$ mixing, they are called *mass* and *decay* matrices. The B-meson is not a stable particle, and therefore $\mathcal{H}_\mathrm{eff}$, which does not include the decay products, can not be hermitian. The matrices M and Γ instead, are hermitian, and, because of CPT invariance, their diagonal elements are equal: they represent the mass $M_{\mathrm{B_q}}$ and width $\Gamma_{\mathrm{B_q}}$ of the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ flavour states. The effective Hamiltonian can therefore be written as

$$\mathcal{H}_\mathrm{eff} = \begin{pmatrix} M_{\mathrm{B_q}} & M_{12} \\ M^*_{12} & M_{\mathrm{B_q}} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{\mathrm{B_q}} & \Gamma_{12} \\ \Gamma^*_{12} & \Gamma_{\mathrm{B_q}} \end{pmatrix}\,. \tag{2.9}$$

The diagonalization of the effective Hamiltonian gives the two mass eigenstates

$$|\mathrm{B}_{1,2}\rangle = p|\mathrm{B_q}\rangle \pm q|\overline{\mathrm{B}}_\mathrm{q}\rangle \ , \tag{2.10}$$

and the corresponding eigenvalues

$$\lambda_{1,2} = \left(M_{\mathrm{B_q}} - \frac{i}{2}\Gamma_{\mathrm{B_q}}\right) \pm \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right). \tag{2.11}$$

In the above equations two complex coefficients q and p have been introduced, they are defined as

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}} \ , \qquad \text{with} \qquad |q|^2 + |p|^2 = 1 \ \ . \tag{2.12}$$

The time dependence of these eigenstates of well-defined masses $M_{1,2} = \mathrm{Re}(\lambda_{1,2})$ and widths $\Gamma_{1,2} = -2\mathrm{Im}(\lambda_{1,2})$ is given by the phases $e^{-i\lambda_{1,2}t}$. Correspondingly, the time-evolution of pure states is given by

$$\begin{aligned} |\mathrm{B_q}(t)\rangle &= g_1(t)|\mathrm{B_q}(0)\rangle + \frac{q}{p}g_2(t)|\overline{\mathrm{B}}_\mathrm{q}(0)\rangle \ , \quad \text{and} \\ |\overline{\mathrm{B}}_\mathrm{q}(t)\rangle &= g_1(t)|\overline{\mathrm{B}}_\mathrm{q}(0)\rangle + \frac{p}{q}g_2(t)|\mathrm{B_q}(0)\rangle \ , \end{aligned} \tag{2.13}$$

where

$$g_{1,2} = \frac{1}{2}\left(e^{-i\lambda_1 t} \pm e^{-i\lambda_2 t}\right) \ . \tag{2.14}$$

The above equations mean that the flavour states oscillate into each other with time-dependent probabilities:

$$\begin{aligned} \mathcal{P}_{\mathrm{B_q}\to\mathrm{B_q}}(t) &= N\frac{e^{-\Gamma_{\mathrm{B_q}}t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_{\mathrm{B_q}}}{2}t\right) \pm \cos\left(\Delta M_{\mathrm{B_q}}t\right)\right] \ , \quad \text{and} \\ \mathcal{P}_{\mathrm{B_q}\to\overline{\mathrm{B}}_\mathrm{q}}(t) &= N\left|\frac{q}{p}\right|^2\frac{e^{-\Gamma_{\mathrm{B_q}}t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_{\mathrm{B_q}}}{2}t\right) \pm \cos\left(\Delta M_{\mathrm{B_q}}t\right)\right] \ , \end{aligned} \tag{2.15}$$

where

$$\Delta M_{\mathrm{B_q}} = |M_1 - M_2| \, , \qquad \Delta\Gamma_{\mathrm{B_q}} = |\Gamma_1 - \Gamma_2| \ , \tag{2.16}$$

and N is the normalization factor defined by

$$\int_0^\infty dt \left[\mathcal{P}_{\mathrm{B_q}\to\mathrm{B_q}}(t) + \mathcal{P}_{\mathrm{B_q}\to\overline{\mathrm{B}}_\mathrm{q}}(t)\right] = 1 \ . \tag{2.17}$$

Approximations

Two approximations can be taken to simplify the formulae shown above which still give expressions precise enough for the purpose of the experimental testing.

1. **CP conservation in mixing**
   CP conservation implies (Eqs. 2.12, 2.13, 2.14)

$$\begin{array}{rclcrcl} \langle \mathrm{B_q}|M|\overline{\mathrm{B}}_\mathrm{q}\rangle & = & \langle \overline{\mathrm{B}}_\mathrm{q}|M|\mathrm{B_q}\rangle & \text{and therefore} & M_{12} & = & M_{12}^* \,, \\ \langle \mathrm{B_q}|\Gamma|\overline{\mathrm{B}}_\mathrm{q}\rangle & = & \langle \overline{\mathrm{B}}_\mathrm{q}|\Gamma|\mathrm{B_q}\rangle & \text{and therefore} & \Gamma_{12} & = & \Gamma_{12}^* \,, \end{array} \tag{2.18}$$

   so the matrix elements M_{12} and Γ_{12} are real. Furthermore, it can also be shown that

$$p = q \,, \qquad \Delta M_{\mathrm{B_q}} = 2M_{12} \,, \qquad \text{and} \qquad \Delta\Gamma_{\mathrm{B_q}} = 2\Gamma_{12} \,. \tag{2.19}$$

2. **Negligible lifetime difference**
   In the approximation of negligible CP violation in the mixing process, the ratio between the width and the mass differences, $\Delta\Gamma_{\mathrm{B_q}}/\Delta m_{\mathrm{B_q}}$, is equal to Γ_{12}/M_{12}. In the Standard Model, this quantity is expected to be of the order of the ratio $m_\mathrm{b}^2/m_\mathrm{t}^2$. It is therefore a small quantity, and the same for the two neutral B meson systems. It can be calculated with lattice QCD, with typical results $\sim 5 \times 10^{-3}$. The width difference, $\Delta\Gamma_{\mathrm{B_q}}$, is caused by the existence of final states to which both the $\mathrm{B_q}$ and $\overline{\mathrm{B}}_\mathrm{q}$ mesons can decay. Such decays involve $\mathrm{b} \to \mathrm{c\bar{c}q}$ quark-level transitions, which are Cabbibo-suppressed if q = d and Cabbibo-allowed if q = s. Standard Model computations thus predict $\Delta\Gamma/\Gamma$ to be very small for the $\mathrm{B_d}$ system (below 1% [13]) and somewhat higher for the $\mathrm{B_s}$ system, up to $\sim 20\%$. Experimental results show $\Delta\Gamma_{\mathrm{B_s}}/\Gamma_{\mathrm{B_s}} = 0.16^{+0.08}_{-0.09}$ [29]. There is no experimental constraint for $\Delta\Gamma_{\mathrm{B_d}}/\Gamma_{\mathrm{B_d}}$. Neglecting $\Delta\Gamma_{\mathrm{B_q}}$ in Eq. 2.15 means replacing the hyperbolic cosine by unity. Even for the $\mathrm{B_s}$ system, this approximation is reasonable and is used along this thesis.

With the approximations above, the time dependent probabilities of Eq. 2.15 become

$$\begin{aligned} \mathcal{P}_{\mathrm{B_q}\to\mathrm{B_q}}(t) &= \Gamma_{\mathrm{B_q}} \frac{e^{-\Gamma_{\mathrm{B_q}} t}}{2} \left[1 + \cos(\Delta m_{\mathrm{B_q}} t)\right] \,, \\ \mathcal{P}_{\mathrm{B_q}\to\overline{\mathrm{B}}_\mathrm{q}}(t) &= \Gamma_{\mathrm{B_q}} \frac{e^{-\Gamma_{\mathrm{B_q}} t}}{2} \left[1 - \cos(\Delta m_{\mathrm{B_q}} t)\right] \,, \end{aligned} \tag{2.20}$$

which are the expressions used along this thesis. The probability density function for a $\mathrm{B_q}$ ($\overline{\mathrm{B}}_\mathrm{q}$) is therefore the exponential decay modulated by a cosine of frequency $\Delta m_{\mathrm{B_q}}$, the mass difference between the two mass eigenstates. In Eq. 2.20, the convention $\hbar = 1$, $c = 1$ is taken, and consequently $\Delta m_{\mathrm{B_q}}$ is expressed in units of inverse time. The conversion factor between ps^{-1} and eV/c^2 is $1\,\mathrm{ps}^{-1} \simeq 6.6 \times 10^{-4}\,\mathrm{eV}/c^2$. The mass difference $\Delta m_{\mathrm{B_q}}$ is always expressed in ps^{-1} in this thesis.

2.2.1 B mixing and the CKM matrix

In the framework of the Standard Model, neutral B-meson mixing is only possible at the one (or more) loop level, via Feynman *box-diagrams* involving up-type quark exchange. Two such interfering diagrams, shown in Fig. 2.2, are involved.

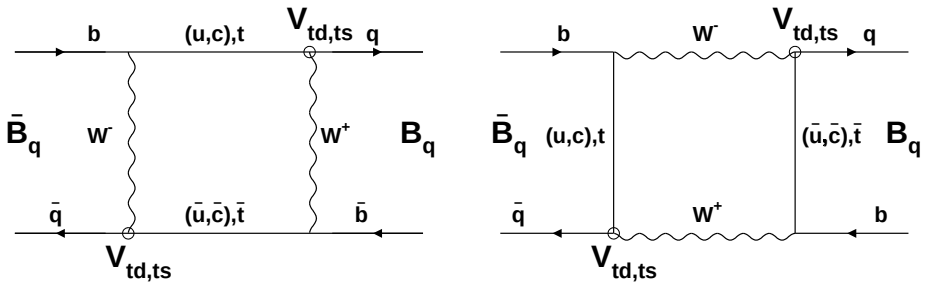

Figure 2.2: Feynman diagrams for B mixing.

The off-diagonal term of the mass matrix, $M_{12}^{(q)}$, can be determined from the computation of these diagrams. And therefore, recalling Eq. 2.19, the mass difference, $\Delta m_{\mathrm{B_q}}$ between the two neutral B-meson mass eigenstates is obtained as well. These diagrams are proportional to the square of the CKM matrix elements involved, $|V_{\mathrm{ib}}V_{\mathrm{iq}}|^2$, where $i = \mathrm{u, c, t}$, and to the squared mass of the up-type quark in the internal loop. Using the parametrization of Eq. 2.5, all three terms involving CKM matrix elements are found to be of similar size for $\mathrm{q} = \mathrm{d}$. In the case of $\mathrm{q} = \mathrm{s}$, the differences between the CKM terms are not big enough to be relevant either. The quark masses are highly hierarchical, *i.e.*, $m_{\mathrm{t}} \gg m_{\mathrm{c}} \gg m_{\mathrm{u}}$. As a consequence, the computation of these diagrams is completely dominated by the top-quark exchange. The basic equation to be solved is

$$2m_{\mathrm{B_q}}|M_{12}^{(q)}| = |\langle \overline{\mathrm{B}}_{\mathrm{q}}|\mathcal{H}_{\mathrm{eff}}^{(\Delta B=2)}|\mathrm{B_q}\rangle| \ , \tag{2.21}$$

where $\mathcal{H}_{\mathrm{eff}}^{(\Delta B=2)}$, relevant for scales $\mu_{\mathrm{b}} = \mathcal{O}(m_{\mathrm{b}})$, is the effective Hamiltonian for a change in two units of the beauty quantum number. This Hamiltonian can be expressed as [30]

$$\mathcal{H}_{\mathrm{eff}}^{(\Delta B=2)} = \frac{G_{\mathrm{F}}^2}{16\pi^2} M_{\mathrm{W}}^2 |V_{\mathrm{tq}}V_{\mathrm{tb}}^*|^2 S(x_{\mathrm{t}}) C(\mu_{\mathrm{b}}) \hat{Q}(\mu_{\mathrm{b}})^{(\Delta B=2)} + h.c. \tag{2.22}$$

Where G_{F} is the Fermi constant, M_{W} is the W-boson mass, $x_{\mathrm{t}} = m_{\mathrm{t}}^2/M_{\mathrm{W}}^2$, $S(x)$, the Inami-Lim function reads [31]

$$S(x) = x\left[\frac{1}{4} + \frac{9}{4}\frac{1}{(1-x)} - \frac{3}{2}\frac{1}{(1-x)^2} - \frac{3}{2}\frac{x^2 \ln x}{(1-x)^3}\right] \ , \tag{2.23}$$

and can be approximated with $0.784x^{0.76}$. The function $C(\mu_{\mathrm{b}})$ and the operator $\hat{Q}(\mu_{\mathrm{b}})$ include all the QCD corrections, *i.e.*, the computation of internal quark and gluon loops. The study of B-mixing from a theoretical point of view involves strong interactions at two distinct energy scales. The electroweak scale, of order M_{W}, and the quark confinement scale, of order Λ_{QCD}. In the first case, perturbative QCD techniques can be applied, to obtain $C(\mu_{\mathrm{b}})$. Non-perturbative QCD is used to get an estimate of the matrix element $|\langle \overline{\mathrm{B}}_{\mathrm{q}}|\hat{Q}(\mu_{\mathrm{b}})|\mathrm{B_q}\rangle|$ [14, 30]. The dependence on the energy scale μ_{b} disappears when both QCD terms are computed.

The expression for $\Delta m_{\mathrm{B_q}}$ is obtained from Eq. 2.21 and Eq. 2.22, and can be written as [14]

$$\Delta m_{\mathrm{B_q}} = \frac{G_{\mathrm{F}}^2}{6\pi^2} M_{\mathrm{W}}^2 m_{\mathrm{B_q}} |V_{\mathrm{tq}}V_{\mathrm{tb}}^*|^2 S(x_{\mathrm{t}}) \eta_B B_{\mathrm{B_q}} F_{\mathrm{B_q}}^2 \ . \tag{2.24}$$

The factor η_B stands for the perturbative QCD corrections. It has been evaluated to be $\eta_B = 0.55 \pm 0.01$ and it is the same for the B_d and B_s systems: short distance QCD calculations do not depend on the light quark in the B-meson [32]. The uncertainty on its estimate comes from the uncertainty on the top-quark mass and from the fact that the calculation is done only up to the next-to-leading order. The term B_{B_q} is called the *Bag*-parameter, it measures deviations from the vacuum saturation approximation [33] which consists in ignoring gluon exchanges between B_q and $\overline{B}_q$ quarks. This approximation would correspond to $B_{B_q} = 1$. The last term, F_{B_q}, is the leptonic decay constant defined by the expression [34]

$$\langle 0|\bar{b}\gamma_\mu\gamma_5 q|B_q(p)\rangle = ip_\mu F_{B_q} \ . \tag{2.25}$$

Both these non-perturbative QCD quantities are computed using lattice QCD. A vast literature on these calculations is available with slightly different results depending on the approach followed. The "average" values quoted in Ref. [35], for instance, for the decay constants are

$$F_{B_d} = (200 \pm 40)\,\text{MeV}\,, \quad F_{B_s} = (230 \pm 40)\,\text{MeV}\,, \quad \text{and} \quad \frac{F_{B_s}}{F_{B_d}} = 1.15 \pm 0.07\ . \tag{2.26}$$

In the same reference, the value given for the *Bag*-parameter does not distinguish q = d from q = s (although it is done in other references like [34]) because all methods used for the *Bag*-parameter estimate predict B_{B_s}/B_{B_d} very close to one. The values quoted in Ref. [35] are

$$B_{B_q}(m_b) = 0.91 \pm 0.06\,, \quad \text{and} \quad \frac{B_{B_s}}{B_{B_d}} = 1.00 \pm 0.03\ , \tag{2.27}$$

The mass difference Δm_{B_q} is an experimentally accessible parameter. Indeed, from Eq. 2.20 it can be interpreted as the B_q-meson oscillation frequency. As seen in Eq. 2.24, Δm_d and Δm_s are directly related to CKM matrix elements. Their measurement, modulo theoretical calculations, would allow poorly known terms of the CKM matrix to be determined and the matrix properties to be tested. However, the theoretical uncertainties are, for the time being, too large to give relevant constraints on the CKM matrix elements involved. Nevertheless, if the ratio of the mass differences of the two neutral B-meson systems is taken, the dependences on m_t and some QCD corrections cancel out, and the following expression is obtained

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}}\xi^2\left|\frac{V_{ts}}{V_{td}}\right|^2, \tag{2.28}$$

where all the theoretical uncertainties are contained in the parameter ξ,

$$\xi \equiv \frac{F_{B_s}\sqrt{B_{B_s}}}{F_{B_d}\sqrt{B_{B_d}}}\ . \tag{2.29}$$

From Eq. 2.26, the decay constants have $\sim 20\%$ uncertainty, whereas their ratio, F_{B_s}/F_{B_d}, is better understood from a theoretical point of view, with only 7% uncertainty. As a consequence, the theoretical uncertainties on the parameter ξ are also under control, with a present best value of $\xi = 1.15 \pm 0.07$ [35]. Therefore, if both Δm_d and Δm_s were measured, the quantity $|V_{ts}|/|V_{td}|$ could be determined with small theoretical uncertainties from their ratio. Getting a measurement of $|V_{ts}|/|V_{td}|$ is one of the main motivations for pursuing a measurement of the Δm_s parameter.

2.2.2 B mixing and CP violation

A measurement of both Δm_{d} and Δm_{s} would have an important impact on the understanding of CP violation in the Standard Model. From Eq. 2.7 and Eq. 2.28, one of the sides of the unitarity triangle (Fig. 2.1) can be expressed as

$$\overline{\mathrm{BA}} = \frac{1}{\lambda}\left|\frac{V_{\mathrm{td}}}{V_{\mathrm{cb}}}\right| = \frac{1}{\lambda}\left|\frac{V_{\mathrm{td}}}{V_{\mathrm{ts}}}\right|\left|\frac{V_{\mathrm{ts}}}{V_{\mathrm{cb}}}\right| = \frac{1}{\lambda}\sqrt{\frac{m_{\mathrm{B_s}}}{m_{\mathrm{B_d}}}\frac{\Delta m_{\mathrm{d}}}{\Delta m_{\mathrm{s}}}}\,\xi\left|\frac{V_{\mathrm{ts}}}{V_{\mathrm{cb}}}\right| . \tag{2.30}$$

The CKM matrix elements $|V_{\mathrm{ts}}|$ and $|V_{\mathrm{cb}}|$ are, to a very good approximation, $\mathcal{O}(\lambda^4)$, equal. Therefore, from Eq. 2.30, the measurement of both neutral B-meson systems oscillation frequencies is directly related to one of the sides of the unitarity triangle. Part of the CP violation description in the Standard Model can thus be obtained from the experimental results on Δm_{d} and Δm_{s}.

The values of the parameters $\bar{\rho}$ and $\bar{\eta}$, from the unitarity triangle, can be extracted from a global fit to a particular set of experimental measurements and theoretical computations [36]. The list of inputs to this global fit is: Δm_{d}, Δm_{s}, $|V_{\mathrm{ub}}|$, $|V_{\mathrm{cb}}|$, m_{t}, $|\varepsilon_K|$ (from CP violation measurements in the kaon system), B_{K} (the *Bag*-parameter for the kaon system), $F_{\mathrm{B_d}}\sqrt{B_{\mathrm{B_d}}}$, and ξ. The results presented in Ref. [36]:

$$\bar{\rho} = 0.224 \pm 0.038 \qquad \bar{\eta} = 0.317 \pm 0.040 \qquad \sin(2\beta) = 0.698 \pm 0.066 \tag{2.31}$$

are compatible with the first direct measurements of the angle β of the unitarity triangle at the asymmetric B factories. An average of the preliminary results reads, $\sin(2\beta) = 0.48 \pm 0.16$ [37].

If the experimental input on $\mathrm{B_s}$ oscillations is removed from the global CKM matrix fit described above, constraints on the still unmeasured oscillation frequency for $\mathrm{B_s}$ mesons can be obtained. One of the many examples of such an analysis is presented in Ref. [38]. The probability density function for Δm_{s} obtained in that reference is shown in Fig. 2.3.

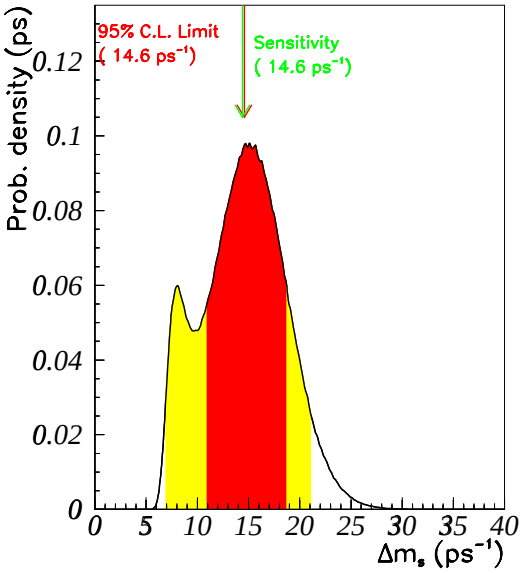

Figure 2.3: Probability density function for Δm_{s}.

Figure 2.3 shows that constraints on the CKM matrix elements favour a value of Δm_s around $15\,\mathrm{ps}^{-1}$ and values above $25\,\mathrm{ps}^{-1}$ would be difficult to accommodate in the absence of new physics.

2.2.3 State-of-the-art

Evidence for neutral B-meson mixing was first observed by the UA1 Collaboration at CERN, Ref. [39]. An excess of like-sign dimuon events was observed. The predicted ratio for like-sign to unlike-sign muon pairs was 0.26 ± 0.03, coming from events where one muon arises from a B-decay and the other from a cascade charm decay. However, the measured ratio was $0.42 \pm 0.07 \pm 0.03$, which constituted the first experimental evidence for B^0-mixing. At that time, b-production fractions were poorly known and it was therefore not possible to disentangle the contribution of B_d and of B_s mixing in the result obtained. The theoretical computations indicated already that the oscillation frequency for B_s was higher than that of B_d. Moreover, the ARGUS and CLEO [40, 41] Collaborations at Desy and CESR had, at that time, put limits on r_d, $r_d < 0.12$ at 90% C.L. and $r_d < 0.24$ at 90% C.L. respectively (r_d is related with χ_d –the mixing parameter, Section 3.1– through: $r_d = \chi_d/(1-\chi_d)$). Results from MARKII [42] were consistent with a very low value of χ_d as well. All these results together made the UA1 Collaboration favour an interpretation of their result in terms of mainly B_s mixing. However, thanks to the much more precise information presently available, it is now known that the effect measured at UA1 was dominated by B_d mixing. Few months after the UA1 publication, ARGUS [43] reported on the first direct observation of $B_d - \overline{B}_d$ mixing at the $\Upsilon(4S)$ resonance. The mass of the $\Upsilon(4S)$ is not high enough to create a $B_s - \overline{B}_s$ pair. Therefore, there is no uncertainty on the neutral B-meson composition, and no possibility of misinterpretation. The first measurement of the B_d mixing parameter at the $\Upsilon(4S)$ was obtained by the ARGUS Collaboration in 1987[43], $r_d = 0.21 \pm 0.08$, followed by a slightly more precise result from the CLEO Collaboration in 1989, $r_d = 0.19 \pm 0.06 \pm 0.06$ [44]. A direct measurement of B_d oscillations had to wait for LEP data and came some years later, in 1994, from the ALEPH Collaboration, $\Delta m_d = (0.50^{+0.07\ +0.11}_{-0.06\ -0.10})\,\mathrm{ps}^{-1}$ [45].

$B_d - \overline{B}_d$ Mixing

The measurement of B_d oscillations has today become a precision measurement. The four LEP Collaborations, SLD and CDF have published results on the subject giving an average value of $\Delta m_d = 0.486 \pm 0.015\,\mathrm{ps}^{-1}$ [48]. The CLEO and ARGUS Collaborations have performed measurements of the mixing parameter χ_d. Their results can be combined with the measurements of Δm_d with a non-trivial procedure explained in Ref. [8]. The present world combined average is $\Delta m_d = 0.487 \pm 0.014\,\mathrm{ps}^{-1}$ [8] (including results from LEP, SLD, CDF, ARGUS and CLEO).

The asymmetric B-factories have been running for more than a year now and the first results presented include measurements on B_d oscillations. These results are still preliminary but show that, with some more time, the precision of the world average for the B_d mixing result will further improve and will soon be dominated by the results from the Babar and Belle experiments taking data on the asymmetric B-factories.

$B_s - \overline{B}_s$ Mixing

Studies on B_s oscillations have only been possible until now at the Z resonance (LEP and SLC) and with lower sensitivity at the Tevatron. The B_s oscillation frequency is expected to be much larger than that of B_d mesons, as it can be deduced from Eq. 2.28 and Fig. 2.3. Up to now, B_s oscillations have not been resolved and only lower limits have been set. The latest world average lower limit which includes some of the results presented in this thesis is $\Delta m_s > 15\,\mathrm{ps}^{-1}$ at 95% C.L. [50].

Some of the analyses both at SLD and at the LEP Collaborations are still expected to improve slightly. It is not excluded that the final combination of the Z peak data analyses be able to resolve B_s oscillations. If with the Z peak data only a lower limit is set, the first measurement will most probably come from the experiments at the Tevatron collider in the next years [51].

The width difference $\Delta\Gamma_s$ could be large enough to be measured before the mass difference Δm_s [52]. The theoretical estimate of $\Delta\Gamma_s/\Delta m_s$ suffers from no CKM uncertainty, it has only QCD contributions which can be computed on the lattice. Therefore, a reliable estimate of Δm_s could be inferred a the measured value of $\Delta\Gamma_s$. The present measurement of $\Delta\Gamma_s/\Gamma_s$ is however not accurate [29], and only provides a mild constraint on the allowed region for Δm_s: $\Delta m_s = 29\,^{+16}_{-21}\,\mathrm{ps}^{-1}$ [50].

2.3 B mixing beyond the Standard Model

In general, most of the extensions of the Standard Model which are being studied do not have sizeable implications on processes at tree level, and therefore have negligible contributions to B decays. The B-mixing phenomenon is only possible at the one (or more) loop level. Contributions from physics beyond the Standard Model can enter these loops, and therefore B-mixing can be a test bench for new theoretical models.

It is not possible in the scope of this thesis to give a comprehensive review of the possible extensions of the Standard Model which have implications on mixing parameters. A selection of examples found in the literature is given. A more extensive discussion is presented, for example, in Ref. [53].

One of the models which is being presently studied is called *extra-dimensions*. In this model the ordinary four-dimensional Standard Model arises as a low energy effective theory of models defined in five or more dimensions. One of the simplest generalizations has one single extra-dimension. The model contains fermions which live in four dimensions, and gauge bosons and one scalar doublet propagating in five dimensions [54]. This extended model allows phenomenological constraints derived from one loop processes to be extracted, and in particular some predictions related with B-mixing can be performed. The correction induced by the extra-dimension to the Inami-Lim function $S(x_t)$ has been estimated giving the range $1 < S(x_t) < 10$, while the Standard Model value is $S(x_t) \sim 2.6$. Experimentally, the value of $S(x_t)$ has still some uncertainties, which makes impossible yet to extract any bound on extra-dimension models from the B-mixing data. However, if in the future a value of $S(x_t)$ larger than that of the Standard Model is found, extra-dimensions could easily accommodate it.

Probably the most popular framework for physics beyond the Standard Model is supersymmetry. Many contributions to the $B - \overline{B}$ mixing are possible. However the parameter space of the different models is vast and firm constraints are difficult to establish. The effect of supersymmetry on $B - \overline{B}$ mixing could be very significant: for certain values of the parameters, for example, the total supersymmetric contribution to $B_d - \overline{B}_d$ mixing can be twice as large as that of the Standard Model [53].

Finally, there are theories with a real (orthogonal) CKM matrix, which therefore explain CP violation in some other means, *superweak theories* [55]. These theories require that Standard Model fits to the unitarity triangle gives $\eta = 0$. Experimental limits on B_s oscillations have now almost excluded a negative value of ρ, and in general do not favour the region $\eta = 0$ with positive ρ either. In the framework of the pure superweak theories, Δm_s is expected to be larger than in the Standard Model: between $10\,\mathrm{ps}^{-1}$ and $32\,\mathrm{ps}^{-1}$ at 95% C.L., which is still compatible with the present experimental results.

Chapter 3

Experimental Method

The B mixing can be studied with two distinct methods. In a first method, only the total rate of events in which one B changes flavour between the time of production and that of decay is determined independently of the time at which the decay occurs. The particle-antiparticle oscillation is therefore integrated over time. In the second method, instead, the time dependence shown in Chapter 2, Section 2.2, is measured, and the actual oscillation frequency is extracted from that distribution. These methods are known as *time-integrated mixing* and *oscillation* measurements. Both methods were successfully used for B_d mesons studies [45, 46, 47, 48, 49]. While the slow oscillation frequency of B_d mesons allows information to be extracted from the time-integrated mixing measurements, the situation is different in the case of B_s mesons. As a consequence of their rapid oscillation frequency with respect to the decay time, half of the B_s mesons are found, when they decay, in a flavour different from that at production. The time-integrated measurement is therefore insensitive to the oscillation frequency.

To start this Chapter, the basic features of the time-integrated and oscillation measurement methods are given, and general techniques used in B-mixing analyses are explained. The main experimental issues are reviewed, with special emphasis on the factors which limit the sensitivity. The most important differences between the experimental studies on B_d and B_s oscillations are explained. The impact of the experimental resolution on the reconstructed B-meson proper time (*i.e.*, the B meson decay time in its rest frame, see Section 3.1.1) is discussed in detail, for instance. In the case of B_s mesons, contrary to B_d mesons, the proper time resolution is crucial. Although the experimental description is centred on relevant issues for data taken at the Z peak (at LEP and SLC), a brief description of the B-mixing measurements performed at CLEO or CDF is also given.

In the case of oscillation studies, a fit to the expected proper time distribution is performed. The fitting procedure is explained in Section 3.2. The fit for B_s oscillations is most often performed with the *amplitude method* [9, 56]. In this method, an amplitude $\mathcal{A}$ is introduced in front of the oscillating term in the probability density function for B_s mesons, (Eq. 2.20), and measured as a function of the oscillation frequency. The method is explained in detail in Section 3.3, along with a complete discussion of the expected dependence on the frequency of the measured amplitude and its uncertainty. The interpretation of the results in terms of the probability of being due to a statistical fluctuation is given.

3.1 Basic features of a B mixing measurement

Time-integrated measurement

The simplest experimental approach, and first used ever, for a measurement of the $B^0 - \overline{B}^0$ mixing consists in the estimation of the B^0 mesons fraction with a different flavour at the times of production and decay. If this fraction is zero, it means that there is no mixing in the system studied. If it is 50% it means that the oscillation is rapid enough to saturate the integrated method, rendering it insensitive to the actual oscillation frequency. Values in between can be turned into a measurement of the oscillation frequency (or equivalently, of the mass difference between the two mass eigenstates, Δm_{B_q}) as follows. The mixing parameter χ_q is defined as the average probability of a B_q^0 meson to be found in a $\overline{B}_q^0$ state at decay:

$$\chi_q = \frac{\int_0^\infty P_{B\to\overline{B}} dt}{\int_0^\infty P_{B\to B} dt + \int_0^\infty P_{B\to\overline{B}} dt} \ . \tag{3.1}$$

With the approximation of $\Delta\Gamma_q = 0$ and negligible CP violation in the mixing, as discussed in Section 2.2, χ_q reads

$$\chi_q \simeq \frac{(\Delta m_{B_q}/\Gamma_{B_q})^2}{2[1 + (\Delta m_{B_q}/\Gamma_{B_q})^2]} \ . \tag{3.2}$$

The time-integrated method has been successfully used for B_d mixing studies done with data taken at accelerators running at the energy of the $\Upsilon(4S)$ resonance [49]. This energy is just enough to allow the $\Upsilon(4S)$ to decay in a coherent $B_d^0 - \overline{B}_d^0$ or B^+-B^- pair, but it is too small for a $B_s^0 - \overline{B}_s^0$ pair to be produced. The first observation of B_d^0 mixing, by the ARGUS experiment, was obtained from the complete reconstruction of a single $B_d^0 - \overline{B}_d^0$ event in which the $\overline{B}_d^0$ had oscillated [43]: two B_d^0 were reconstructed, with the following decay chains: $B_d^0 \to D^{*-}\mu^+\nu_\mu$, with $D^{*-} \to \pi^-\overline{D^0}$, and $\overline{D^0} \to K^+\pi^-$, and the other, $B_d^0 \to D^{*-}\mu^+\nu_\mu$, with $D^{*-} \to \pi^0 D^-$, and $D^- \to K^+\pi^-\pi^-$. A display of this event is shown in Fig. 3.1. The presence of two B_d^0 in the final state is an unambiguous proof of the oscillation. However, complete reconstruction selections are very inefficient and therefore very few such fully-reconstructed events are available, and most of the sensitivity to χ_d for the experiments running at the $\Upsilon(4S)$ resonance, comes from events partially reconstructed. In most of the cases, two semileptonic decays are looked for, with like-sign leptons to identify events with a B_d^0 meson having oscillated.

Time-integrated mixing studies with similar techniques were performed in $p\overline{p}$ colliders (UA1 and CDF) [46] and at the Z resonance (SLD and LEP experiments) [47]. In both cases, however, B_d mesons are not the only contribution to mixing, B_s mesons also contribute to the integrated mixing rate. As a result, χ_d is not directly measurable, and the measured quantity $\overline{\chi}$ is an average of the B_d and B_s mesons integrated-mixing. The average mixing parameter is related to χ_d through,

$$\overline{\chi} = f_d\chi_d + f_s\chi_s \ , \tag{3.3}$$

where f_d and f_s are the fractions of $B_{d,s}$ in $b\overline{b}$ decays, and $\chi_s = 0.5$.

As already stated, B_s oscillations are too rapid to be studied in a time-integrated manner: only analyses which exploit the B-meson proper time information are sensitive to Δm_s.

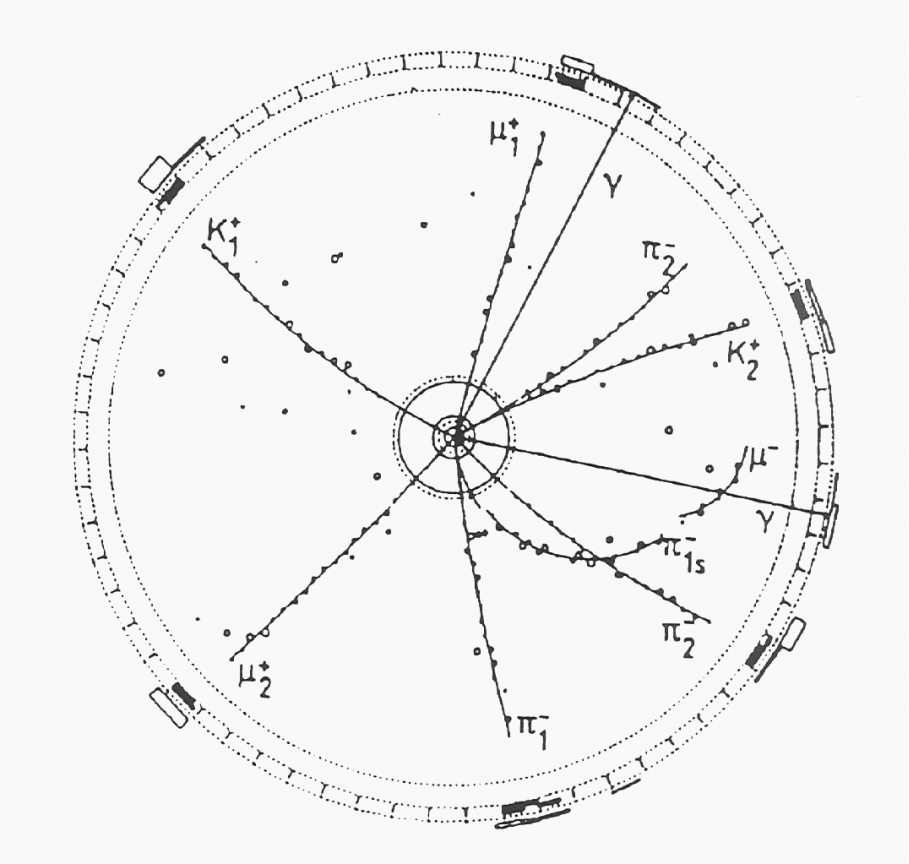

Figure 3.1: The gold plated event, first evidence for B_d oscillations from the ARGUS Collaboration.

Oscillation measurement

Oscillation measurements aim at reconstructing the proper time distribution of B^0 meson decays shown in Eq. 2.20. To do so, more experimental information on the decay than in the case of time-integrated measurements is needed.

First, the proper time of the B^0 decay studied needs to be determined, although it is not a quantity directly measurable. Indeed, it can only be inferred from two independent measurements, *i*) the decay length l, *i.e.*, the distance of flight before decay; and *ii*) the momentum p of the B^0 meson; both measured in the laboratory rest frame. These two measurements are described in Section 3.1.1.

It is also necessary to determine if the meson oscillated or not. To this end, the b-quark flavour (b or $\overline{\mathrm{b}}$) of the B^0 meson has to be determined both at production (initial state) and decay (final state); if identical, the event is labelled as unmixed, and as mixed otherwise. In general, the analysis event selection is strongly determined by the technique used for the final state flavour determination. Details are given in Section 3.1.2.

The proper time and the mixed/unmixed label are the two main pieces of information needed for an oscillation measurement, and would actually be the only needed in a pure

sample of B_s (or B_d) mesons. However, such pure samples cannot be isolated at the Z-boson pole or at the $p\overline{p}$ colliders. Further selection must be performed to isolate as much as possible the kind of events desired. Some of the most common approaches for signal enrichment and background evaluation are given in Section 3.1.3.

Although oscillation studies are done for both B_d an B_s mesons systems, the oscillation frequency Δm_d and Δm_s are of different enough magnitude to render the experimental key issues in both cases quite different. To illustrate this point, the proper time distributions expected in both cases are presented in Fig. 3.2. In the case of B_d oscillations, the frequency,

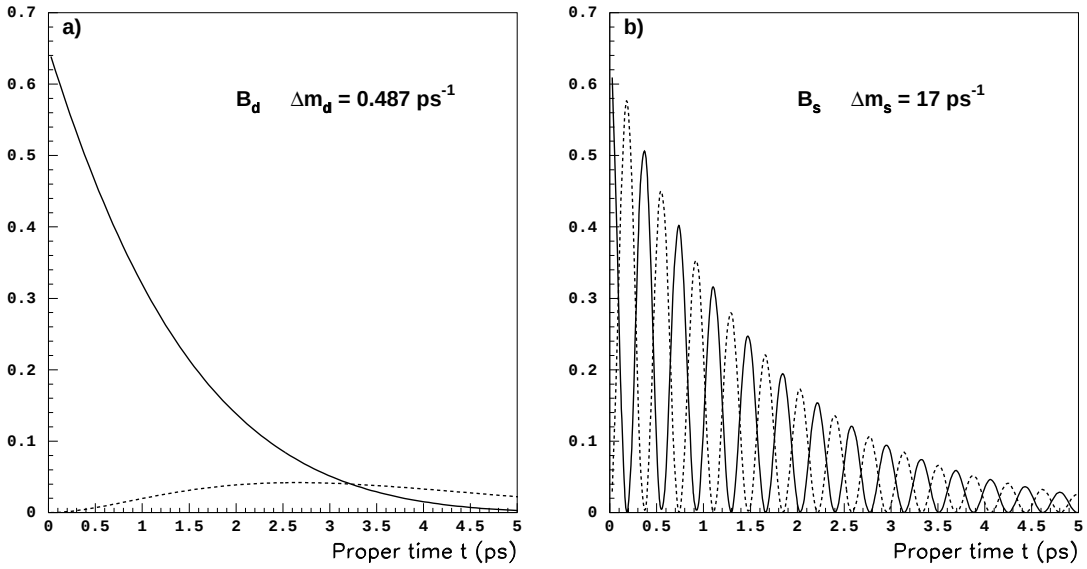

Figure 3.2: Proper time distribution for mixed/unmixed B_d and B_s mesons. Full curve holds for unmixed and dotted for mixed events.

already measured with good accuracy, is small. As it is shown Fig. 3.2a, the proper time distribution expected for mixed and unmixed events can be disentangled even with a moderate proper time resolution (typical values of the proper time resolution achieved at LEP are of the order of 0.3 ps, already excellent for that purpose). For B_s mesons, however, the rapid oscillation frequency (for illustration $\Delta m_s = 17\,\mathrm{ps}^{-1}$ is taken) makes the analysis complicated and renders the proper time resolution crucial. The distributions shown in Fig. 3.2 refer to the expectations in the case of perfect mixed/unmixed labelling. When the experimental uncertainty on this labelling is taken into account, the difference between mixed and unmixed proper time distributions is diluted. A B_s oscillation frequency measurement, never achieved yet, requires excellent proper time resolution as well as very good performance in flavour tagging.

All the ingredients of such an oscillation analysis are described in detail in the following Sections. The experimental techniques described apply to B_s oscillations analyses, but in most of the cases they are similar to those used for B_d studies.

3.1.1 Proper time measurement

The determination of the B meson proper time, *i.e.*, the B meson lifetime as evaluated in its own rest frame, requires the measurement of its decay length and its momentum in the laboratory rest frame. The experimental technique often depends on the event selection. Some of the general aspects are given in this Section. More details, closely related to the main analysis described in this thesis, can be found in Chapter 6.

The decay length of the B meson is the three-dimensional distance between the point where the primary interaction takes place (*i.e.*, the e^+e^- annihilation, in the case of LEP), and the point where the flying meson decays. The main interaction point determination is not specific to a B oscillation analysis and depends on the experimental environment where the data are taken. The method used by the ALEPH experiment is explained in Section 5.4. The secondary vertex, *i.e.*, the point where the B meson decays, is determined with different methods depending on the event selection. The best resolutions are achieved for exclusive decay reconstructions. In this configuration all the secondary particles are identified and their common vertex position can be determined with high precision. For more inclusive selections, topological techniques lead to much worse resolutions. In some cases, for semileptonic decays for instance, the decay vertex is obtained from hybrid techniques, in which a fraction of the decay particles are identified.

The B meson momentum measurement depends on the event selection as well. If the decay is exclusively reconstructed, it is determined as the sum of the momenta of all decaying particles and the resolution is excellent by construction. The momentum of the jet in which the secondary vertex is found is taken as the meson momentum for an inclusive event selection. In the case of semileptonic decays, a fraction of the B meson momentum is carried by a neutrino which escapes the experiment without detection. The neutrino momentum is estimated from the requirement of energy and momentum conservation in the event (the quality of this technique depends strongly on the experimental environment where the data are taken, it is successfully used at the ALEPH experiment, as explained in Section 6.7.1). The B momentum is then estimated as the sum of the lepton and the charm candidate momenta, complemented with the neutrino momentum estimate.

Even though the exclusive reconstruction of decays has the advantage of an almost perfect resolution, it implies a very tight selection. The resulting data samples available are therefore statistically poor. In contrast, the event selection for an ultimate B oscillation measurement has to be a compromise between exclusiveness to obtain good resolution and inclusiveness to gain from statistics. The analysis described in Chapter 6 is based on an inclusive semileptonic selection, which constitutes the best compromise in the LEP environment. For the SLD experiment the situation is different. The data sample is ten times smaller than in a LEP experiment but the intrinsic resolution of the tracking is better by a factor of 3. That makes SLD data more adequate for analyses based on a fully inclusive selection. At the CDF experiment, being on a $p\overline{p}$ collider, inclusive or semi-inclusive techniques are not viable, only exclusive (or close to exclusive) methods are possible.

The proper time t is determined from the decay length l and the momentum p of the B meson as

$$t = \frac{l\,m}{p\,c}\ , \tag{3.4}$$

where m is the B meson mass, and c is the speed of light. The proper time resolution can be therefore expressed, by differentiation, as a function of the decay length and momentum resolutions as

$$\sigma_t = \frac{m}{p\,c}\,\sigma_l \oplus t\,\frac{\sigma_p}{p}\ , \tag{3.5}$$

where σ_l is the decay length resolution, and σ_p/p the relative momentum resolution. The second term of Eq. 3.5 is proportional to the measured proper time itself. In the case of B_s mesons, the period of the oscillation is much shorter than the lifetime, therefore the sensitivity comes mostly from short lived mesons, for which the contribution of the momentum resolution vanishes. The sensitivity of an inclusive analysis is then limited by the resolution on the decay length. In the case of B_d oscillations, a significant mixing rate is observed only at proper times larger than a lifetime ($t > \tau_{B_d}$). Therefore the momentum resolution term dominates the proper time reconstruction uncertainty, which is anyway not a crucial issue for those analyses.

3.1.2 Flavour tagging

The b-quark flavour of the B meson needs to be determined at production and decay and compared to each other to establish the unmixed/mixed label of each data event.

Three different techniques have been used to date for the final state (decay) flavour determination, depending on the event selection performed.

- In the case of exclusive reconstruction of flavour eigenstates there is no ambiguity. The decay chain establishes if the parent meson was a B^0 or a $\overline{B}^0$. Examples of exclusive reconstruction oscillation analyses are found in Refs. [57, 58] and in Section 7.2.

- For semileptonic decays, the electric charge of the lepton is used as a tag. Figure 3.3 shows the Feynman diagrams for semileptonic decays of both B^0 and $\overline{B}^0$. A lepton

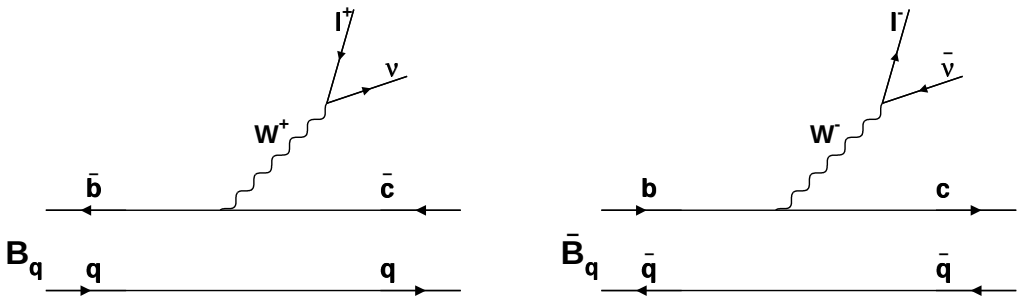

Figure 3.3: Semileptonic B^0 meson decay diagrams.

with negative electric charge originates from $\overline{B}^0$ and a positively charged lepton from a B^0. However, the presence of a lepton in the final state does not guarantee a direct semileptonic decay. Cascade decays with $b \to c \to \ell$ are also frequent, and have the opposite lepton-quark electric charge sign correlation (the corresponding diagrams are

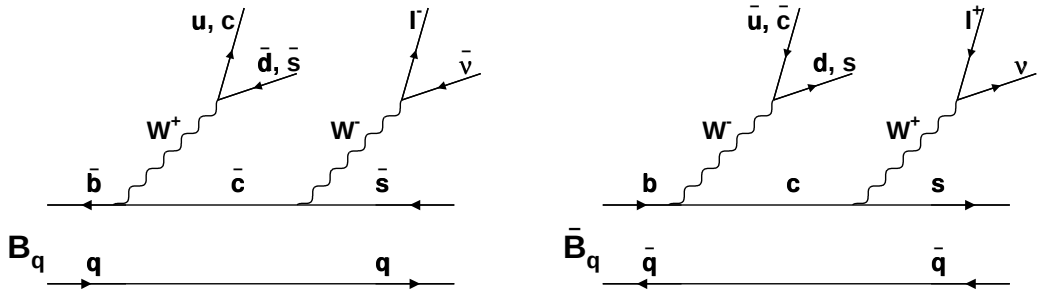

Figure 3.4: Cascade B^0 meson decay diagrams.

shown in Fig. 3.4). These cascade decays need to be rejected from the final data sample. Additional cascades decays with $b \to \overline{c} \to \ell$ are also present (Fig. 3.5). The charge correlation between the lepton and the b quark is the same as in direct semileptonic decays, therefore no specific rejection for these decays is required. Examples of oscillation

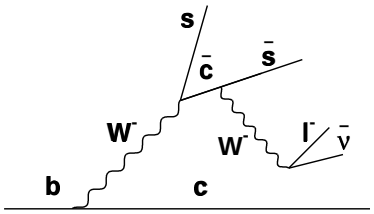

Figure 3.5: Cascade $b \to \bar{c} \to \ell$ decay diagram.

analyses using semileptonic selections are found in Refs.[58, 59].

- In the case of an inclusive selection, no decay particle is identified. A different method has to be used. An inclusive variable sensitive to the charge flow between the secondary and the tertiary vertices is used to statistically separate B^0 from $\overline{B}^0$ decays. In the case of SLD, their excellent tracking resolution is exploited which allows the tertiary vertex to be inclusively reconstructed. A dipole, $|Q_D - Q_B| \times |\vec{V}_D - \vec{V}_B|$, is formed and used as a final state tagging variable [60].

In e^+e^- and $p\overline{p}$ collisions, b-quarks are produced in $b\overline{b}$ pairs. Therefore, in addition to the B meson studied for mixing, the event contains another b-hadron with a b-flavour opposite to that of the B meson at production. The flavour of the B meson at production can therefore be determined from the flavour of the second b-hadron in the event. Some additional information is available from the fragmentation products of the B meson: the s-quark of the B_s meson (or d-quark for B_d) is created in a $s\overline{s}$ ($d\overline{d}$) pair, most of the times the second s-quark forms a kaon (pion) the electric charge of which (if non zero) is correlated with that of the B meson. The two pieces of information can be combined to obtain a single estimator of the initial b-flavour of the B meson. At SLC, the e^+e^- beams are polarized, which allows the b-forward-backward asymmetry to be used to further improve the initial

state b-flavour tagging. Typical effective values of the probability of incorrect initial state flavour assignment, the initial mistag, at LEP are around $\eta_I \sim 25\%$ while at SLD, mostly thanks to the polarized beams, $\eta_I \sim 15\%$ can be reached.

The final label mixed/unmixed obtained for a particular B meson decay has a certain probability of being incorrect, which is called the total mistag probability, η_T. This quantity is evaluated from a combination of the mistag for the initial, η_I, and final state, η_F, of the B meson as

$$1 - 2\eta_T = (1 - 2\eta_I)(1 - 2\eta_F) \; . \tag{3.6}$$

The combined mistag probability can be evaluated on average in the selected sample, or, as it is done in most of the cases, computed event by event. Simulated events are usually used to convert the combined tagging variable of each event into a mistag probability. The use of event-by-event information increases significantly the statistical power of an oscillation analysis.

The mistag probability, η_T, is used in the fitting procedure to weight events according to their probability of being correctly tagged.

3.1.3 Signal enrichment and background evaluation

Inclusive or semi-inclusive event sample selections result on poor signal purity: as in an unbiased $b\overline{b}$ sample, only about 10% of the events contain B_s meson decays. Some decay properties like the estimated electric charge of the secondary vertex, or the presence of kaons both in the fragmentation and the decay of the B_s candidate, can be used to statistically increase the effective signal purity. All the relevant discriminant variables are combined, and the result is used to define an event-by-event signal purity. More detailed explanations on this subject are given in Section 6.9.

Semi-exclusive selections, like the $D_s\ell$ reconstruction explained in Section 7.1, have an average signal purity of the order of 40%. The effective signal purity can still be increased with techniques similar to those used in inclusive selections. In the case of hadronic reconstruction of specific decays, the average purity is even higher, of the order of 60%, and can be increased in an analogous manner with the B_s constructed mass as a key discriminant variable.

The non-signal components in the selected sample, the background, need to be identified and properly characterized. In the case of inclusive or semi-inclusive selections, the background contains mostly other b-hadrons (B_d, $B^{+/-}$, and b-baryons), but also a small fraction of hadrons formed with lighter quarks, called udsc-background in the following. The estimation of the relative amount of the background components is explained in Section 6.9, for the b-hadron species, and in Section 6.4, for the udsc background, in the case of the inclusive semileptonic selection. The background treatment of semi-exclusive or exclusive selections is slightly different and explained in Sections 7.1, 7.2.

3.1.4 Effect of mistag, background, and resolution

The non-perfect flavour tagging, the presence of background, and the finite detector resolution introduce a dilution on the oscillation signal studied. Figure 3.6 is used to illustrate

this point. On the three plots the expected fraction of events tagged as mixed is plotted

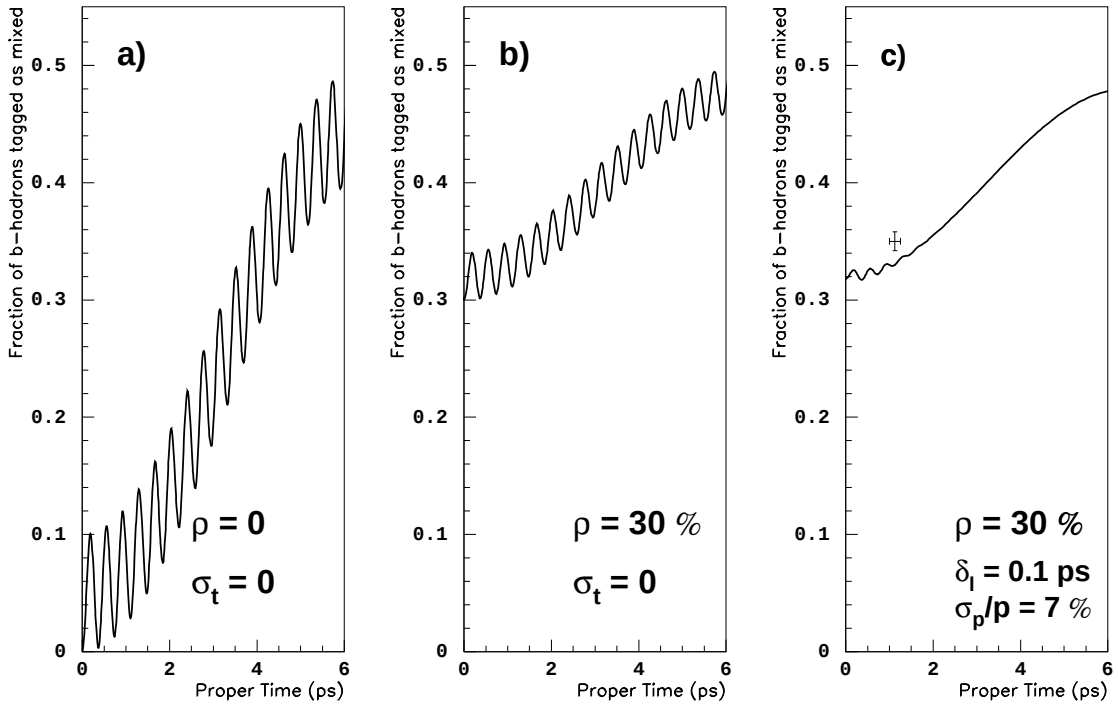

Figure 3.6: B_s oscillations signal resolution as a function of analysis parameters. The typical error bar for proper time values below 2 ps in an inclusive semileptonic analysis in ALEPH is shown.

as a function of the measured proper time. An inclusive selection with 10% B_s, 40% B_d, 40% B^+ and 10% b-baryons is considered, with oscillation frequencies $\Delta m_d = 0.487\,\mathrm{ps}^{-1}$ and $\Delta m_s = 17\,\mathrm{ps}^{-1}$. In Fig. 3.6a) perfect flavour tagging, both at production and decay, is assumed, as well as negligible proper time resolution. The rapid B_s meson oscillation is seen convoluted with the B_d mesons oscillation. In Fig. 3.6b) the effect of a global mistag probability of 30% is shown. The effective B_s oscillation amplitude is reduced rendering it more difficult to resolve. In Fig. 3.6c) the effect of the experimental resolution on the proper time measurement is also introduced, with momentum resolution $\sigma_p/p = 7\%$ and decay length resolution $\delta_l \equiv \sigma_l\, m/p_0\, c = 0.1\,\mathrm{ps}$ (these values are better than the average in most of the inclusive B_s oscillation analyses). As a consequence, B_s oscillations become only a marginal distortion on the B_d oscillation pattern, only visible at small proper time values. The typical experimental uncertainty for proper time values below 2 ps is shown both for the proper time and the asymmetry measurements in the case of an inclusive semileptonic analysis with the ALEPH data (Chapter 6). This last plot, compared to the others, is a clear illustration of the importance of the proper time resolution on B_s oscillations, as well as of the experimental difficulty of the measurement. The proper time resolution expression in Eq. 3.5 shows that the momentum resolution does not contribute at very small proper time values, the crucial issue for B_s oscillations analyses is therefore the decay length resolution.

3.2 Fitting procedure

Once the proper time of the selected B mesons and the mixed/unmixed label are determined, a fit to the expected proper time distribution is performed, to tentatively determine the value of the oscillation frequency.

A selected data sample is a mixture of signal, B_s mesons in this case, and different type of backgrounds. The probability density function (*p.d.f.*) has therefore to take into account the expected proper time properties of all the components in the sample, *i*) the oscillating signal component (B_s mesons); *ii*) the oscillating background component (B_d mesons); *iii*) the non-oscillation background component ($B^{+/-}$ and b-baryons); and *iv*) the udsc-background. The different parts of the *p.d.f.* are explicited below for events labelled as unmixed and mixed.

- Oscillating B mesons (B_s and B_d mesons):

$$\begin{aligned} p.d.f.^{\rm unmix}(t) &= f_{\rm B_q} \left[(1-\eta_T) \frac{1+\cos\Delta m_{\rm B_q} t_0}{2\,\tau_{\rm B_q}}\, e^{-t_0/\tau_{\rm B_q}} + \right. \\ &\qquad \left. \eta_T \frac{1-\cos\Delta m_{\rm B_q} t_0}{2\,\tau_{\rm B_q}}\, e^{-t_0/\tau_{\rm B_q}} \right] \otimes \mathcal{R}(t_0,t)\,, \\ p.d.f.^{\rm mix}(t) &= f_{\rm B_q} \left[(1-\eta_T) \frac{1-\cos\Delta m_{\rm B_q} t_0}{2\,\tau_{\rm B_q}}\, e^{-t_0/\tau_{\rm B_q}} + \right. \\ &\qquad \left. \eta_T \frac{1+\cos\Delta m_{\rm B_q} t_0}{2\,\tau_{\rm B_q}}\, e^{-t_0/\tau_{\rm B_q}} \right] \otimes \mathcal{R}(t_0,t)\,. \end{aligned} \tag{3.7}$$

- Non oscillating B hadrons ($B^{+/-}$ and b-baryons):

$$\begin{aligned} p.d.f.^{\rm unmix}(t) &= f_{\rm B} \left[(1-\eta_T) \frac{e^{-t_0/\tau_{\rm B}}}{\tau_{\rm B_d}} \right] \otimes \mathcal{R}(t_0,t)\,, \\ p.d.f.^{\rm mix}(t) &= f_{\rm B} \left[\eta_T \frac{e^{-t_0/\tau_{\rm B}}}{\tau_{\rm B}} \right] \otimes \mathcal{R}(t_0,t)\,. \end{aligned} \tag{3.8}$$

- The udsc background is most often parametrized from simulated events.

$$\begin{aligned} p.d.f.^{\rm unmix}(t) &= f_{\rm udsc} \left[(1-\eta_T)\, F^{\rm simul}(t) \right]\,, \\ p.d.f.^{\rm mix}(t) &= f_{\rm udsc} \left[\eta_T\, F^{\rm simul}(t) \right]\,. \end{aligned} \tag{3.9}$$

In the above expressions f_i ($i = {\rm B_q}, {\rm B}, {\rm udsc}$) are the fractions of each component in the selected sample, with $\sum_i f_i = 1$, and η_T is the mistag probability, which can be different for each component. The function $\mathcal{R}(t_0,t)$ is the proper time resolution function, built as a convolution of the decay length and momentum resolutions. All lifetimes, τ_i, are experimentally known, $\Delta m_{\rm d}$ is measured as well. The only free parameter, to be fit to the data, is the B_s oscillation frequency $\Delta m_{\rm s}$.

The above expressions are used, convoluted with the experimental proper time efficiency, and resolution in a global likelihood function of the data sample minimized with respect to $\Delta m_{\rm s}$. To date, no minimum deep enough for a $\Delta m_{\rm s}$ measurement has ever been found. Only lower limits on $\Delta m_{\rm s}$ have been set.

3.3 The amplitude method

Many different B_s oscillations analyses are being (or have been) performed by the LEP Collaborations, SLD and CDF. A method to easily combine their results and therefore increase the global final reach on Δm_s, known as the *amplitude method*, was first used by the ALEPH Collaboration [61] and is presented in Refs. [56, 9]. In this Section the basics of the method are explained, along with its interpretation.

The B_d oscillations are slow enough to allow the oscillation frequency Δm_d to be determined by several analyses in all experiments where b-hadrons are produced. An average value of Δm_d is obtained from the individual results of these analyses. In the case of B_s mesons, instead, as no analysis has yet been able to resolve the oscillation, the analogous approach is not applicable.

In the amplitude method the probability density function for the B_s signal is transformed by the introduction of the amplitude, $\mathcal{A}$, in front of the oscillation term as

$$p.d.f.^{\mathrm{u,m}}(t) = \frac{\Gamma_s e^{-\Gamma_s t}}{2}\left[1 \pm \cos(\Delta m_s t)\right] \;\Rightarrow\; p.d.f.^{\mathrm{u,m}}(t) = \frac{\Gamma_s e^{-\Gamma_s t}}{2}\left[1 \pm \mathcal{A}\cos(\omega t)\right] \;. \quad (3.10)$$

The method consists in measuring $\mathcal{A}$, *i.e.*, maximizing the likelihood with respect to $\mathcal{A}$, as a function of the test frequency ω (throughout this thesis ω stands for the frequency folded in the fitting function, whereas Δm_s or Δm indicates the frequency of the oscillations in the sample studied). A value of $\mathcal{A} = 1$ is expected to be measured at $\omega = \Delta m_s$. Far below the true oscillation frequency $\mathcal{A} = 0$ is expected; a more detailed description of the expected values of the amplitude $\mathcal{A}$ as a function of the test frequency is derived in Section 3.3.2.

The idea behind the amplitude method is to define a measurable quantity, with well defined uncertainties, in such a way that an average between different analyses can be easily performed. The amplitude $\mathcal{A}$ fulfils these requirements.

The frequency range for which the amplitude is found to be compatible with zero and incompatible with one can be excluded at 95% C.L., *i.e.*,

$$\omega_0 \;\text{ excluded at }\; 95\%\text{ C.L. }\; \text{if }\; \mathcal{A}_{(\omega_0)} + 1.645\,\sigma_{\mathcal{A}} < 1 \;. \quad (3.11)$$

The sensitivity of an analysis is defined as the expected limit at 95% C.L. given the analysis performance, *i.e.*, the limit which would be set if the measured value of $\mathcal{A}_{(\omega_0)}$ was zero. The corresponding sensitivity frequency ω^{95} fulfils,

$$\omega^{95} \;\Leftrightarrow\; 1.645 \;\times\; \sigma_{\mathcal{A}_{\omega^{95}}} \;=\; 1 \;. \quad (3.12)$$

It is equivalent to the frequency for which amplitude values zero and one can be distinguished at 95% C.L.

The amplitude $\mathcal{A}$ has well behaved parabolic uncertainties, and therefore a combination of different analyses is easily performed. A combined limit is obtained applying the above prescription to the combined amplitude scan as a function of ω.

As already mentioned, the amplitude measurements are obtained by maximizing the likelihood L of the proper time distributions of mixed and unmixed events with the amplitude of the oscillating term $\mathcal{A}$ as the free parameter, and ω fixed at a chosen value. Denote

$\mathcal{L} = -\log L$, its expansion at second order around the minimum of $\mathcal{L}$, $\mathcal{L}_\omega(\bar{\mathcal{A}})$, can be approximated by

$$\mathcal{L}_\omega(\mathcal{A}) \;\simeq\; \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{\mathcal{A}-\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega , \tag{3.13}$$

where $\bar{\mathcal{A}}$ is the measured value of the amplitude, and $\sigma_{\mathcal{A}}$ is the uncertainty on $\bar{\mathcal{A}}$. This approximation turns out to be very accurate in reality, $\mathcal{L}_\omega(\mathcal{A})$ being parabolic in a wide range around $\bar{\mathcal{A}}$.

From Eq. 3.13 it follows that, again for each value of ω:

$$\mathcal{L}_\omega(\mathcal{A}=1) \;\simeq\; \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{1-\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega . \tag{3.14}$$

The oscillation vanishes for $\mathcal{A}=0$ on the one hand, and it averages to zero for $\omega \to \infty$ due to finite resolution on the other. Therefore, the following equality can be written

$$\mathcal{L}_{\omega\to\infty}(\text{any } \mathcal{A}) = \mathcal{L}_{\text{any } \omega}(\mathcal{A}=0) \;(\equiv \mathcal{L}_\infty), \tag{3.15}$$

and therefore, from Eq. 3.13,

$$\mathcal{L}_\infty \;= \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega . \tag{3.16}$$

If Eq. 3.16 is subtracted from Eq. 3.14, the following formula is obtained

$$\Delta\mathcal{L}(\omega) \;\equiv\; \mathcal{L}_\omega(\mathcal{A}=1) - \mathcal{L}_\infty \simeq \left[\frac{1}{2}\left(\frac{1-\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2 - \frac{1}{2}\left(\frac{\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2\right]_\omega , \tag{3.17}$$

which allows the value of $\Delta\mathcal{L}$ to be calculated, for each ω, from the fit amplitude and its uncertainty. This formula was already given in Ref. [56] and allows a minimum of the likelihood of the world combined data (or any other combination) to be looked for.

3.3.1 Interpretation

The properties of the amplitude method for B_s oscillations were studied in detail in a paper (Ref. [9]) written by Duccio Abbaneo and myself. The three main Sections of this paper are reproduced below (some changes of notation are introduced for coherence with the rest of this thesis). In Section 3.3.2 analytical expressions for the expected shape of the measured amplitude and its uncertainty are, for the first time, derived. The small and large frequency limits are discussed, and an approximate interpretation in terms of Fourier transformations is proposed. Several mis-concepts present in the previous literature are clarified. The probability of observing statistical fluctuations which would fake a signal in a sample with the oscillation frequency far beyond the sensitivity is also discussed. In Section 3.3.3 the structure and the features of a toy-experiment generator used throughout the paper, and this thesis, are described. A procedure to tune the parameters of the simulation in order to reproduce the observed uncertainties in a particular analysis or combination of analyses is given. In

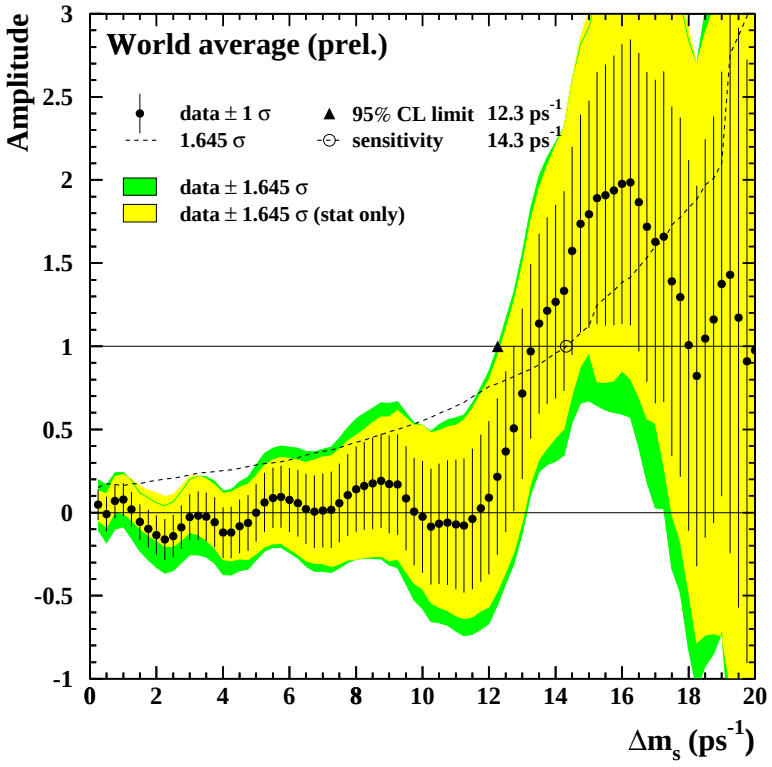

Figure 3.7: Winter 1999 combined amplitude measurements as a function of ω, from the B Oscillation Working Group.

Section 3.3.4 a procedure to extract a confidence level value from the likelihood function is presented and discussed. The uncertainty arising from the lack of a detailed simulation is investigated. The method described in this Section is used in Chapter 8 for the interpretation of the results of this thesis.

At the time of the paper publication, the latest results on B_s oscillations available were those presented in the 1999 winter conferences, and therefore these are the results used for illustration throughout the paper and in its reproduction below. For the sake of completeness, the combined amplitude measurements obtained from published and preliminary analyses available at the time of the 1999 winter conferences [8] are presented in Fig. 3.7.

The likelihood difference $\Delta\mathcal{L}(\omega)$ obtained for these data is shown in Fig. 3.8. A good parametrization for the shape of $\Delta\mathcal{L}$ is obtained with a function $f(\omega) \propto 1/\omega^{\alpha}$ with $\alpha = 1.64$, plus some Gaussian functions to describe the deviations. A parabolic fit of the three lowest points of the plot gives a minimum for $\omega = 14.8\,\mathrm{ps}^{-1}$, with a value $\Delta\mathcal{L}_{\mathrm{min}} = -2.9$. As discussed in the following sections, the significance of this minimum cannot be extracted in an analytical way, but needs to be determined with toy experiments.

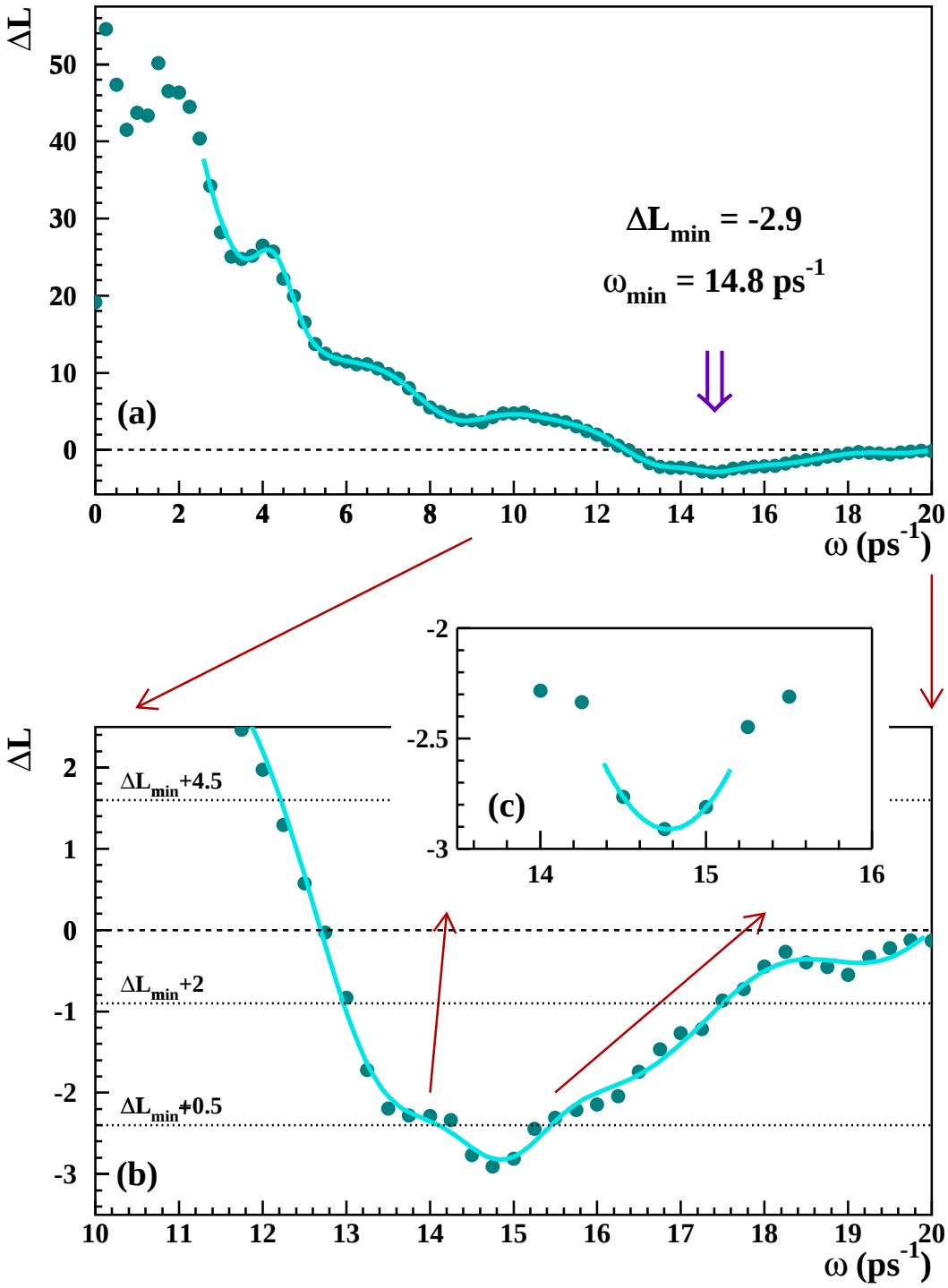

Figure 3.8: Likelihood as a function of ω derived from the combined amplitude measurements. A minimum is observed for $\omega = 14.8\,\mathrm{ps}^{-1}$. The parametrization described in the text is shown in (a) and (b); the parabolic fit to the three lowest points in (c).

3.3.2 The amplitude analysis

The true proper time distribution of mixed and unmixed B meson decays is written as follows:

$$p.d.f._{\rm u,m}^{0}(t_0) \;=\; \Gamma \exp\left(-\Gamma t_0\right) \frac{1 \pm \cos \Delta m \, t_0}{2} \;\equiv\; \frac{E^0(t_0) \pm f^0_{\Delta m}(t_0)}{2}\,, \tag{3.18}$$

where $f^0_{\Delta m}(t_0)$ contains the oscillation term. The plus (minus) sign holds for unmixed (mixed) events. Any difference in the decay widths of the two mass eigenstates has been neglected.

The reconstructed proper time distributions can then be written as:

$$p.d.f._{\rm u,m}(t) \;=\; \int_0^\infty dt_0 \; \frac{E^0(t_0) \pm f^0_{\Delta m}(t_0)}{2} \, \mathcal{R}(t_0, t) \;\equiv\; \frac{E(t) \pm f_{\Delta m}(t)}{2}\,. \tag{3.19}$$

For the sake of simplicity, no time dependent selection efficiency has been considered in the calculations. In what follows, it is assumed that the *relative uncertainty* on the b-hadron momentum, and the *absolute uncertainty* on the decay length are Gaussian. This approximation follows what typically happens in real analyses, where the uncertainty on the reconstructed b-hadron momentum is found to roughly scale with the momentum itself, while the uncertainty on the decay length does not. This fact has important consequences in the way the two resolution components affect the amplitude shape.

Under these assumptions, the resolution function $\mathcal{R}(t_0, t)$ can be written as:

$$\begin{aligned}
\mathcal{R}(t_0, t) \;&=\; \int_{-\infty}^{\infty} dp \; \frac{1}{\sqrt{2\pi}\,\sigma_p} \exp\left(-\frac{(p-p_0)^2}{2\,\sigma_p^2}\right) \frac{1}{\sqrt{2\pi}\,\sigma_l} \exp\left(-\frac{(pct - p_0 ct_0)^2}{2\,(m\sigma_l)^2}\right) \frac{pc}{m} \\
&\approx\; \frac{1}{\sqrt{2\pi\,[\delta_l^2 + (\delta_p t)^2]}} \exp\left(-\frac{(t-t_0)^2}{2\,[\delta_l^2 + (\delta_p t)^2]}\right)\,,
\end{aligned} \tag{3.20}$$

where $\delta_l \equiv \sigma_l \, m/(p_0\, c)$, $\delta_p \equiv \sigma_p/p_0$. The approximation is valid if δ_p is significantly smaller than one, which is anyway required to assume Gaussian uncertainties, since the reconstructed momentum cannot be negative. Furthermore, p_0 is not accessible in real data; the reconstructed momentum is therefore used in the evaluation of the uncertainty from the decay length resolution: $\delta_l \approx \sigma_l \, m/(p\, c)$.

A set of parameters is chosen here for the purpose of illustration. Resolution values of $\delta_p = 0.15$ and $\delta_l = 0.14\,$ps are used; the latter one would correspond to a monochromatic sample of $\rm B_s$ mesons with $p_0 = 32$ GeV/c and $\sigma_l = 250\,\mu$m. In a real analysis the normalization of the non–oscillating component is the total number N of b-hadron decays (differences in lifetime are neglected), while the oscillation term is multiplied by $N\, f_{\rm B_s}\,(1 - 2\,\eta_T)$, $f_{\rm B_s}$ being the fractions of $\rm B_s$ in the sample, and η_T the global mistag rate. For an inclusive analysis, $f_{\rm B_s}\,(1 - 2\,\eta_T)$ is typically about 0.05. The curves obtained with these parameters, normalization factors omitted, are shown in Fig. 3.9. As the frequency increases, the oscillation amplitude is damped because of the resolution. For very large frequencies only the first period can be resolved.

The fitting technique commonly used in the amplitude analysis is a simultaneous maximum–likelihood fit to the proper time distributions of mixed and unmixed events, as explained in Section 3.2. Alternatively, the difference of the two distributions, *i.e.* the oscillating term, can be fit with a binned χ^2 method. These two methods are discussed in the following.

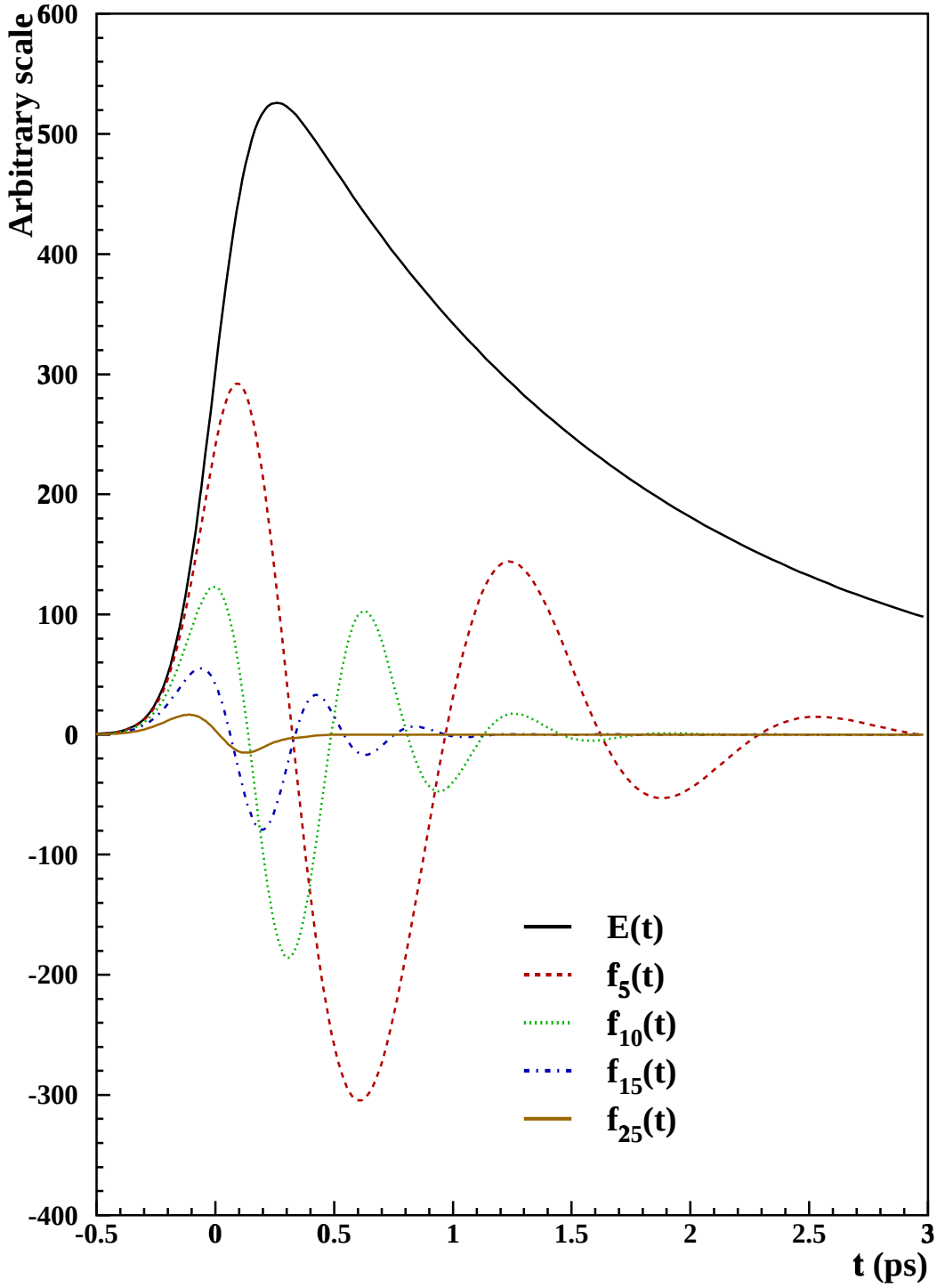

Figure 3.9: Reconstructed proper time distributions for the non–oscillating component, $E(t)$, and the oscillating component, $f_{\Delta m}(t)$, at different values of Δm. Resolutions of $\delta_p = 0.15$ and $\delta_l = 0.14$ps are assumed.

The maximum likelihood fit. Using the aforementioned formalism, the likelihood function can be written as:

$$\begin{aligned} -\log L \;=\; \frac{1}{2}\int_{-\infty}^{\infty} dt \quad & [E(t)+f_{\Delta m}(t)]\,\log\left[E(t)+\mathcal{A}f_{\omega}(t)\right] \\ + \quad & [E(t)-f_{\Delta m}(t)]\,\log\left[E(t)-\mathcal{A}f_{\omega}(t)\right] \;+\; \text{Const}\,, \end{aligned}$$

where again Δm is the frequency of the oscillations in the sample analysed, and ω is the value chosen in the fitting function. The minimization with respect to $\mathcal{A}$ leads to the condition

$$\int_{-\infty}^{\infty} dt\; \frac{f_{\omega}(t)\, f_{\Delta m}(t) - \mathcal{A} f_{\omega}^2(t)}{E(t)\left(1-\mathcal{A}^2\frac{f_{\omega}^2(t)}{E^2(t)}\right)} \;=\; 0\,, \tag{3.21}$$

which allows $\mathcal{A}$ to be determined.

The χ^2 fit. Similarly, the χ^2 can be written as

$$\chi^2 \;=\; \int_{-\infty}^{\infty} dt\; \frac{\left[f_{\Delta m}(t)-\mathcal{A}f_{\omega}(t)\right]^2}{E(t)}\,, \tag{3.22}$$

the minimization of which gives

$$\int_{-\infty}^{\infty} dt\; \frac{f_{\omega}(t)\, f_{\Delta m}(t) - \mathcal{A} f_{\omega}^2(t)}{E(t)} \;=\; 0. \tag{3.23}$$

Equations 3.21 and 3.23 both give $\mathcal{A}=1$ for $\omega=\Delta m$. For $\omega\neq\Delta m$ they are equivalent if $\mathcal{A}f_{\omega}(t)$ is negligible compared to $E(t)$.

The expression of $\mathcal{A}_{\Delta m}(\omega)$ can be derived from Eq. 3.23 as

$$\mathcal{A}_{\Delta m}(\omega) \;=\; \frac{\displaystyle\int_{-\infty}^{\infty} dt\; \frac{f_{\omega}(t)\, f_{\Delta m}(t)}{E(t)}}{\displaystyle\int_{-\infty}^{\infty} dt\; \frac{f_{\omega}^2(t)}{E(t)}}\,. \tag{3.24}$$

The resulting amplitude curves for $\Delta m = 5,\ 10,\ 15\ \text{ps}^{-1}$ are shown in Fig. 3.10a. On top of the curves, values obtained from the likelihood fit (Eq. 3.21) are also shown. The two fitting methods are indeed equivalent for $\omega\approx\Delta m$, as expected, while some difference appears for $\omega\neq\Delta m$.

The expected amplitude is unity at $\omega=\Delta m$. For $\omega>\Delta m$, the behaviour depends on Δm (for given resolutions). In this example, for $\Delta m = 15\,\text{ps}^{-1}$ the expected amplitude increases monotonically.

The expressions derived for the χ^2 fit allow also the expected uncertainty on the measured amplitude to be extracted,

$$\chi^2(\mathcal{A}+\sigma_{\mathcal{A}}) - \chi^2(\mathcal{A}) \;=\; 1, \tag{3.25}$$

which in turn gives

$$\sigma_{\mathcal{A}}(\omega) \;=\; \frac{1}{\sqrt{\displaystyle\int_{-\infty}^{\infty} dt\; \frac{f_{\omega}^2(t)}{E(t)}}}\,. \tag{3.26}$$

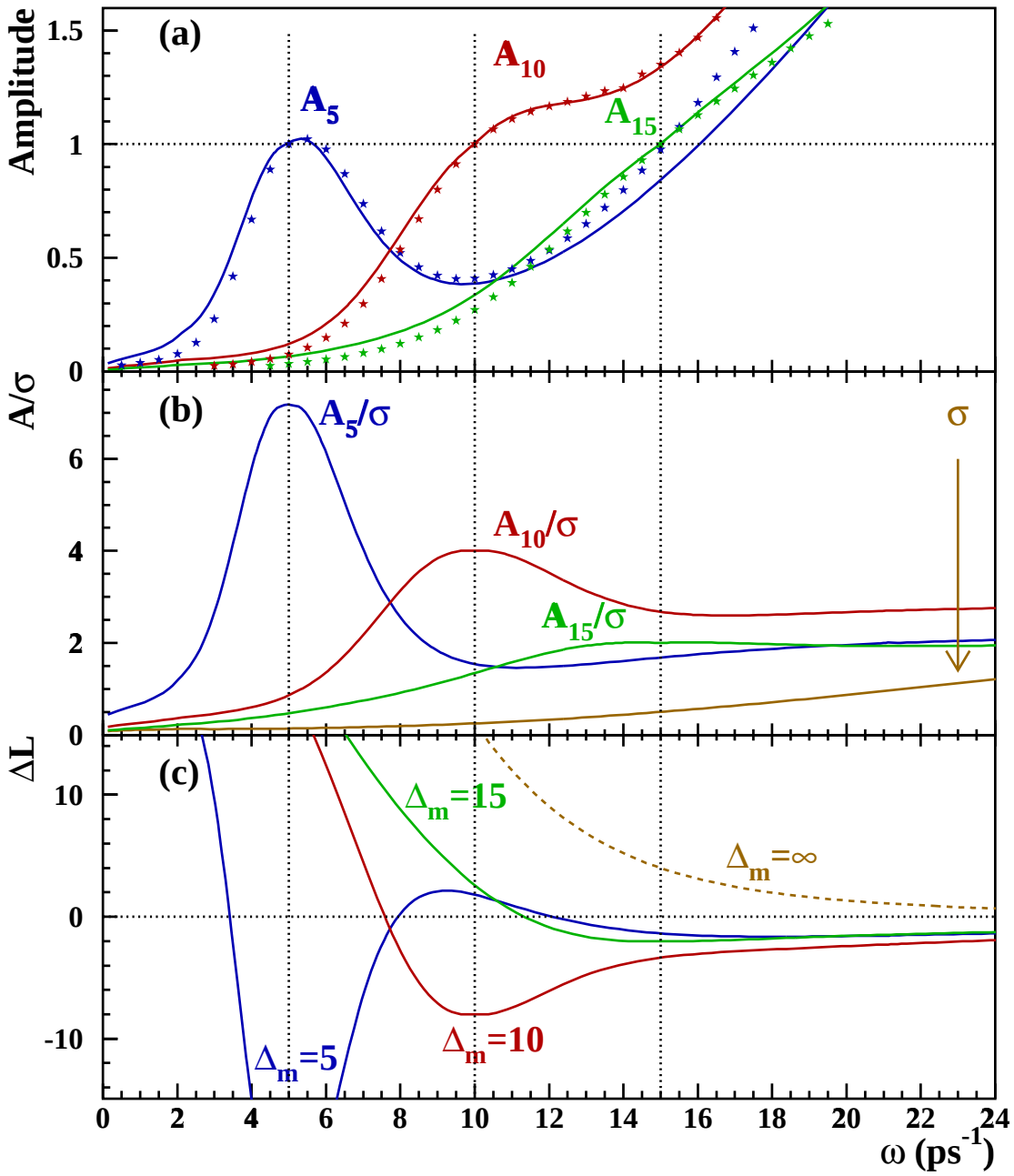

Figure 3.10: (a) Expected amplitude values for Δm = 5, 10, 15, ps^{-1}. The curves refer to the χ^2 minimization, the points to the likelihood fit.
(b) Amplitude significance curves (χ^2 fit). The expected shape of $\sigma(\omega)$ is also shown.
(c) Expected shape of the likelihood, derived from the amplitude and its uncertainty. The dashed line corresponds to the the limit $\Delta m \to \infty$.
Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu$m are assumed for the three plots.

If a global mistag probability η_T, and the B_s signal fraction in the data f_s are introduced, the expression for the amplitude curve remains unchanged, but the expected uncertainty on the amplitude becomes, for a statistics of N events,

$$\sigma_{\mathcal{A}}^{-1}(\omega) \;=\; \sqrt{N}\,(1-2\eta_T)\,f_s\,\sqrt{\int_{-\infty}^{\infty} dt\;\frac{f_\omega^2(t)}{E(t)}}\;. \tag{3.27}$$

The significance of the measured amplitude is therefore:

$$S_{\Delta m}(\omega) \;=\; \frac{\mathcal{A}_{\Delta m}(\omega)}{\sigma_{\mathcal{A}}^{\Delta m}(\omega)} \;=\; \frac{\displaystyle\int_{-\infty}^{\infty} dt\;\frac{f_\omega(t)\,f_{\Delta m}(t)}{E(t)}}{\displaystyle\sqrt{\int_{-\infty}^{\infty} dt\;\frac{f_\omega^2(t)}{E(t)}}}\,\sqrt{N}\,(1-2\eta_T)\,f_s\;. \tag{3.28}$$

This latter equation is correct only because $\mathcal{A}$ and $\sigma_{\mathcal{A}}$ are independent.

The amplitude significance curves for $\Delta m = 5,\ 10,\ 15\,\mathrm{ps}^{-1}$ are shown in Fig. 3.10b. The normalization of the uncertainty, in the same figure, is arbitrarily chosen to have $\sigma_{\mathcal{A}} = 0.5$ at $\omega = 15\,\mathrm{ps}^{-1}$.

The expected significance is maximal at $\omega = \Delta m$. For $\omega > \Delta m$ it decreases without reaching zero in the range explored. The decrease is more smooth for high values of Δm.

The expected shape of the likelihood, as calculated from the amplitude and its uncertainty using Eq. 3.17, is shown in Fig. 3.10c.

Limits for small and large Δm

In the limit of very small or very large Δm, some approximations can be made in the formulae, which yield simplified expressions of easier interpretation.

Small Δm. If $\delta_l \ll 1/\Delta m,\ \delta_p/\Gamma \ll 1/\Delta m$, the oscillation is slow and marginally affected by the resolution. This limit holds in the case of B_d oscillations. If the resolution effects are neglected, Eq. 3.24 can be rewritten as

$$\mathcal{A}_{\Delta m}(\omega) \;=\; \frac{\displaystyle\int_0^{\infty} dt\;\Gamma\exp(-\Gamma t)\,\cos\omega t\,\cos\Delta m t}{\displaystyle\int_0^{\infty} dt\;\Gamma\exp(-\Gamma t)\,\cos^2\omega t}\,, \tag{3.29}$$

which gives

$$\mathcal{A}_{\Delta m}(\omega) \;\approx\; \frac{\dfrac{\Gamma^2}{\Gamma^2+(\omega+\Delta m)^2} + \dfrac{\Gamma^2}{\Gamma^2+(\omega-\Delta m)^2}}{1+\dfrac{\Gamma^2}{\Gamma^2+4\,\omega^2}}\,. \tag{3.30}$$

The resulting shape is shown in Fig. 3.11. The dots superimposed are obtained with toy experiments generated at the same value of the frequency, including resolution effects (for details on the simulation see Section 3.3.3; the parameters used in the generation are those of samples **S** there defined). The two shapes are in qualitatively good agreement.

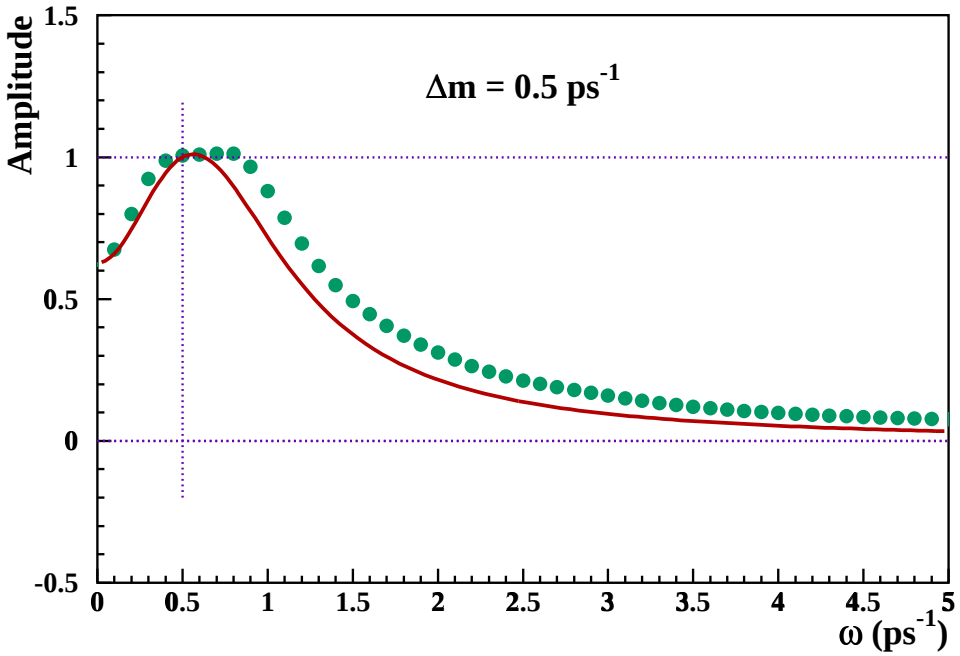

Figure 3.11: The full curve gives the expected shape of the amplitude for a signal at $\Delta m = 0.5\,\mathrm{ps}^{-1}$ when all resolution effects are neglected. The dots are obtained with a toy experiment in which resolution effects are simulated (where $\delta_p = 0.15$ and $\sigma_l = 250\,\mu\mathrm{m}$). The two shapes are in agreement.

Large Δm. In this limit, which corresponds to the regime of B_s oscillations, the resolution effects dominate. If $\Delta m_s \approx 15\,\mathrm{ps}^{-1}$ and $\delta_p = 0.15$, then $\delta_p/\Gamma \simeq 0.23\,\mathrm{ps}$, which is larger than $1/\Delta m_s \approx 0.07\,\mathrm{ps}$ and therefore implies that only events with small proper time contribute to the sensitivity. Similarly, taking $\delta_l = 0.14\,\mathrm{ps}$ gives $\delta_l > 1/\Delta m_s$, which implies a substantial damping of the amplitude of the oscillating term due to the decay length resolution. In this case, a useful approximation is to assume that the term $E(t)$ in Eq. 3.24 varies slowly compared to the fast oscillating term $f_\omega(t)$, which is non-zero in a limited time range (Fig. 3.9), and take it out of the integral. In this way the expression can be simplified and rewritten in terms of the Fourier transformation of the oscillating components,

$$\mathcal{A}_{\Delta m}(\omega) \;\approx\; \frac{\int_{-\infty}^{\infty} dt\;\, f_\omega(t)\, f_{\Delta m}(t)}{\int_{-\infty}^{\infty} dt\;\, f_\omega^2(t)} \;=\; \frac{\int_{-\infty}^{\infty} d\nu\; \tilde{\mathrm{f}}_\omega(\nu)\, \tilde{\mathrm{f}}_{\Delta m}(\nu)}{\int_{-\infty}^{\infty} d\nu\; \tilde{\mathrm{f}}_\omega{}^2(\nu)}\,. \tag{3.31}$$

The approximation is valid only if both ω and Δm are large. The functions $\tilde{\mathrm{f}}_\omega$ are shown in Fig. 3.12a for a few different values of $\omega \geq 10\,\mathrm{ps}^{-1}$. Fig 3.12c shows the product of two of these Fourier transformations to illustrate the behaviour of the ratio in Eq. 3.31.

With increasing Δm, the frequency spectra, $\tilde{\mathrm{f}}_\omega$, become broader and smaller in amplitude. High *true* frequencies, Δm, have their spectrum damped faster than low frequencies, and the peak at $\omega \approx \Delta m$ disappears for Δm well beyond the sensitivity (Fig. 3.12b). Due to the broadening of the spectra, the product $\tilde{\mathrm{f}}_{\omega_1}(\nu)\tilde{\mathrm{f}}_{\omega_2}(\nu)$ is peaked around the smallest between

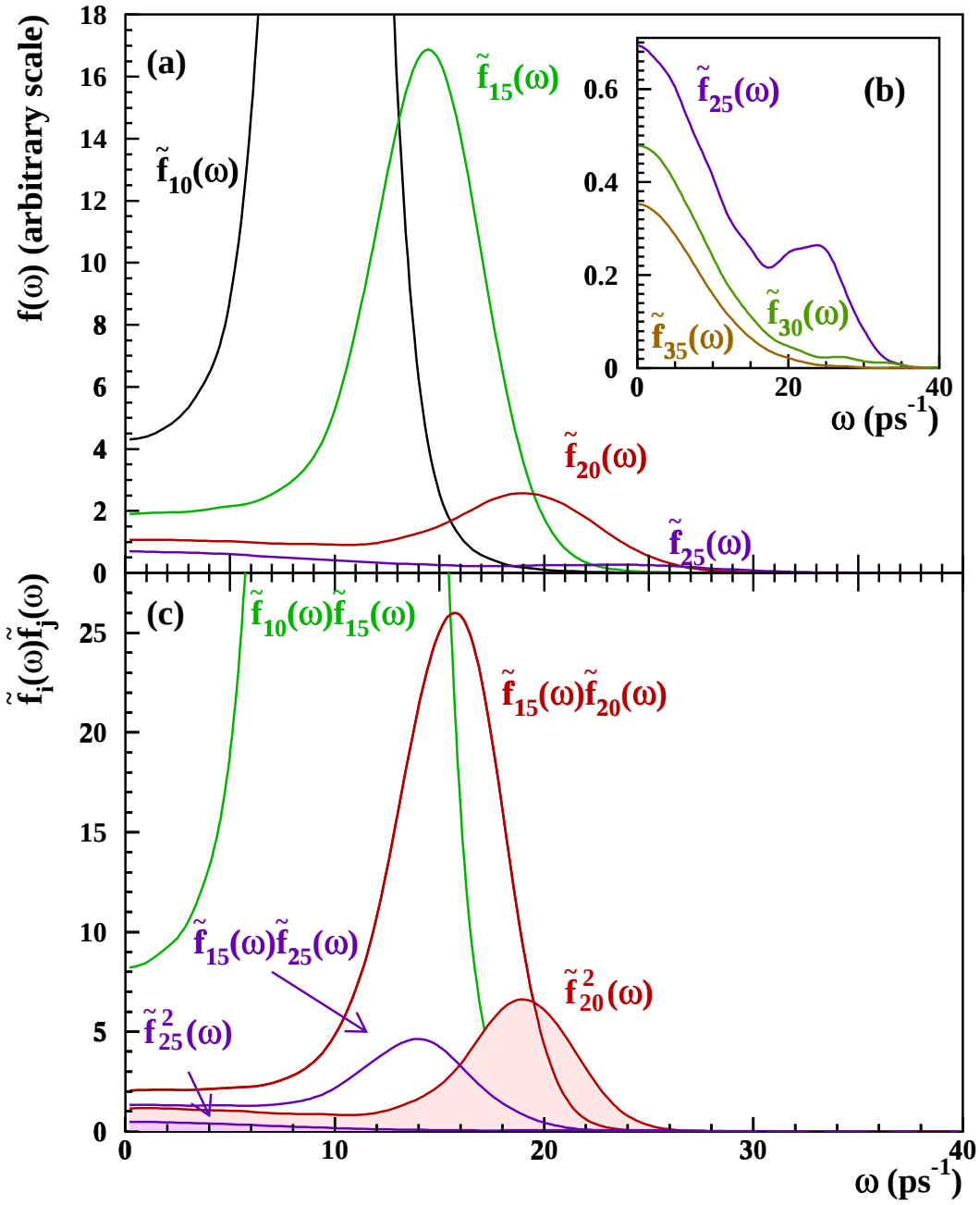

Figure 3.12: (a) Expected shapes of the Fourier spectra $\tilde{f}_{\Delta m}$ for different values of Δm. The spectra become broader and lower in amplitude when Δm increases.
(b) Detail of the spectra for high Δm.
(c) Products of pairs of Fourier spectra. The resulting functions are peaked around the smallest of the two frequency values.
Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu$m are assumed.

ω_1 and ω_2 (Fig. 3.12c). This fact implies that when a sample with oscillations at frequency Δm is analysed with a function containing a frequency $\omega < \Delta m$, the measured amplitude is dominated by the frequencies around ω; therefore the shape of $\mathcal{A}_{\Delta m}(\omega)$ for $\omega < \Delta m$ resembles that of $\tilde{\mathrm{f}}_{\Delta m}(\omega)$. For $\omega > \Delta m$ the frequencies around Δm are always tested, with a normalization factor which increases with ω (in Eq. 3.31, the denominator decreases very fast), therefore $\mathcal{A}_{\Delta m}(\omega)$ increases monotonically.

In order to understand better the effect of the decay length and proper time resolution, it is useful to study them separately. Setting $\delta_p = 0$ in Eq. 3.20, the following simplified expression can be obtained,

$$\tilde{\mathrm{f}}_{\Delta m}(\omega) \;=\; \frac{1}{2}\left[\frac{\Gamma^2}{\Gamma^2+(\omega+\Delta m)^2} + \frac{\Gamma^2}{\Gamma^2+(\omega-\Delta m)^2}\right]\exp\left(-\frac{\delta_l^2\omega^2}{2}\right), \tag{3.32}$$

which shows that the decay length resolution is responsible for the damping of the high frequencies.

Considering the momentum resolution alone, the following expression is obtained,

$$\tilde{\mathrm{f}}_{\Delta m}(\omega) \;=\; \int_{-\infty}^{\infty} d\nu\; \frac{1}{2}\left[\frac{\Gamma^2}{\Gamma^2+(\nu+\Delta m)^2} + \frac{\Gamma^2}{\Gamma^2+(\nu-\Delta m)^2}\right]\exp\left(-\frac{(\omega-\nu)^2}{2\,(\delta_p\omega)^2}\right), \tag{3.33}$$

which shows that the momentum resolution causes the broadening of the frequency spectrum (as intuitively expected, since a shift in the reconstructed momentum is equivalent to a change of scale on the time axis). For $\Delta m = 15\,\mathrm{ps}^{-1}$ and $\delta_p = 0.15$, the width of the frequency spectrum is dominated by the momentum resolution.

A broader frequency spectrum corresponds to a broader structure in the amplitude spectrum, or, equivalently, to higher correlations between values of the amplitude measured at different frequencies. This property is relevant for the confidence level estimation as explained in Section 3.3.4.

Fluctuations

The expected shape of the likelihood for a sample with oscillations at a frequency far beyond the sensitivity is shown in Fig. 3.10c. In a given frequency range, statistical fluctuations of the likelihood can produce values below 0 which can fake a signal. The probability of observing that $\Delta\mathcal{L}$ is lower than a given value $\overline{\Delta\mathcal{L}}$ at a given frequency ω can be estimated from Eq. 3.17, using the fact that the uncertainties on the measured amplitudes are found to be Gaussian with high precision:

$$\mathcal{P}(\Delta\mathcal{L},\omega) \;\equiv\; \mathcal{P}(\Delta\mathcal{L} < \overline{\Delta\mathcal{L}})_\omega \;=\; \frac{1}{2}\,\mathrm{erfc}\left[\left(-\overline{\Delta\mathcal{L}}\,\sigma_{\mathcal{A}}(\omega) + \frac{1}{2\,\sigma_{\mathcal{A}}(\omega)}\right)\Big/\sqrt{2}\right]. \tag{3.34}$$

The function $\mathcal{P}(\Delta\mathcal{L},\omega)$ is shown in Fig 3.13a, where the same parameters and normalization as for Fig. 3.10 are used. This function can be used as an estimator of the signal–ness of a given sample. Estimator contours, equidistant on a logarithmic scale, are drawn in Fig. 3.13b. Small negative values are most probable at high frequencies, while for larger negative values the maximum of the probability is found at lower frequencies.

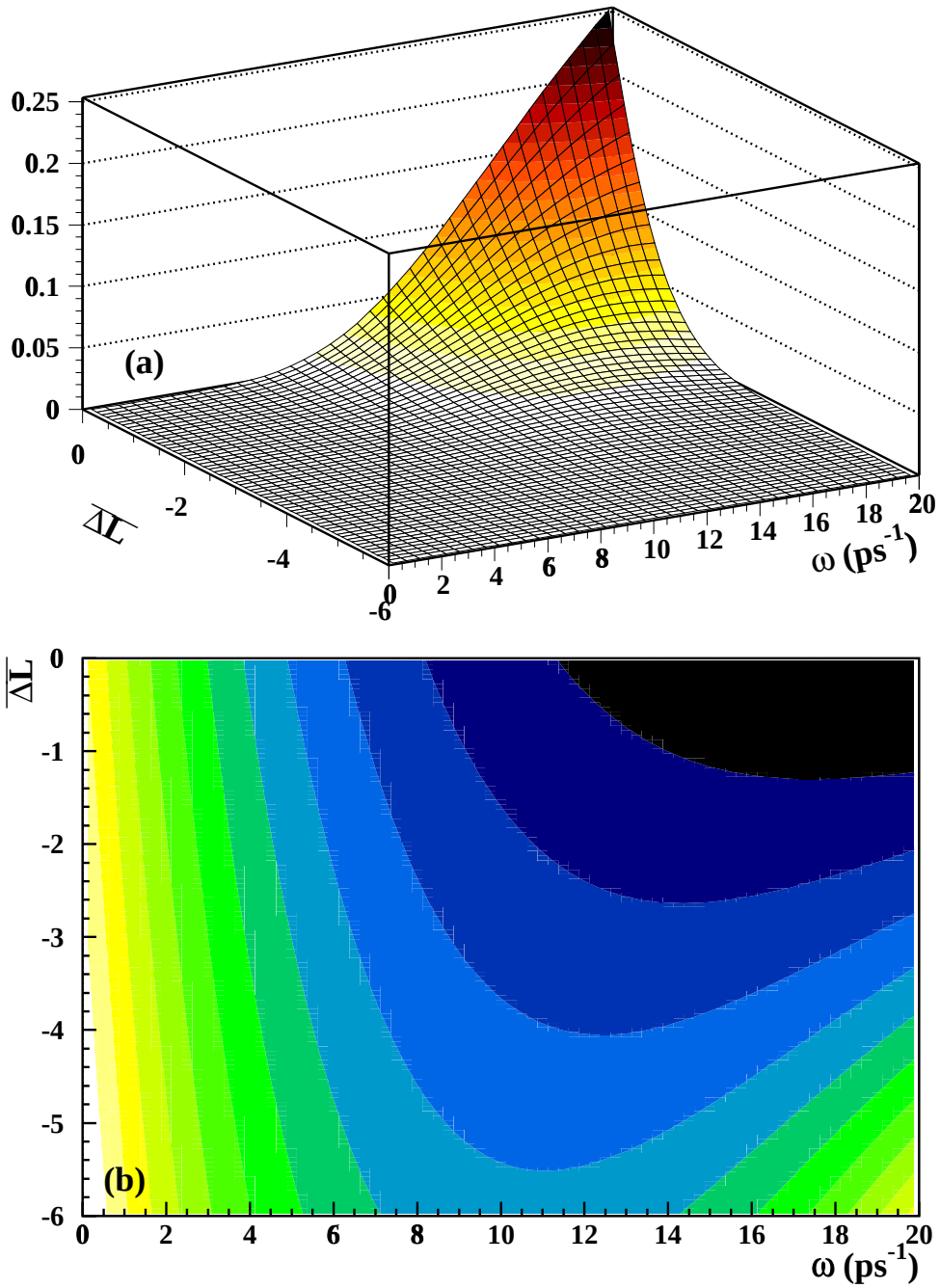

Figure 3.13: (a) Probability to find a $\Delta\mathcal{L}$ value lower than $\overline{\Delta\mathcal{L}}$ at a frequency ω. (b) Contours of equal probability in the $(\overline{\Delta\mathcal{L}}, \omega)$ plane. Values between -3 and -5 are found with higher probability when the uncertainty on the amplitude is from 0.5 (at $\omega = 15\,\mathrm{ps}^{-1}$ in this example) to 0.3 ($\omega \approx 11\,\mathrm{ps}^{-1}$). The contours shown are equidistant on a logarithmic scale.
Resolutions of $\delta_p = 0.15$ and $\delta_l = 0.14\,\mathrm{ps}$ are assumed.

3.3.3 The toy experiments

The probability that the minimum observed in the likelihood (Fig. 3.8) is caused by a fluctuation can be evaluated by means of toy experiments with the above estimator (Eq. 3.34).

In a general case, the depth of the likelihood minimum can be translated to a statistical significance in the approximation that the likelihood is parabolic, which is not the case here.

At each frequency point, the probability that the measured $\Delta\mathcal{L}$ is lower than a given value $\overline{\Delta\mathcal{L}}$ can be calculated as explained in Section 3.3.2, starting from the uncertainties on the measured amplitude. This procedure cannot be applied to the minimum, since $\Delta\mathcal{L}_{\rm min}$ is not an "unbiased" value, but it is chosen as the lowest value found over a certain frequency range explored.

The sum of the probabilities of obtaining a likelihood value lower than observed at all points where the amplitude is measured does not provide a good estimate either. The different points are highly correlated and they cannot fluctuate independently, therefore the sum of the individual probabilities would give a gross overestimate of the overall probability of finding a minimum as or more unlikely than the one observed.

The only viable possibility is to calibrate the significance of the structure observed with the help of toy experiments. The worldwide combination includes many analyses, and a detailed simulation of each of them is highly impractical. The procedure adopted here is to choose a set of parameters for the generation of the toy experiments such that each experiment is equivalent to the world average. The set of parameters cannot be uniquely determined from the data: it turns out that some parameters need to be fixed *a priori*, and therefore the dependence of the result obtained upon the particular choice adopted needs to be understood. The possible effects of the lack of a detailed simulation are investigated in Section 3.3.4 by studying the dependence of the correlations in the amplitude measurements upon the parameters chosen to generate the toy experiments.

Generation

The basic features of the toy experiments used to estimate the significance of the likelihood minimum can be summarized as follows.

- Bottom hadron species are generated according to a chosen composition.
- For each species, the *true* proper time t_0 of each b-hadron is generated according to an exponential with decay constant equal to its width, Γ, multiplied by a given efficiency function.
- Neutral B mesons are allowed to mix. Mixed and unmixed particles have their proper time distributions modified by the appropriate oscillating term, with given frequency.
- The *true* momentum p_0 is generated according to a Peterson distribution, tuned to reproduce a given mean scaled energy $\langle x_b \rangle$.
- The *true* decay length is then obtained, for each b-hadron, from

$$l_0 = \frac{t_0\, p_0}{m}\, c\,.$$

- A smearing is applied to the *true* decay length and momentum according to given resolution functions, to obtain the *measured* decay length and momentum, l and p.
- The *measured* proper time is hence calculated as

$$t = \frac{l\,m}{p\,c}\,.$$

- A mixed/unmixed tag is assigned to the generated hadrons using specified mistag rates.
- The udsc background is neglected.

The choice of the parameters

The only information, at the level of the world combination, which can drive the choice of the parameters for the simulation is provided by the uncertainty on the measured amplitude as a function of the frequency ω. The step (seen in Fig. 3.7) at $\omega = 15\,\mathrm{ps}^{-1}$ is due to some analyses in which the scan was not performed beyond that value of the frequency. The step at $\omega = 19\,\mathrm{ps}^{-1}$ is due to the SLD analyses, for which no measurement was provided for $\omega > 19\,\mathrm{ps}^{-1}$. In all what follows the four points with $\omega > 19\,\mathrm{ps}^{-1}$ are ignored, in order to reduce the pathologies in the uncertainty shape.

The uncertainties on the measured amplitudes can be formally written as (see also Eq. 3.27)

$$\sigma_{\mathcal{A}}^{-1}(\omega) \;=\; \sqrt{N}\,f_{\mathrm{B_s}}\,(1-2\,\eta_T)\,\Sigma(\delta_p,\sigma_l,\omega)\,. \tag{3.35}$$

The factor $\kappa = \sqrt{N}\,f_{\mathrm{B_s}}\,(1-2\,\eta_T)$ gives the normalization of the uncertainty distribution, without affecting the shape, and obviously the three parameters can not be disentangled. It is chosen to fix $f_{\mathrm{B_s}}$ and η_T to some "typical" values (namely $f_{\mathrm{B_s}} = 0.15$, $\eta_T = 0.25$), and adjust N to fit the uncertainties in the data. The effect of a different choice which yields the same κ value is investigated in Section 3.3.4.

The decay length and momentum resolution terms both affect the shape of the measured uncertainty as a function of ω. The sensitivity is not enough to get a reliable simultaneous determination of both. It is thus chosen to fix δ_p, again, to a "typical" value of $\delta_p = 0.15$, and tune the value of σ_l. This choice is preferred because, as explained later, δ_p plays an important rôle in the determination of the confidence level, and needs anyway to be varied over a wide range to check the stability of the result obtained.

Samples are generated at three starting points for σ_l, which are chosen to be $200\,\mu$m, $250\,\mu$m and $300\,\mu$m, each with 30000 events and the other parameters as described above.

For each value of σ_l, the number of events in the toy experiments is tuned by minimizing the sum of the differences with the data uncertainties,

$$\sum_i \left(\sigma_{\mathcal{A}}^{\mathrm{data}} - \sigma_{\mathcal{A}}^{\mathrm{toy}}\right)_i^2\,, \tag{3.36}$$

where the scaling low of Eq. 3.35 is used (Fig. 3.14a). The three minima found are then compared and interpolated with a parabolic fit (Fig. 3.14b) to find the absolute minimum, which turns out to be very close to $250\,\mu$m. The number of events needed at this point is 34000.

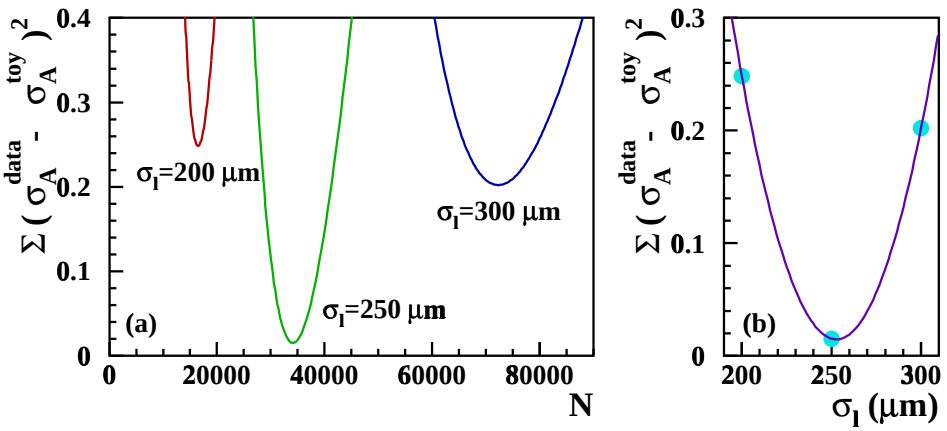

Figure 3.14: (a) Optimization of the number of events for three different values of the decay length resolution.
(b) Choice of the optimal decay length resolution.

Samples description

On the basis of the procedure described above, a set of parameters **S** is defined as follows:

- a single *purity* class: $15\%\,\mathrm{B_s}$ $38\%\,\mathrm{B_d}$ $38\%\,\mathrm{B^+}$ $9\%\,\Lambda_\mathrm{b}$;
- a single *tagging* class: mistag rate $\eta_T = 25\%$ for all species;
- a single *resolution* class:

$$\left.\begin{array}{l}\sigma_l = 250\,\mu\mathrm{m} \\ \sigma_p/p_0 = 0.15\end{array}\right\} \text{ both Gaussian with no tails;}$$

- Monte Carlo parametrized efficiency (taken from the analysis of Ref. [10]). The curve is shown in Fig. 3.15;
- b-hadron masses and lifetimes, and Δm_d from Ref. [13];
- $\langle x_b \rangle = 0.7$;
- Δm_s fixed at different values, according to the study considered;
- statistics of 34000 b-hadron decays.

A second set of parameters $\mathbf{S'}$ is defined to generate a second family of toy experiments. The momentum resolution is chosen to be $\sigma_p/p_0 = 0.07$, which is significantly better than what is typically achieved in inclusive analyses. In order to keep the agreement with the world average uncertainties on the measured amplitudes, the number of events is reduced to 29000 (obtained with the procedure described above). The other parameters are left unchanged. These experiments are used in the following to investigate the dependence of the confidence level upon the momentum resolution.

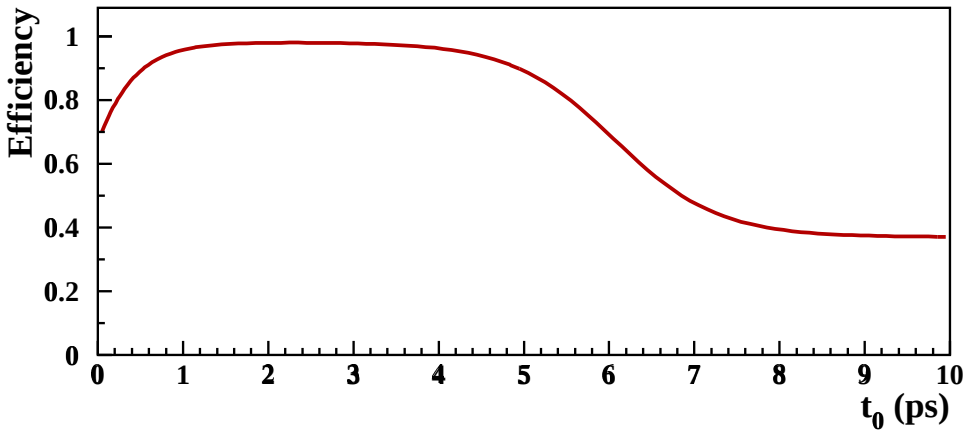

Figure 3.15: Shape of the reconstruction efficiency as a function of true proper time. The normalisation of the vertical scale is arbitrary.

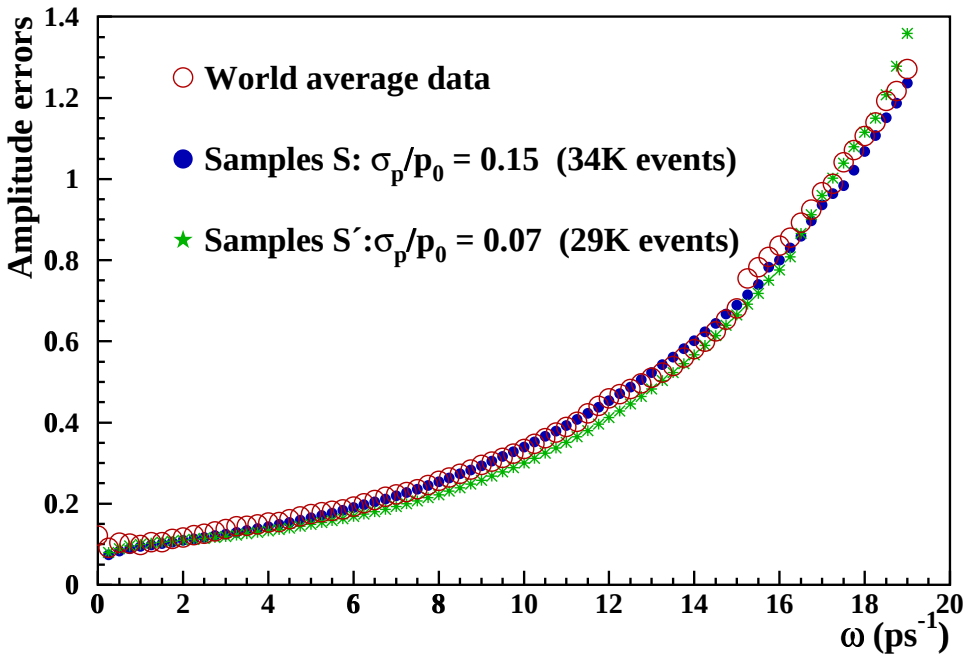

Figure 3.16: Amplitude uncertainties comparison: simulated experiments *versus* world average data.

The uncertainties on the amplitude, $\sigma_{\mathcal{A}}$, obtained with these two sets of experiments are compared to the uncertainties from the world combined data in Fig. 3.16. The step at $\omega = 15\,\mathrm{ps}^{-1}$ could be reproduced by averaging, for each "experiment", two "analyses", of which one has its scan stopped at that point. No attempt was made in this direction.

A third set of samples **S″** with $\delta_p = 0.15$, $\sigma_l = 200\,\mu$m and statistics of 16500 decays (which correspond to the optimization of Fig. 3.14) is used to investigate the dependence of

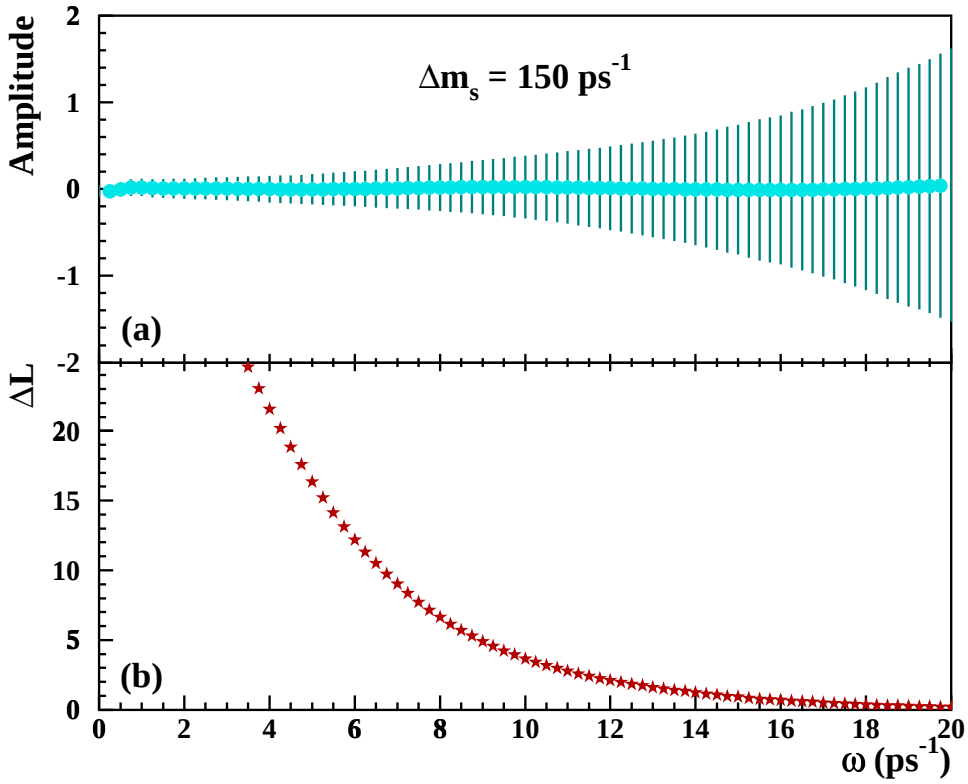

Figure 3.17: (a) Expected amplitude and uncertainty for samples of type **S**, with $\Delta m_{\rm s} = 15\,{\rm ps}^{-1}$, as a function of ω.
(b) Expected likelihood shape. The plots are obtained by averaging 2000 samples. Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu$m are assumed.

the confidence level upon the decay length resolution.

Finally samples of type **s** are defined from samples **S** by increasing $f_{\rm B_s}$ by a factor of five (*i.e.*, having $f_{\rm B_s} = 0.75$) and reducing the statistics by a factor of 25 (which gives 1360 decays).

In Fig. 3.17 the expected shape of the amplitude and the likelihood is shown, as obtained by averaging 1000 samples of type **S**, generated with $\Delta m_{\rm s} = 150\,{\rm ps}^{-1}$. The expected value is consistently zero, and the uncertainties on $\mathcal{A}$ are Gaussian, which confirms the validity of the amplitude method to set limits on the oscillation frequency.

3.3.4 The estimate of the significance

As demonstrated in Section 3.3.2, the probability that, at a given point in the frequency scan, a value of the likelihood $\Delta\mathcal{L} < \overline{\Delta\mathcal{L}}$ be found, can be calculated, given $\overline{\Delta\mathcal{L}}$, from the uncertainty on the measured amplitude, which is available from the data.

For the purpose of establishing the significance of the minimum, however, this probability is not enough, since what is needed is the probability that anywhere in the range explored

a configuration more unlikely than the one observed may appear (in the hypothesis of large Δm_s). This significance is driven not only by the uncertainties, but also by the correlations between the amplitude measurements at different frequencies, which are not controlled from the data, and might depend on the particular combination of parameters chosen for the simulation. It is therefore mandatory to identify the most relevant sources of systematic uncertainty which might affect the extraction of the confidence level. This point is investigated in what follows.

Correlations

From the discussion of Section 3.3.2, it turns out that the momentum resolution is the most critical parameter to determine the point–to–point correlation in the amplitude scan. In a sample with better momentum resolution, correlations are smaller and therefore the probability of having significant deviations from $\mathcal{A} = 0$ in a sample with no signal is larger, in a given frequency range explored.

In order to investigate the dependence of the point–to–point correlation upon the parameters used in the generation, a sensitive quantity is the average difference between amplitudes measured at two given points in the frequency scan. If there were no correlations, this difference could be written in terms of the uncertainties on the amplitude as

$$\langle \mid \mathcal{A}_i - \mathcal{A}_j \mid \rangle \;=\; \sqrt{\frac{2}{\pi}}\,\sqrt{{\sigma^i_{\mathcal{A}}}^2 + {\sigma^j_{\mathcal{A}}}^2}\,. \tag{3.37}$$

Correlations reduce this value if i and j are close enough. A scan in steps of $0.25\,\mathrm{ps}^{-1}$ is assumed, as for the data analyses.

For each of the four set of parameters, **S**, **S′**, **S″** and **s** defined in Section 3.3.3, 150 samples are produced, and the quantity $\langle \mid \mathcal{A}_i - \mathcal{A}_j \mid \rangle$ is calculated, for $i - j = 1,\ 4,\ 7,\ 10$. The results are shown in Fig. 3.18, where they are compared with the expectation for no point–to–point correlation.

Compared to the most "realistic" samples, **S**, the largest deviation is observed, as expected, when the momentum resolution is changed (samples **S′**). At low ω, the difference between the no-correlation limit (curve) and the values found in the simulation (markers), decreases rapidly as the distance between the points increases: for $i-j = 4$ ($\Leftrightarrow\ \Delta\omega = 1\,\mathrm{ps}^{-1}$) it is reduced by about a factor of two compared to $i - j = 1$, so $\Delta\omega = 1\,\mathrm{ps}^{-1}$ can be taken as an estimate of the "correlation length" at small frequencies. When ω increases, the difference between the curve and the simulation remains substantial even when the points are a few inverse pico-seconds apart, demonstrating the increase of the correlation length with ω.

Samples **S′** can be used to estimate a "systematic uncertainty" on the confidence level obtained, coming from the specific choice of the parameters used in the simulation.

The Confidence Level

The significance of the minimum observed in the $\Delta\mathcal{L}$ distribution (Fig. 3.8) is estimated by computing the probability that a structure as or more unlikely is observed in a sample with Δm_s far beyond the sensitivity.

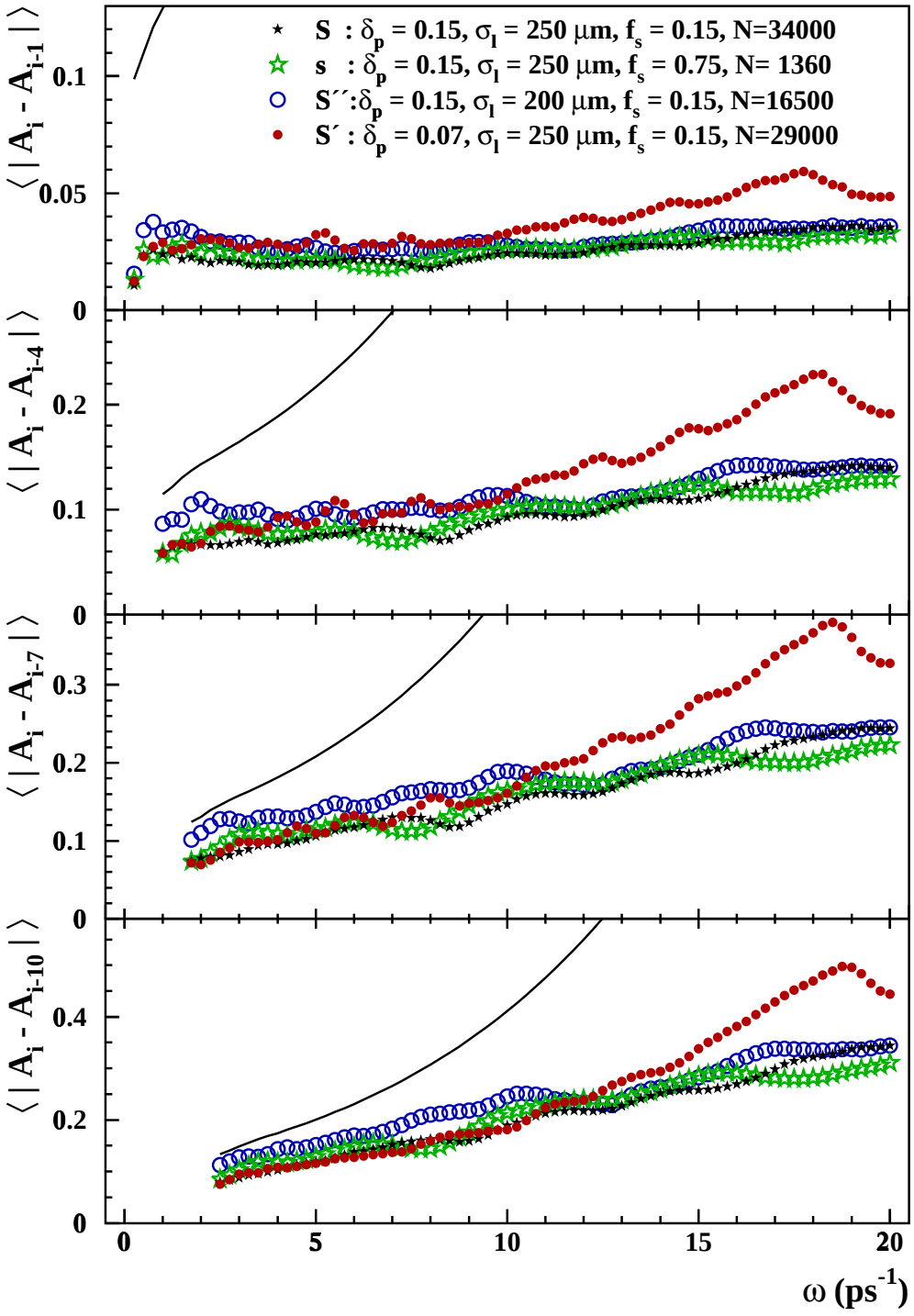

Figure 3.18: Point–to–point fluctuations for four sets of samples. From top to bottom, the average difference between points distant 1–4–7–10 steps in the amplitude scan are shown. The lines correspond to the limit of no correlation between points.

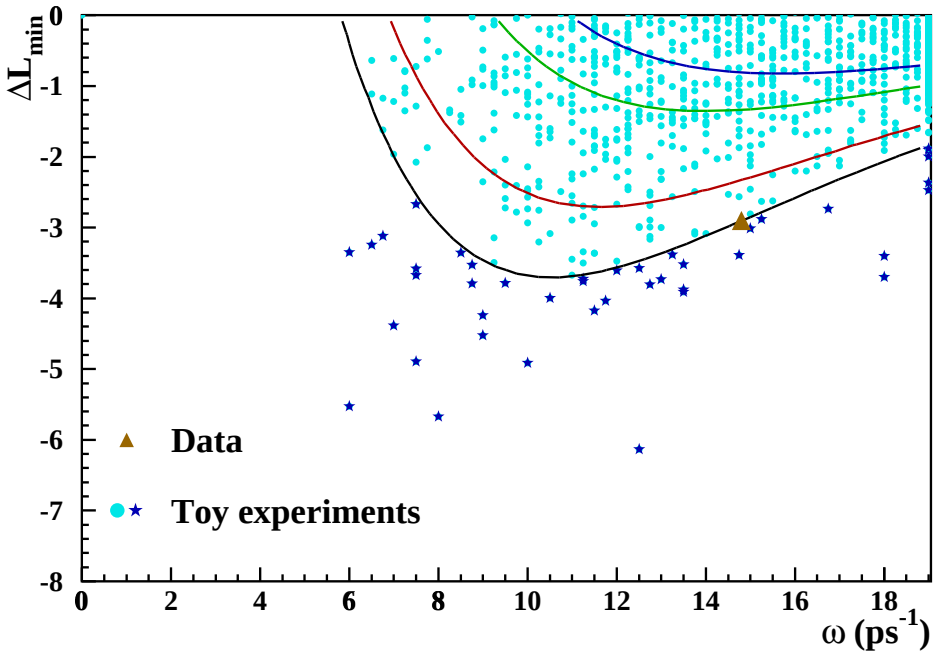

Figure 3.19: Minima of $\Delta\mathcal{L}$ for 2000 samples of type **S**, with $\Delta m_s = 150\,\mathrm{ps}^{-1}$. The curves represent contours of equal probability of observing a value of $\Delta\mathcal{L}$ smaller than $\overline{\Delta\mathcal{L}}$, as a function of ω (as in Fig. 3.13).

In order to do that, it is taken into account that the probability of observing a given value of $\Delta\mathcal{L}$ is a non–trivial function of ω. Probability contours in the $(\Delta\mathcal{L}, \omega)$ plane (as in Fig. 3.13b) are built from the data uncertainties. The contour corresponding to the data sample is computed. $N = 2000$ samples of type **S** with $\Delta m_s = 150\,\mathrm{ps}^{-1}$ are analysed and the number $N_{\mathrm{exp}}^{\mathrm{out}}$ of those that give a minimum $\Delta\mathcal{L}_{\mathrm{min}} < 0$ outside the contour corresponding to the data is recorded. Since the expected value of the likelihood is positive for all frequencies (see Fig. 3.17b), occasionally the minimum in the range $0 - 19\,\mathrm{ps}^{-1}$ is also positive. These minima are not counted, independently of the frequency at which they occur, since they can not be interpreted as a signal of oscillations.

The population of the toy experiments in the $(\Delta\mathcal{L}, \omega)$ plane along with the point corresponding to the data sample, is shown in Fig. 3.19.

The confidence level is computed as

$$1 - \mathrm{C.L.} \equiv \frac{N_{\mathrm{exp}}^{\mathrm{out}}}{N_{\mathrm{exp}}} = 0.021 \pm 0.003\,. \tag{3.38}$$

The study is repeated with 2000 samples of type **S′**, and yields

$$1 - \mathrm{C.L.} = 0.033 \pm 0.004\,. \tag{3.39}$$

This value has to be understood as a conservative estimate of the probability of statistical fluctuations, since it is obtained with experiments built to have lower point–to–point correlations than that expected for the average of real analyses. The distribution of the minima

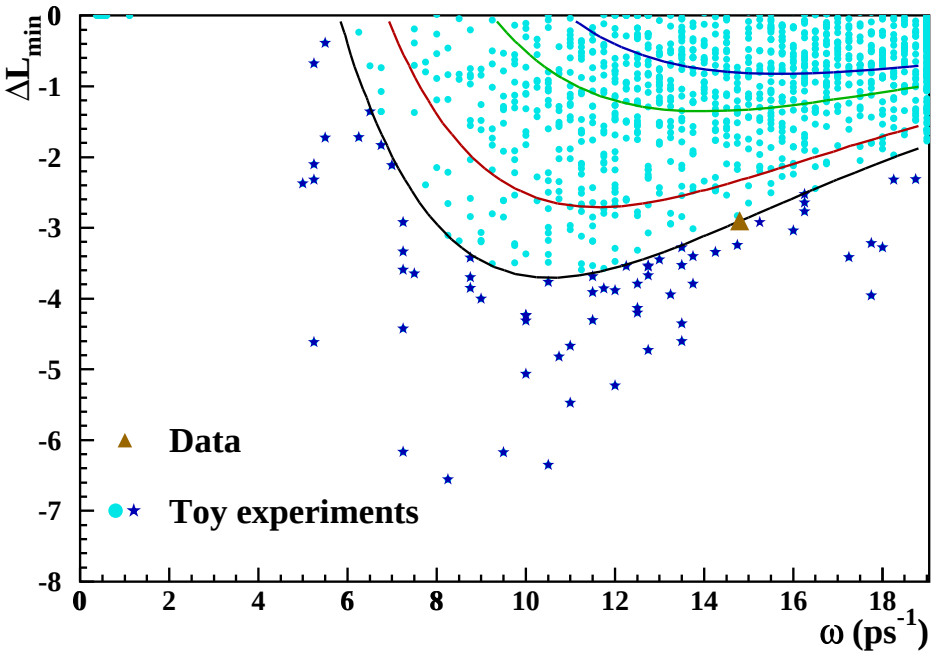

Figure 3.20: Minima of $\Delta\mathcal{L}$ for 2000 samples of type $\mathbf{S'}$ with $\Delta m_s = 150\,\mathrm{ps}^{-1}$. The same curves as in Fig 3.19 are shown.

for this case is shown in Fig. 3.20. The difference between the values of Eq. 3.38 and Eq. 3.39 gives an upper limit for the uncertainty coming from the lack of a detailed simulation.

The probability that the winter 1999 result of the world combination of B_s oscillation analyses was due to a statistical fluctuation can be therefore quantified to be around 3%. The uncertainty on this number coming from the inaccuracies of the simulation is below 1%.

Comparison with the oscillation hypothesis

In order to check that the amplitude spectrum observed in the data is in qualitative agreement with the hypothesis of oscillations, 500 samples of type **S** have been produced, with input frequency $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$. The expected amplitude and uncertainty at each frequency value are shown in Fig. 3.21, with the data points superimposed. The agreement is good over the whole frequency range.

A quantitative study of the compatibility of the data with the signal hypothesis would require to perform a fine scan on Δm_s with many samples at each value, in order to define a probability that the results observed are produced by an oscillation with a frequency in the range explored. This kind of study is not attempted here.

A simple check is performed instead. The 500 samples with the true oscillation at a value $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$ are analysed in terms of their incompatibility with the no–oscillation hypothesis. The scatter plot of the likelihood minima in the $(\overline{\Delta\mathcal{L}}, \omega)$ plane, as for the samples with $\Delta m_s = 150\,\mathrm{ps}^{-1}$, is presented in Fig 3.22.

An enhanced density in the region $14\,\mathrm{ps}^{-1} < \omega < 16\,\mathrm{ps}^{-1}$, $-3 < \Delta\mathcal{L}_{\mathrm{min}} < -1$ is shown

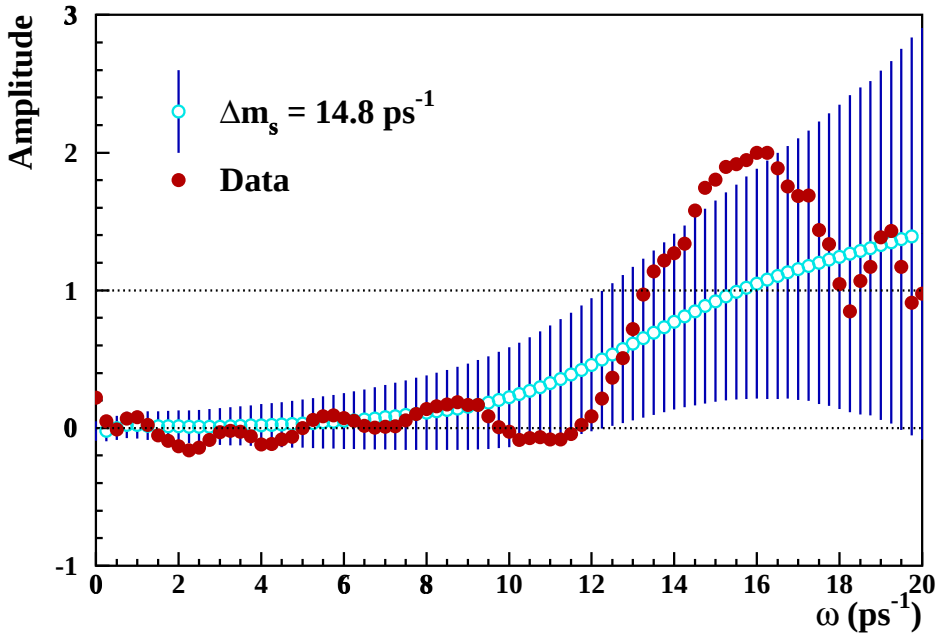

Figure 3.21: Average amplitude and expected uncertainty as a function of ω for a signal at $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$. The amplitude values, obtained by averaging 500 toy experiments, are in good agreement with the data measurements (solid points).

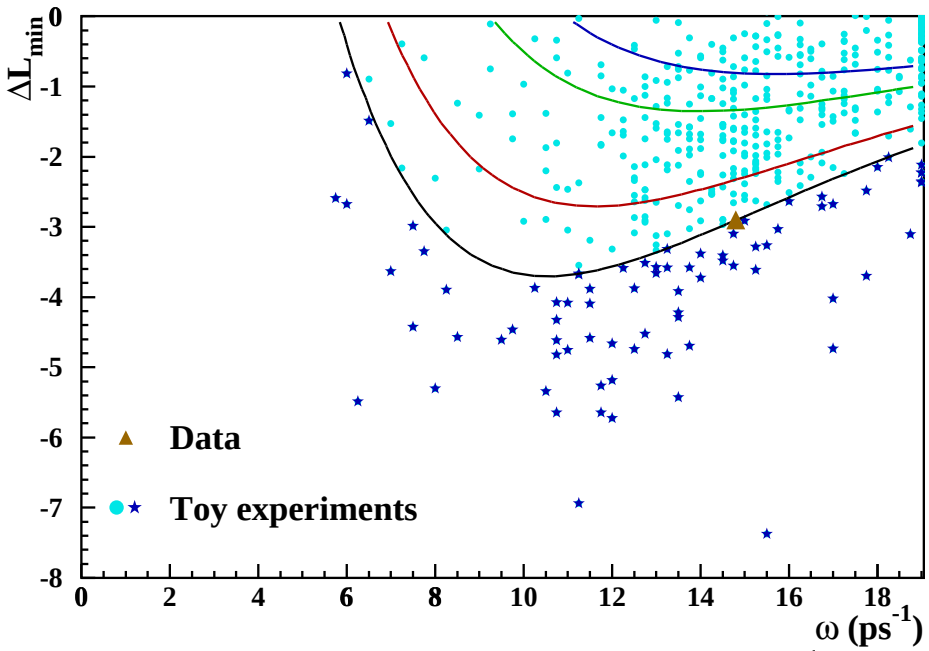

Figure 3.22: Minima of $\Delta\mathcal{L}$ for 500 samples of type **S**, with $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$. The same curves as in Fig. 3.19 are shown.

in the plot. A cluster of experiments with minima at $\omega = 19\,\mathrm{ps}^{-1}$ is also clearly visible: for these experiments the lowest point of the likelihood was at the boundary of the region analysed. Experiments with $\Delta\mathcal{L}_{\mathrm{min}} < -5$ appear at frequencies lower than the true one, where fluctuations which can produce deep minima are more likely.

Out of these 500 samples, 80 were found outside the estimator contour corresponding to the data, which gives a probability of 16%. If the data results were perfectly "typical" compared to the toy samples, the expected result would be 50%.

3.3.5 Sensitivity limiting factors

Different factors determine the performance of a B_s oscillations analysis and its relevance at high values of ω, where the present interest is. Toy experiments are very useful to simulate the impact of each analysis ingredient (such the number of events, the resolution on the momentum or on the decay length) on the final shape of the amplitude statistical uncertainty as a function of the test frequency ω. The toy experiments described in Section 3.3.3 are also used here.

For a B_s oscillations analysis with N selected events, a B_s fraction f_s, a total mistag probability η_T, and momentum and decay length resolutions σ_p/p and δ_l, the statistical uncertainty on the measured amplitude as a function of the test frequency ω is given by Eq. 3.27. This equation shows that the number of events, the global mistag probability, and the fraction of signal events in the sample contribute to the normalization of the uncertainty distribution without affecting its shape.

Proper time resolution

The proper time resolution affects the shape of the amplitude statistical uncertainty as a function of ω. In general, a better resolution corresponds to a smaller degradation of the precision on the amplitude at high frequency. The two components from which the proper time is estimated have a different impact on the shape of the uncertainty. Toy experiments can be used in this case to illustrate separately the effect of σ_l and σ_p/p. In Fig. 3.23 all uncertainty lines are obtained from the average of 20 samples with $N = 20000$, $f_s = 15\%$, $\eta_T = 25\%$, the resolutions are varied as follows.

- The reference sample, full line: $\sigma_l = 200\,\mu\mathrm{m}$, $\sigma_p/p = 15\%$.
- Varying momentum resolution, dash-dotted line: $\sigma_p/p = 7\%$.
- Varying momentum resolution, dash-dotted line: $\sigma_p/p = 4\%$.
- Varying decay length resolution, dashed line: $\sigma_l = 160\,\mu\mathrm{m}$.
- Varying decay length resolution, dashed line: $\sigma_l = 250\,\mu\mathrm{m}$.

For values of the frequency smaller than $\sim 1.5\,\mathrm{ps}^{-1}$, all the curves coincide: the proper time resolution does almost not affect the uncertainty curve in that region. However, as the frequency increases, the effect of varying one of the two resolution components becomes visible. An improvement of the momentum resolution leads to a flatter curve at mid-range

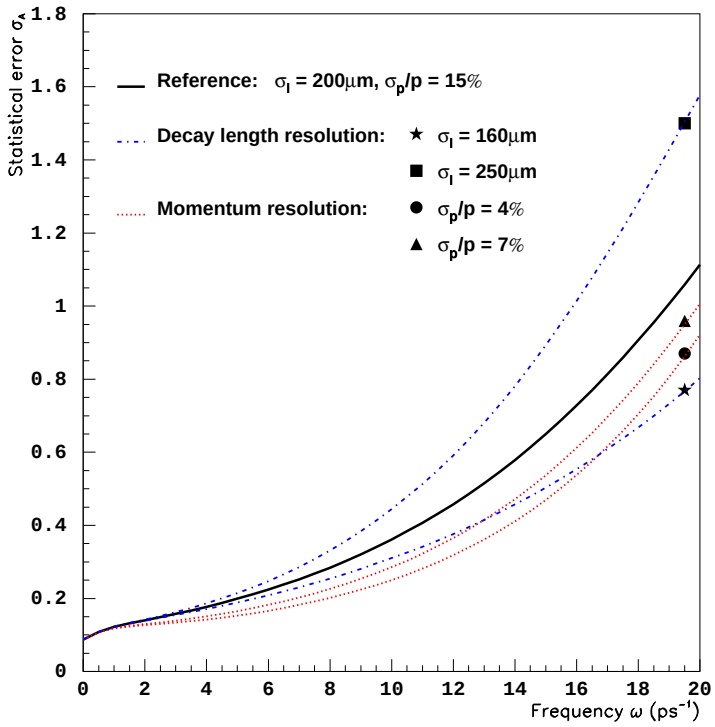

Figure 3.23: Proper time resolution effect on the amplitude uncertainty. The decay length and momentum resolutions are varied in turn to illustrate their separate impact on the shape of the amplitude uncertainty as a function of ω.

frequencies $\sim 1.5 - 12\,\mathrm{ps}^{-1}$, but the uncertainty curve remains as steep as the reference one, at high frequency. A change of the decay length resolution results in a change of the slope of the uncertainty curve for all values of the frequency ω, this effect is particularly important in the region of interest, at high frequencies. The different impact observed for σ_l and σ_p/p, shows the crucial rôle of the decay length resolution in a B_s oscillation analysis.

Event-by-event decay length uncertainty

In all toy experiments presented up to here, the same decay length resolution is taken for the smearing of all events in the generated sample and this resolution is then used in the fitting function. However, in a real analysis not all events have their secondary vertex reconstructed with the same uncertainty. It is important to use an event-by-event estimate of the uncertainty and not the average of the sample. To illustrate this last point, 20 samples of 20000 events with different decay length resolution values (18% at 80 μm, 35% at 150 μm, 27% at 210 μm, 10% at 350 μm, 10% at 1.1 mm) were generated. Two different fits were performed; in the first case, each event was fit using the decay length resolution it had been generated with, in the second, all events were fit with the average resolution of the whole sample ($\sim 270\,\mu$m). The other parameters σ_p/p, f_s, and η_T were chosen as in the reference sample used in the previous Section.

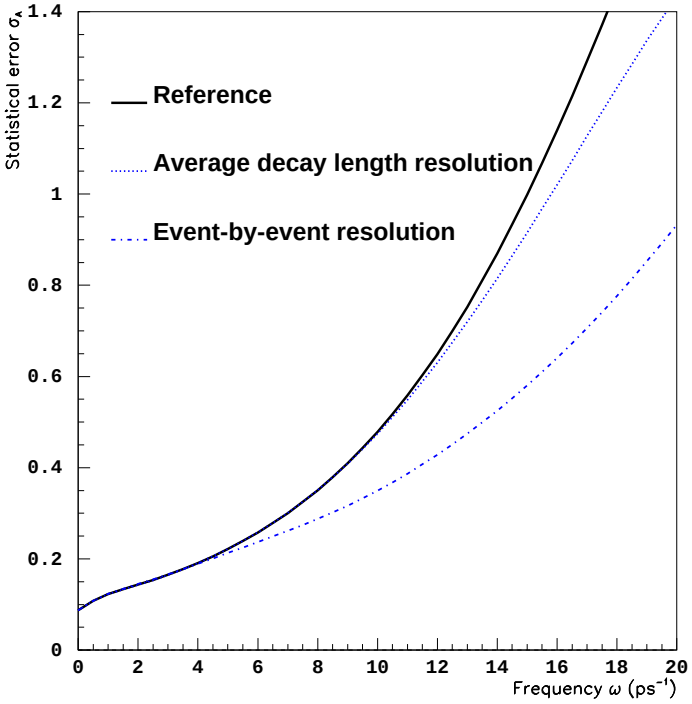

Figure 3.24: Effect of the decay length resolution treatment on the amplitude uncertainty. The use of event-by-event uncertainties is compared to the use of the average uncertainty.

The result is shown in Fig. 3.24. The dash-dotted line corresponds to the fit where each event was treated with its own decay length resolution. The same fitting function was used for the reference line and for the dotted line, with a single resolution value of 270 μm. The difference between these last two uncertainty curves (full and dotted) is due to the presence of some events with very small decay length resolution in the second sample. The fit is sensitive to these events, even if they are not treated at their best. However, the big difference is appreciated between the dotted and the dash-dotted lines. The event samples are exactly the same: the difference illustrates that an adequate event-by-event characterization increases significantly the potential of a B_s oscillation analysis.

This little exercise shows how much can be gained just by using event-by-event uncertainties, or dividing the sample in resolution classes at least. This point is also illustrated with the real data sample used for the inclusive semileptonic analysis presented in this thesis, in Chapter 8.

Chapter 4

Experimental Setup

The B_s oscillation analyses presented in this thesis were performed with the data sample collected by the ALEPH experiment between 1991 and 1995, during the phase 1 of LEP at CERN. A short reminder of LEP characteristics and achievements, in particular in its first running phase, are given to start this Chapter in Section 4.1. The ALEPH detector, one of the four multipurpose detectors of LEP, took data between August 1989 and November 2000. The detector and its performance for physics analyses are described in Section 4.2.

4.1 The LEP Storage Ring

The LEP [62] (Large Electron Positron Collider) storage ring is an e^+e^- collider located at CERN (European Laboratory for Nuclear Research). The LEP project was proposed in 1975, approved six years later, and started its operation in 1989. The main goal of the machine was to perform precision measurements of electroweak observables, in order to test the Standard Model at the level of radiative corrections. Such corrections take contributions from off-shell particles, and are therefore sensitive to physics beyond the production threshold. The LEP accelerator also provided the possibility to directly search for heavy, yet unobserved particles in a range of energies substantially higher than previous machines.

The accelerator programme consisted of two phases. In the first phase, LEP stored, accelerated, and collided electrons and positrons at a centre-of-mass energy around the Z boson mass (91.2 GeV), with a typical instantaneous luminosity of $10^{31}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Several million of Z bosons were produced and their decays studied (as in the analyses described in this thesis). Since 1995, the centre-of-mass energy was increased every year. In 2000, the last year of running, a centre-of-mass energy of 209 GeV was reached, shortly before the machine was shutdown for the last time. The LEP dismantling started in November 2000.

The LEP ring is located underground, inside a nearly horizontal tunnel of 26.67 Km of circumference, at a depth which varies between 50 m and 150 m, spanning the French and Swiss territories (Fig. 4.1). The plane of the tunnel is inclined by 1.4% to ensure that all underground caverns and the main part of the tunnel are located in solid rock while, at the same time, the maximum depth of the shafts is limited to less than 150 m. The LEP accelerator consists of eight arcs alternated with eight straight sections. The beams are

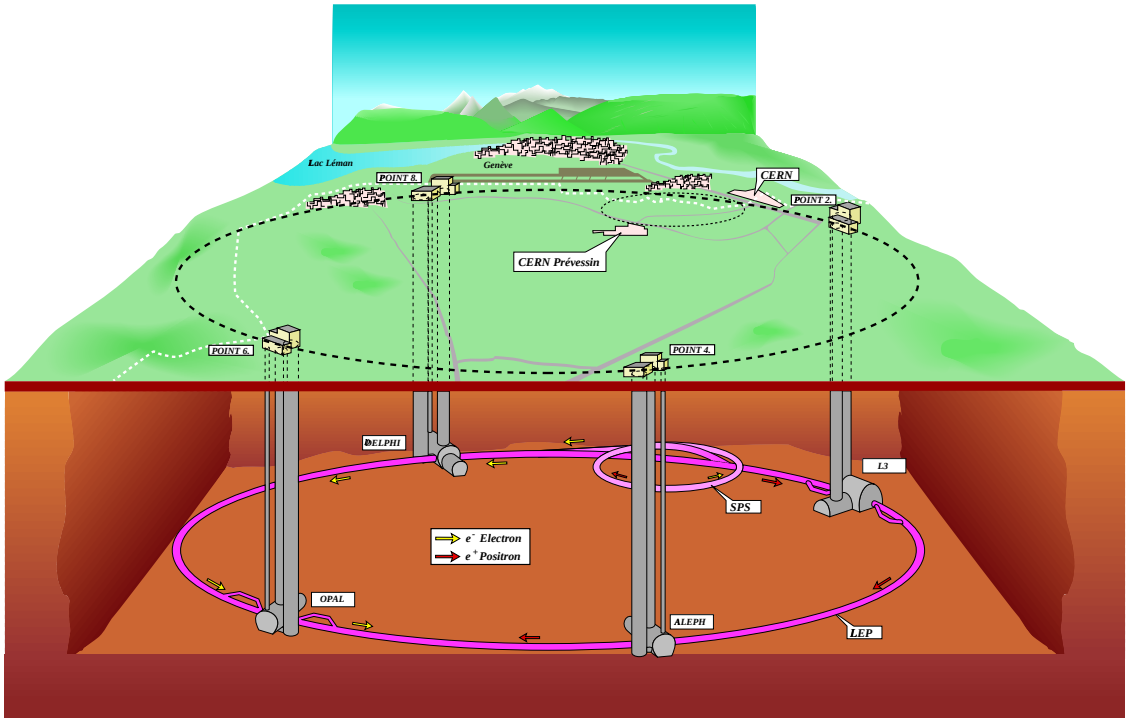

Figure 4.1: Schematic view of LEP and the Surrounding Area.

formed by bunches of electrons and positrons which circulate inside the beam pipe at nearly the speed of light for many hours (typically 10 hours at the Z, 1h30 at 207 GeV, and 15 minutes at 209 GeV). The beams are accelerated in opposite directions and cross in eight points, although they are steered to collide only on the four *Interaction Points* (I.P.). The straight sections, about 500 m long, are located around these eight points. The four LEP experiments, ALEPH, DELPHI, L3, and OPAL are sited in the middle of the interaction regions (Figs. 4.1 and 4.2).

The LEP injection chain can be seen in Fig. 4.2: it is formed by five accelerators, each of which handles the same electrons and positrons. Electrons are generated by the electron gun. A fraction is used to create positrons by collision with a tungsten target. Two linear accelerators (LINAC) of 200 MeV and 600 MeV are used to accelerate the lepton beams. Following the LINAC's, the particles are injected into a small circular electron positron accumulator, the EPA. They are then injected into the PS (proton synchrotron) accelerator, where the energy is increased up to 3.5 GeV. The PS injects the beams into the Super Proton Synchrotron (SPS), which operates as a 20 GeV electron-positron injector for LEP (since 1997, this energy has been increased up to 22 GeV). These last two machines (the PS and SPS), are operating at CERN since 1959 and 1976 respectively. Finally electron and positron beams are accelerated in LEP to reach the nominal energy.

The acceleration in LEP is provided by radio-frequency (RF) cavities and the magnetic

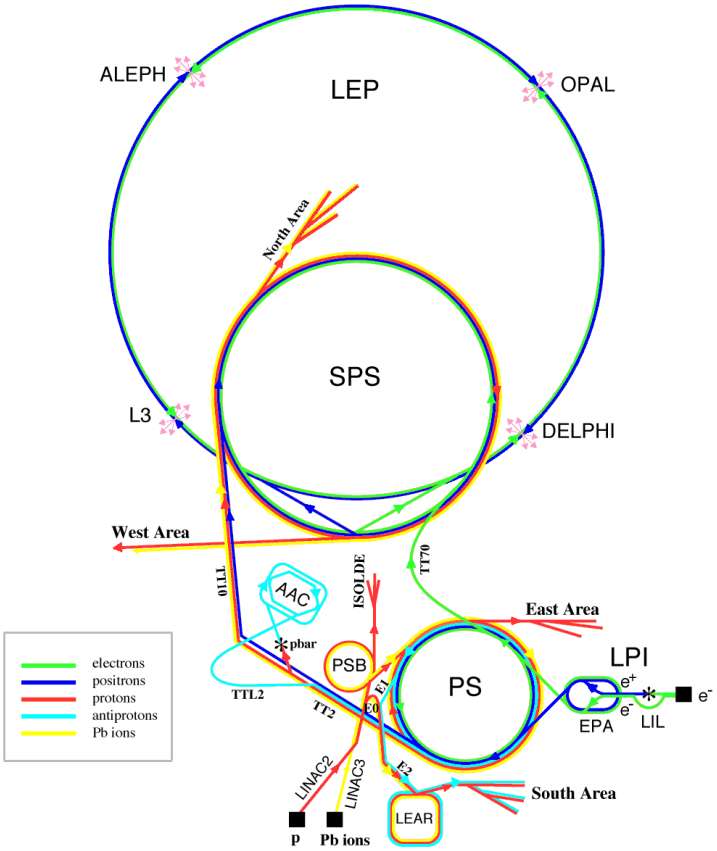

Figure 4.2: CERN Accelerators, LEP injection chain. The accelerators are not drawn at scale.

field. The RF system provides the energy needed to compensate for the synchrotron radiation loss in the curved trajectories, while the magnetic field is used to increase the energy from the initial $\sim 20\,\mathrm{GeV}$ to the nominal energy.

The performance of an accelerator is expressed with the centre-of-mass energy at which the collisions are made, and with the *luminosity* achieved. The luminosity is a quantity which relates the interaction rate for a given process to its cross-section:

$$\frac{dn}{dt} = \sigma \mathcal{L} \tag{4.1}$$

where dn/dt is the production rate of events according to the corresponding cross-section σ for a luminosity $\mathcal{L}$. The *luminosity* is defined by equation 4.1.

The luminosity provided by a collider can be expressed and computed in terms of its parameters. Therefore, if these parameters are known, the expected luminosity can be computed. The corresponding expression in the case of a circular collider like LEP is [63]

$$\mathcal{L} = \frac{k_b \; I_{\mathrm{e}^+} I_{\mathrm{e}^-}}{4\pi \; e^2 \; f_{\mathrm{rev}} \; \sigma_x^* \; \sigma_y^*} \; , \tag{4.2}$$

where k_b is the number of bunches per beam, I_{e^+}, I_{e^-} are the electron and positron beam currents, f_{rev} is the revolution frequency of the beam collisions, and σ_x^*, σ_y^* are the widths

of the Gaussian distributions representing the r.m.s. beam sizes at the collision point (it is assumed that the current intensity can be expressed as, $I = Nef_{\mathrm{rev}}$, e being the electric charge, and N the number of electrons). Typical values for the instantaneous LEP luminosity which can be obtained with this expression are of the order of $10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, in agreement with the actual LEP performance at its first phase.

The total luminosity delivered by LEP to the four experiments in each of the 12 years of running is shown in Fig. 4.3. The integrated luminosity delivered at the Z pole from

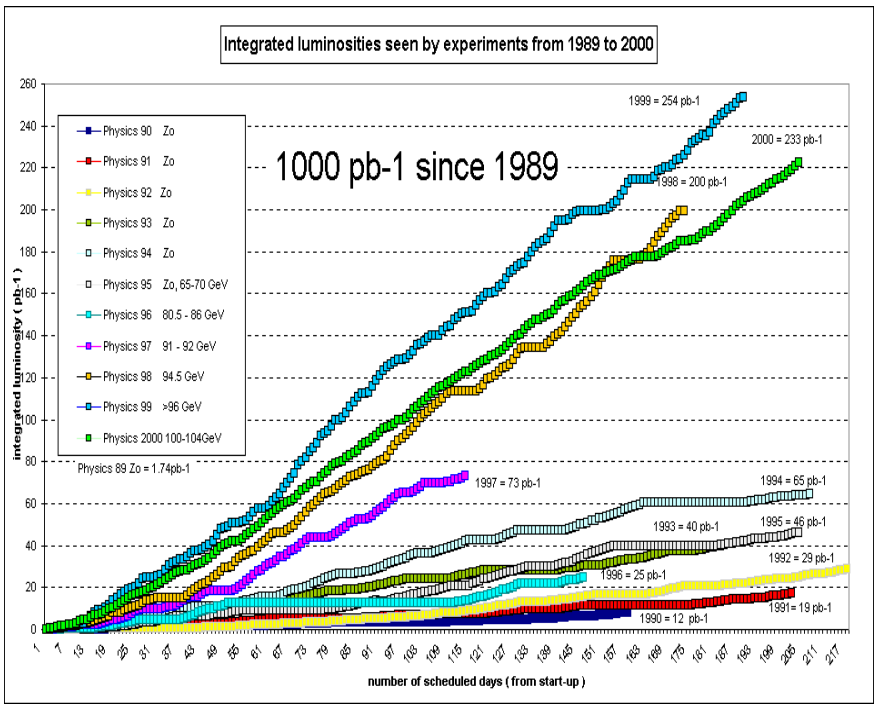

Figure 4.3: LEP Integrated luminosity from 1989 to 2000.

1991 to 1995 is $\sim 190\,\mathrm{pb}^{-1}$. During the second phase of LEP, some luminosity was also delivered at the Z pole, this data was mostly foreseen for detector calibration purposes. It is not included in the data sample analysed for the present thesis. Some parts of the ALEPH detector (mainly VDET, see Section 4.2.1) were upgraded for the second LEP phase, and therefore the analysis of the additional Z data requires specific studies. This fact, together with the relatively small fractional increase in statistics ($\sim 10\%$), discouraged the analysis of the additional Z decays.

4.2 The ALEPH Detector and Performance

The ALEPH (Apparatus for LEP Physics) detector, is one of the four experiments at the LEP collider of CERN. It is located at the interaction point number four in a cavern at 143 m under the surface. The approximated shape of the apparatus is a 12 m diameter by 12 m length cylinder positioned around the LEP beam pipe. Its axis coincides with the beam axis

and it is parallel to the magnetic field. It covers as much of the 4π solid angle as possible. It is designed to measure the momenta of charged particles, to measure the energy deposited in calorimeters by charged and neutral particles, to identify the three lepton flavours, and to measure the distance of travel of short-lived particles such as the tau lepton and the b and c hadrons. Particular emphasis has been given to momentum resolution up to the highest energies (by means of a large tracking system in a magnetic field), to electron identification (by means of a highly segmented, projective electromagnetic calorimeter, as well as ionization measurement in the tracking system), and to muon identification (with continuous tracking inside sufficient iron absorber to eliminate the hadrons).

In the ALEPH reference system, the z-direction is tangent to the beam pipe, positive in the direction followed by the electron beam. The positive x-direction points to the centre of the LEP ring and the positive y-direction is orthogonal to z and x (pointing to the top). A sketch of the detector with all the subdetectors shown is displayed in Fig. 4.4.

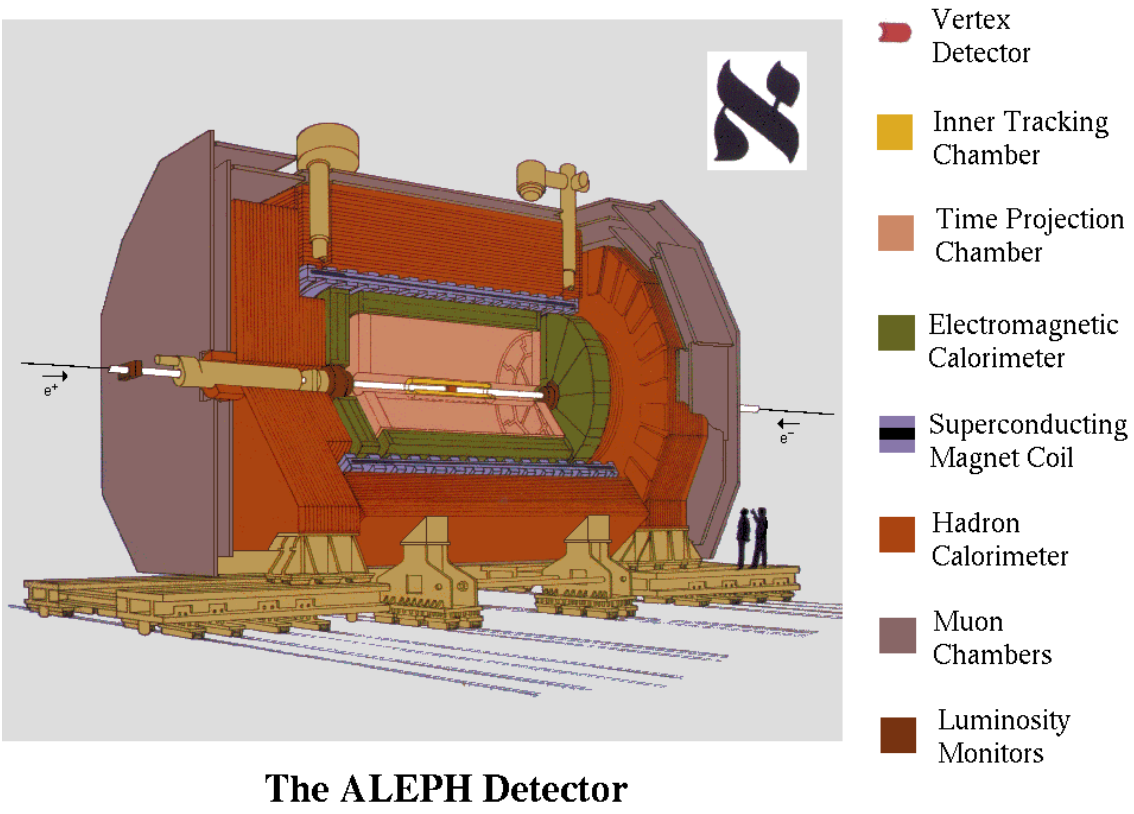

Figure 4.4: Schematic view of the ALEPH Detector, with all sub-detectors.

The ALEPH detector is described in details in Refs. [64] and [65], and its performance in Ref. [66]. Here only a brief discussion is presented, with particular emphasis on the aspects most relevant for the analysis described in this thesis.

4.2.1 The tracking system

The tracking system of a particle detector aims at reconstructing charged particles trajectories inside the system (tracks). It is a non-destructive system, which most of the particles traverse with almost no variation on their momentum. In some cases (if immersed in a magnetic field), it is also capable of separating positively from negatively charged particles, and to some extent, to provide particle identification.

The ALEPH tracking system involves three detectors. From the inner to outer part those are, the vertex detector (VDET), the inner tracking chamber (ITC) and the time projection chamber (TPC). They are immersed on a 1.5 T axial magnetic field created by a super conducting solenoid. In an axial and uniform magnetic field, B, charged particles have helix trajectories with a curvature, ρ, proportional to their transverse momentum, p_T, as

$$p_T \;=\; q\rho \mathrm{B}\;, \tag{4.3}$$

where q is the electric charge of the particle. The electric charge sign of reconstructed particles is given by the rotation sense.

The vertex detector

The ALEPH silicon vertex detector (VDET) was installed in 1990 and was fully operational in 1991. It provides high-precision measurements of charged particles trajectories close to the interaction point, without which the analyses presented in this thesis would not be possible.

The configuration of the VDET is shown in Fig. 4.5. Two layers of silicon strip detectors

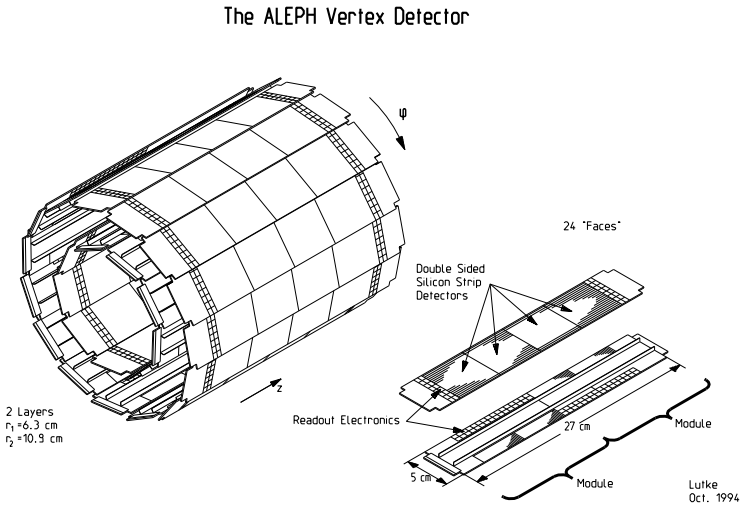

Figure 4.5: Configuration of the vertex detector.

are arranged in two concentric barrels around the beam pipe, with an average radius of $r_1 \approx 6.3\,\mathrm{cm}$ for the inner layer, and $r_2 \approx 10.9\,\mathrm{cm}$ for the outer layer. The strips on one side are parallel to the beam direction and measure the azimuthal angle ϕ. The strips on the other side are perpendicular to the beam and measure the z-coordinate. The radius, r, of the

individual wafers is given by the mechanical holding frame. The position of a single particle hit is then determined in cylindrical coordinates, $(r\,,\phi\,,z)$.

The vertex detector hits are used to improve the precision of the track reconstruction obtained with the outer tracking. The overall tracking performance is given at the end of this Section. The effective point resolution achieved by the vertex detector alone in both directions, $r\,\phi$ and $r\,z$, is $\sim 12\,\mu$m. The VDET solid angle coverage is 87% and 70% for the inner and the outer layers respectively.

The first-generation ALEPH vertex detector pioneered the use of double-sided silicon microstrip technology. In autumn 1995 it was replaced by an upgraded longer version which provides better solid-angle coverage and more radiation resistance. All the data analysed for this thesis were obtained before the vertex detector upgrade.

The Inner Tracking Chamber

The inner tracking chamber (ITC) is a cylindrical multiwire drift chamber which provides up to eight accurate $r\,\phi$ points for tracking in the radial region between 16 cm and 26 cm and $|\cos\theta| < 0.99$. As a fast response device, with a reaction time within $1\,\mu$s$-3\,\mu$s, it also provides the only tracking information used for the first-level trigger.

The wires run parallel to the beam direction in a gas mixture made of 80% argon and 20% carbon dioxide at atmospheric pressure. The $r\,\phi$-coordinate is obtained from the drift time measurement. A precision of about $150\,\mu$m is achieved. The z-coordinate is found with the measurement of the difference in the arrival times of pulses at the two ends of each sense wire with a coarse precision of about 5 cm.

The Time Projection Chamber

The time projection chamber (TPC) is the most important element of the ALEPH tracking system. It was chosen for its good momentum and angular resolution capabilities, as well as its ability for particle identification through the $\mathrm{d}E/\mathrm{d}x$ (energy loss by ionization in matter) measurements of charged particles. The chamber has a cylindrical structure, as shown in Fig. 4.6. It is 4.7 m long, with 0.31 m inner radius and 1.8 m outer radius. The chamber is divided in two halves by a high voltage central membrane which produces the longitudinal electric drift field. It is filled by a mixture of argon (91%) and methane (9%) at atmospheric pressure. The closing end-plates are covered by wire chambers organized in "sectors". Charged particles, when traversing the chamber, produce electrons by ionization of the gas. These electrons drift towards one of the closing end-caps, where they are collected. Their arrival position and time are measured by the wire chambers and pads. The z-coordinate is obtained from the drift time and the known drift velocity. The ϕ-coordinate is calculated by the interpolation of the signals induced on cathode pads located on the sectors and the r-coordinate is given by the radial position for the pads involved in the measurement. The chamber measures 21 three-dimensional points for each charge particle traversing the inner and outer field cages.

The azimuthal coordinate resolution obtained with leptonic Z decays is $\sigma_{r-\phi} = 173\,\mu$m. The longitudinal resolution is $\sigma_z = 740\,\mu$m for charged particles close to normal incidence,

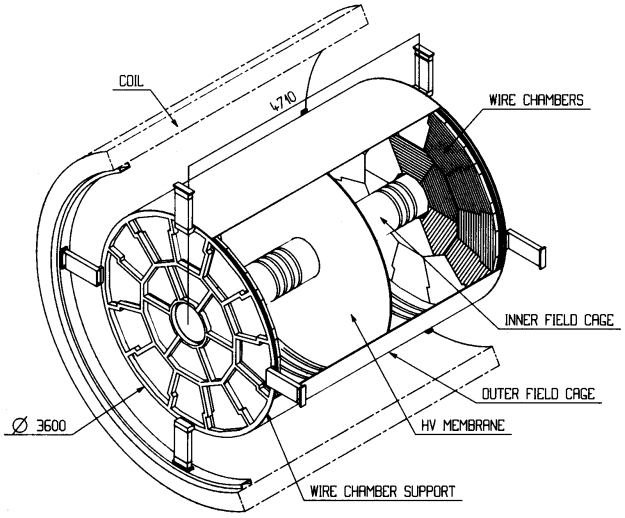

Figure 4.6: Time projection chamber overall view.

$80^\circ < \theta < 100^\circ$.

The ALEPH data from the LEP1 phase were reprocessed in 1998. The reason for this reprocessing was to implement improvements on the track reconstruction [67] and to include new information on $\mathrm{d}E/\mathrm{d}x$ measurement from the TPC. Only the wires measurements used to be available, but the less accurate pads measurements are now used as well. A complete description of the new $\mathrm{d}E/\mathrm{d}x$ combination algorithm and performance can be found in an ALEPH internal note [68].

Tracking performance

All three detectors are used for the track reconstruction in ALEPH. The reconstruction algorithm uses the TPC information first. Close hits (measured three-dimensional points) are linked together to from track segments, and the segments are connected to make tracks by requiring consistency with a helix hypothesis. These track candidates are extrapolated to the inner detectors, ITC and VDET, where consistent hits are assigned. The final track fit takes into account multiple scattering between subsequent points.

The overall performance of the momentum measurement for charged particles with full tracking information is shown in the first column of Table 4.1. The performance obtained with the TPC alone is presented first, followed by the effect of the other two devices. The TPC momentum resolution is improved by a factor two when the measurement is complemented with that of the other tracking devices.

The tracking system is also used for track impact parameter measurement (defined and used in Sections 5.5 and 5.7.1). The impact parameter resolution obtained on data for high momentum charged particles is also shown in Table 4.1 (second and third columns). The vertex detector is essential for the accuracy on the impact parameter measurement. When it

Detector	$\sigma(1/p_T)(\mathrm{GeV}/c)^{-1}$	$r\phi$ (μm)	rz (μm)
TPC	1.2×10^{-3}	310	808
+ ITC	0.8×10^{-3}	107	808
+ VDET	0.6×10^{-3}	23	28

Table 4.1: Momentum (first column), and impact parameter resolution performance. The successive rows show the improvement obtained when the detectors are successively included in the track fit.

is used, the resolution is reduced by more than an order of magnitude.

The ALEPH tracking system has achieved very good performance, without need of too much material which would degrade the final energy measurement. In Fig. 4.7 the thickness of the materials traversed by particles is expressed as a percentage of a radiation length, summed over the various materials traversed. It has a strong dependence on the polar angle θ, especially at small θ where the incidence into the material is oblique (because the material is often arranged in cylindrical layers).

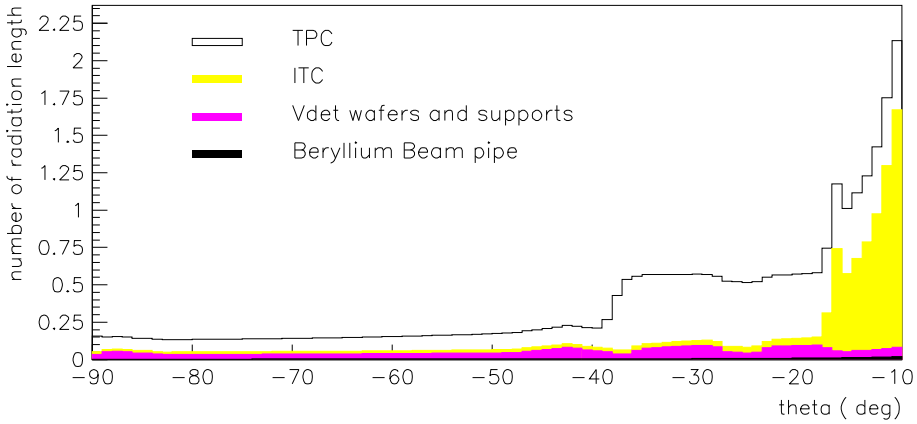

Figure 4.7: Material before the electromagnetic calorimeter measured in number of radiation length.

4.2.2 Calorimetry

Calorimeters are devices which measure the energy of particles by stopping them in a dense material. Characteristic interactions with matter (*e.g.*, atomic excitation, ionization) are used to generate a detectable effect, via charged particles, which is then translated to an energy measurement. Different types of particles have different penetration power in matter. In particular electrons and photons are easily absorbed, unlike hadrons. An electromagnetic calorimeter is meant to be able to stop electrons and photons, while hadrons typically traverse it leaving only a fraction of their initial energy. It is therefore common, like in ALEPH, to build a hadron calorimeter outside the electromagnetic calorimeter. Muons, as minimum ionizing particles, have an even higher penetration power than hadrons and are not stopped after the two calorimeters; the muon chamber detectors are located at the most external

part of ALEPH to detect their presence. Neutrinos almost do not interact with matter, the ALEPH detector is blind to them.

The Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is located between the TPC and the magnet. It is a lead/wire-chamber sampling calorimeter, which measures the energy and position of electromagnetic showers from electrons and photons. Its nominal thickness is 22 radiation lengths (X_0).

The geometrical structure of the detector is approximately cylindrical. The calorimeter consists of a barrel part, closed by flat end-caps. This structure and positioning with respect to the coil and the hadron calorimeter is shown in Fig. 4.8. The barrel is 4.8 m long, and

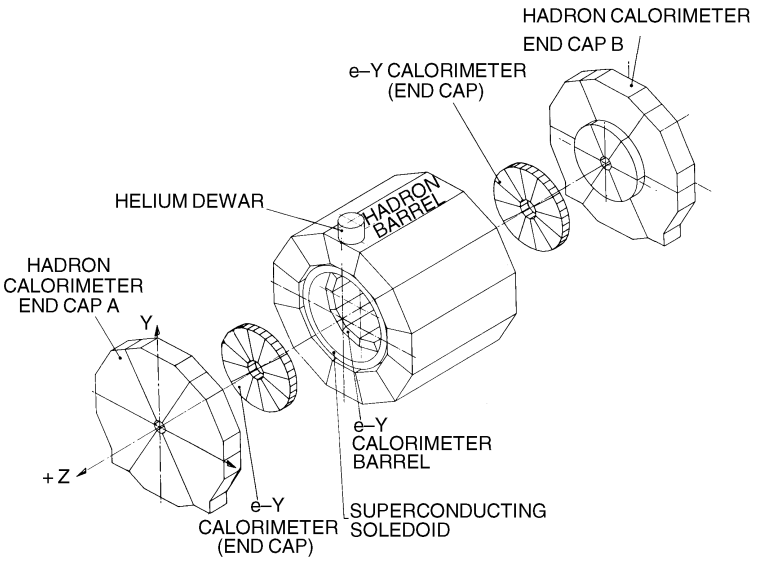

Figure 4.8: Overall geometry of the hadron calorimeter surrounding the superconducting coil and the electromagnetic calorimeter.

its inner radius is 1.54 m, while the outer is 2.53 m. Both end-caps and the barrel consist of 12 modules each. Each module is subdivided in towers pointing towards the centre of the detector. Each of the towers is read in three different depth modules called "storeys". The ensemble of storeys at the same depth is called a "stack". A total of 73728 towers provide a fine granularity of $0.9^\circ \times 0.9^\circ$ (in the barrel), which is very useful for $e/\gamma/\pi^0$ identification, energy, and angular measurement. The entire calorimeter is rotated by -32.7 mrad with respect to the hadron calorimeter to avoid overlapping of uninstrumented separation between modules (cracks). The polar angle coverage is $|\cos\theta_{\rm track}| < 0.98$.

The energy resolution for isolated electrons and photons, determined by comparing the measured energy to the track momentum, is fit to be

$$\frac{\sigma(E)}{E} = \frac{0.18}{\sqrt{E/\mathrm{GeV}}} + 0.009 \; . \tag{4.4}$$

The position of a shower is calculated by an energy-weighted mean of the polar and azimuthal angle of the individual storeys in the cluster. The spatial resolution is evaluated to be

$$\sigma_{\theta,\,\phi} = \left(\frac{2.5}{\sqrt{E/\mathrm{GeV}}} + 0.25 \right) \text{ mrad} . \tag{4.5}$$

The hadron calorimeter and the muon chambers

The hadron calorimeter is located inside a large iron structure which is the main mechanical support of ALEPH and which also serves as a return yoke for the magnetic flux. The iron plays the rôle of the passive material in a sampling calorimeter, while the active part is provided by layers of plastic streamer tubes. The HCAL is used as a muon filter as well as a powerful muon identification tool. It also measures the energy of neutral and charged hadrons which are not stopped in the electromagnetic calorimeter.

As seen in Fig. 4.8, the hadron calorimeter is subdivided in a central barrel and two closing end-caps. The total thickness of the iron is 1.2 m, which amounts to 7.16 interaction lengths for a hadron traversing the calorimeter perpendicular to the slabs.

The energy resolution for pions at normal incidence, when both ECAL and HCAL are used, is found to be:

$$\frac{\sigma(E)}{E} \approx \frac{0.85}{\sqrt{E/\mathrm{GeV}}} . \tag{4.6}$$

Part of the hadron calorimeter readout is used for muon identification (see Section 5.3.2). In addition, two layers of muon chambers are placed outside the hadron calorimeter. Each of these chambers is made of two layers of streamer tubes which are not used here for energy measurement, but as tracking devices for the detection of particles leaking out the hadron calorimeter. The two layers are separated from each other by a sufficient distance so that not only the position but also the angle of a charged particle can be measured. Up to two three-dimensional coordinates are then obtained for each charged particle traversing a muon chamber.

4.2.3 Energy flow

The energy flow algorithm [66] combines the energy and particle identification measurements from the calorimeters and the tracking system. The overall energy measurement is therefore significantly improved.

As a first step, "good" tracks (with a similar selection as the one described in Section 5.1) are identified, and extrapolated to the two calorimeters. The energy deposits around the extrapolated tracks are associated to them. The calorimetric objects are then formed by these charged particles and calorimetric deposits or energy deposits in the ECAL associated with energy deposits in the HCAL. The calorimetric objects are then treated in the following manner:

- Tracks identified as electrons, with their associated deposits in the ECAL are eliminated from the calorimetric objects list. If the difference between the calorimetric energy

and the track momentum is greater than 3 times the ECAL resolution, this energy difference is taken as coming from a Bremsstrahlung photon, and counted as a neutral electromagnetic object.

- Tracks identified as muons are eliminated from the list as well, together with the energy of the associated calorimetric cluster, with a maximum energy of 1 GeV in the ECAL and 400 MeV per HCAL layer in the extrapolated 'road'.

- Photons are eliminated from the calorimetric objects list as well, and are counted as neutral electromagnetic objects.

- Only charged and neutral hadrons remain in the calorimetric objects list. The charged particles and their associated calorimetric objects are eliminated from the list. The remaining calorimetric energy is considered as coming from neutral hadronic particles.

4.2.4 Luminosity measurement

The luminosity received by the ALEPH detector is determined from the reconstruction of small angle elastic e^+e^- scattering (the so-called Bhabha scattering) in a well defined fiducial region. The Bhabha scattering at low angle is a purely QED process and its cross-section can be calculated with high precision. To a good approximation it amounts to,

$$\frac{d\sigma_0}{d\Omega} \simeq \frac{4\alpha^2}{E_b^2\,\theta^4}\ , \tag{4.7}$$

where α is the fine structure constant, E_b is the beam energy, and θ is the polar scattering angle of the electron or positron with respect to the beam direction.

Three lumonisity detectors were in use by the end of the experiment life. The ALEPH luminosity measurement was performed up to 1991 by a specific electromagnetic calorimeter called LCAL. The detector structure is very similar to that of the ECAL. The LCAL is located around the beam pipe on both sides of the interaction region with the first sampling layer at a distance of 2.63 m from the I.P. The detector covers, in the horizontal plane, a polar angle range of $45\,\text{mrad} < \theta < 160\,\text{mrad}$. With this detector the accuracy obtained on the luminosity measurement was of, at best, 0.4 %, which was found not to be enough for the needs of some physics analyses (in particular the Z line-shape studies [69]). It was therefore decided to build and install a new more precise luminosity detector: SiCAL.

The silicon luminosity calorimeter (SiCAL) [70], installed in 1992, consists of two homogeneous cylindrical silicon-tungsten sampling electromagnetic calorimeters which surround the beam pipe at approximately $z = \pm 2.5\,\text{m}$ on either side of the interaction point. The active volume of the detector subtends the polar angular region, $24.3\,\text{mrad} < \theta < 57.5\,\text{mrad}$. The aim of SiCAL was to reduce the total uncertainty in the luminosity measurement down to 0.1%. The SiCAL polar angular coverage is closer to the beam pipe than that of LCAL, to improve the statistical precision. A greater accuracy, with respect to LCAL, in the determination of the position of the detector components allowed the fiducial cross-section estimation to be improved, and therefore the systematic uncertainty on the luminosity measurement was reduced.

For the LEP2 phase, SiCAL was not used as the luminosity monitor anymore. The photon synchrotron radiation having increased dramatically compared to LEP1, new "masks" were installed close to the beam pipe to protect the central sub-detectors. These elements were added on the region where the SiCAL modules were positioned and their shadow reduced the fiducial region which could be used for the Bhabha rate counting below the needs for providing a measurement accurate enough.

The third luminosity detector, BCAL, is used for an online monitoring of the instantaneous luminosity. The two previous detectors could not be used for this purpose because their event rates are too low for a reliable and quick online estimate of the luminosity. The BCAL monitor consists of four modules (sampling calorimeter made of tungsten, scintillator and silicon) located very close to the beam pipe and more than 7 m from the interaction point, to detect Bhabha events at very small angle. The first version of the detector, which was in use till 1996, had a polar angle coverage of $5.1\,\text{mrad} < \theta < 9.0\,\text{mrad}$. The second version installed in 1997 was able to cover angles down to $\approx 4.5\,\text{mrad}$. The uncertainty on the online luminosity was about 5% (every two minutes) for a luminosity of 10^{31} $\text{cm}^{-2}\text{s}^{-1}$.

Chapter 5

General analysis tools

Along the 12 years of the ALEPH data taking, many analysis tools have been developed and improved, to serve general purposes, such as lepton identification and vertex finding. In this Chapter, those analysis tools used in the work presented in this thesis are described. In the first Section, the criteria to select events with a Z-boson decaying into a $q\overline{q}$ pair are explained. The three jet clustering algorithms used in this thesis are introduced in Section 5.2. In Section 5.3, the electron and muon identification procedures are described. The primary vertex finding method is explained in Section 5.4, and the procedure to estimate the track impact parameter is discussed in Section 5.5. Although the final secondary vertex algorithm used in the inclusive semileptonic analysis in Chapter 6 is an original work of mine (described later in Section 6.3), it uses, as a first step, the result of a standard inclusive secondary vertex finder. Two such vertex finders are available in ALEPH, both of them are described in Section 5.6. The last Section of this Chapter is dedicated to the general tools for the selection of events which contain b-hadron decays. The method used in the thesis for b-tagging was especially developed for it, however, the three b-tagging tools described in Section 5.7 are used, together with other information, as explained later in Section 6.4.

5.1 Hadronic selection

Hadronic events, *i.e.*, originating from $Z \rightarrow q\overline{q}$ decays, are selected using charged particle information alone. At least five charged particle tracks must be reconstructed by the TPC for an event to be selected. These tracks need to satisfy the following requirements to be called "good tracks". *i*) Four or more TPC three-dimensional points must be used for the track helix fit; *ii*) the track must originate from a cylinder centred in the beam position, with radius 2 cm and length 10 cm; and *iii*) the track polar angle with the beam axis must satisfy $|\cos\theta| < 0.95$, to ensure that six or more pad rows in the TPC are hit.

A cut on the total measured charged energy is applied in order to remove events for which the hadronic final state is less likely to come from a Z-boson decay than from the interaction of two photons radiated from the incident e^+e^-. The sum of the momenta of all good tracks is required to exceed 10% of the centre-of-mass energy. This cut also removes efficiently events produced by beam-gas interactions. The total efficiency of this selection is 97.48% [69], the

losses originating mainly from events at low polar angle. This selection is especially interesting because its systematic uncertainties are well under control and only account for 0.087%; the dominant background for this selection (which is anyway negligible once combined with other selection criteria like it is done in Chapter 6) comes from $e^+e^- \to \tau^+\tau^-$ and two photon interactions with a hadronic final state. Another good feature of the selection is that it is independent of the quark flavour to better than 0.1%.

The *thrust* axis of an event is defined as a vector $\vec{n}$ such that the quantity T defined as,

$$T \equiv \frac{\sum_{i=1}^{N} |\vec{p_i} \cdot \vec{n}|}{\sum_{i=1}^{N} |\vec{p_i}|} , \tag{5.1}$$

is maximal, where, $\vec{p_i}$ is the momentum of the i-th particle and N is the total number of reconstructed particles (charged and neutral) in the event (see Section 4.2.3). This direction aims at the reconstruction of the direction of the quarks from the Z decay. However, in the case of a hard gluon radiation, $Z \to q\overline{q}g$, this intuitive meaning is lost.

The thrust axis is therefore used to divide the event into two hemispheres with respect to the plane perpendicular to $\vec{n}$ and including the interaction point. With this definition, each hemisphere contains the fragmentation and hadronization products of each of the quarks from the Z decay (except in some rare cases of $Z \to q\overline{q}g$ decay) .

5.2 Jet algorithms

At high energy (approximately above 20 GeV or so), quarks and gluons are produced with a boost large enough to form "jets" of particles from their hadronization.

Three different jet clustering algorithms are used for the analyses described in this thesis. They are described in the following subsections.

5.2.1 The JADE algorithm

The JADE algorithm is described in Ref. [71]. In each event, charged and neutral particles are clustered in jets with the following prescription. Two particles are grouped together if the distance y_{ij} defined as

$$y_{ij} = \frac{2E_iE_j\,(1-\cos\theta_{ij})}{E_{\rm vis}^2} , \tag{5.2}$$

is smaller than a predefined cut value called the jet-resolution parameter, $y_{\rm cut}$. In Eq. 5.2, E_i and E_j are the energies of particles i and j, θ_{ij} is the angle between the momenta of these particles, and $E_{\rm vis}$ is an estimate of the total visible energy in the event. The numerator in Eq. 5.2 is equal to m_{ij}^2, the invariant mass of the system, if particles i and j are considered massless. The particle pair with the smallest value of y_{ij} is kept, the two original particles are replaced by a new pseudo-track, the energy and momentum of which is obtained from

the sum of energy and momenta of the two parent particles. The same procedure is repeated until no pair which fulfils $y_{ij} \leq y_{\rm cut}$ is found. The pseudo-tracks obtained at the last step are the particle jets.

5.2.2 The Nucleated jet algorithm

In the Nucleated jet algorithm [72], the jet is built around an *a priori* track seed (the choice of this charged particle depends on the data analysis to be performed). The quadri-momenta of charged and neutral particles close to the chosen seed are added until the mass of the jet reaches a cut value $m_{\rm cut}$, which has to be optimized depending on the type of jet to be reconstructed (from a b- or c-hadron, for instance). The criterion for a "close" particle is the same as in the JADE algorithm above. In the analysis described in Chapter 6, this algorithm is used to reconstruct the charmed object momentum. In order to reject tracks not originating from the b-hadron decay, particles with momentum smaller than $500\,\mathrm{MeV}/c$ are excluded from the the clustering process.

5.2.3 A cone algorithm: BTCONE

In a cone algorithm, particles are grouped together according to their angular proximity. A track is taken as a seed (the seed choice depends on the data analysis to be performed), charged and neutral particles in a predefined angular acceptance are grouped with the seed to form the jet. This approach was first developed at LEP by the OPAL Collaboration in Ref. [73]. This kind of algorithm is particularly useful for the reconstruction of long-lived hadron jets. The purities in terms of b-hadron (or c-hadron) decay tracks and measured b-hadron (or c-hadron) energy of the subsequent jet are improved with respect to that obtained with JADE based algorithms (see Ref. [74]). The best angular resolution for the b-hadron flight direction is achieved with this algorithm.

5.3 Lepton identification

In this thesis, a *lepton* refers to an electron or a muon. Their typical decay length being much larger than the dimensions of the detector (∞ for electrons, 300 Km for muons), they are reconstructed as charged particles in the tracking detectors, and identified from their typical interaction in all subdetectors (TPC, ECAL, HCAL, and muon chambers) as described in Section 4.2 and in Ref. [66].

All tracks considered by the hadronic selection described in the previous Section are taken as possible lepton candidates. As a first step of the lepton identification, some additional track-quality cuts are applied. *i*) At least five TPC hits are required; *ii*) the distance of closest approach to the beam axis must be smaller than 5 mm; and *iii*) and the z-coordinate of the track at the point of closest approach must be, in absolute, smaller than 5 cm.

In the following Sections the electron and muon specific selection criteria and performance are given.

5.3.1 Electron identification

The electron identification makes use of the ionization measurement in the TPC and the shape of the electromagnetic shower developed in the ECAL. These basic measurements are expressed in terms of normally distributed estimators on which cuts are applied to select electron candidates. The redundancy of the TPC and the ECAL information allows the performance of each estimator to be evaluated directly with the data.

Electron candidates are first required to have a momentum greater than $2\,\mathrm{GeV}/c$. The electron hypothesis is tested in the TPC and ECAL with a procedure as follows.

Electron identification in the TPC

For each track in the TPC, the $\mathrm{d}E/\mathrm{d}x$ is measured from the wires and from the pads. The measurement from the wires is more precise, but it is reliable only if at least fifty wire hits are associated with the track. The pads measurement is always available, it is combined with the wire estimate [68]. The combined $\mathrm{d}E/\mathrm{d}x$ estimator of the measured ionization in the TPC is normalized to 1.0 for a minimum-ionizing particle.

The $\mathrm{d}E/\mathrm{d}x$ estimation of each track is compared with its value expected in the electron-mass hypothesis, $\langle \mathrm{d}E/\mathrm{d}x \rangle_\mathrm{e}$. The final estimator is given by R_I

$$R_\mathrm{I} = \frac{\mathrm{d}E/\mathrm{d}x - \langle \mathrm{d}E/\mathrm{d}x \rangle_\mathrm{e}}{\sigma_{\mathrm{d}E/\mathrm{d}x}} \,, \tag{5.3}$$

where $\sigma_{\mathrm{d}E/\mathrm{d}x}$ is the one standard deviation resolution on the ionization measurement for the electron mass hypothesis. For a pure sample of electrons, the distribution of R_I is very close to a Gaussian of mean zero and width one. In this thesis, the cut applied on the R_I estimator to select electrons is $R_\mathrm{I} > -2.0$. No cut is applied to the positive R_I tail because only electrons are found in that region.

Electron identification in the ECAL

Both the compactness of the electromagnetic energy deposition around the particle direction reconstructed in the TPC, and the longitudinal shape of the shower, distinguish electrons from other particles, and are exploited for their identification.

All good tracks are extrapolated from the end of the TPC through the ECAL. A crossing point is computed at each of the three ECAL stacks.

The shower transverse development in the ECAL is measured with the estimator R_T. This estimator is defined from the track momentum, p, and the energy deposited, E_4, in the four towers closest to the extrapolated track, E_4

$$R_\mathrm{T} = \frac{E_4/p - \langle E_4/p \rangle}{\sigma_{E_4/p}} \,, \tag{5.4}$$

where $\langle E_4/p \rangle$ is the expected value of E_4/p for an electron, and $\sigma_{E_4/p}$ is the expected resolution (one standard deviation). The R_T estimator is most efficient for hadron rejection when the momentum of the track is high.

The longitudinal shower development is expressed with the estimator $R_{\rm L}$. The inverse of the first moment of the longitudinal energy distribution is defined by

$$\chi_{\rm L} \;=\; \frac{E_4}{\sum\limits_{i=j}^{4}\sum\limits_{j=1}^{3} E_i^j S_j} \;, \tag{5.5}$$

where E_i^j is the energy deposited in the i-th storey of the j-th stack, and S_j is the mean depth of the energy deposits in the j-th stack. The quantity $\chi_{\rm L}$ is independent of the angle of the incoming particle and it is computed using the energy measured in the three stacks with an iterative procedure assuming, at each step, that the longitudinal deposition induced by an electromagnetic shower is described by the standard shape [75]. This procedure converges quickly for electrons and it is essentially divergent for hadrons. The longitudinal estimator reads

$$R_{\rm L} \;=\; \frac{\chi_{\rm L} - \langle\chi_{\rm L}\rangle}{\sigma_{\chi_{\rm L}}} \;, \tag{5.6}$$

and has a normal Gaussian distribution for electrons.

The distribution of $R_{\rm T}$ versus $R_{\rm L}$ for a sample of tracks enriched in photon conversions is shown in Fig. 5.1. The electron and hadron contributions are clearly separated, the accu-

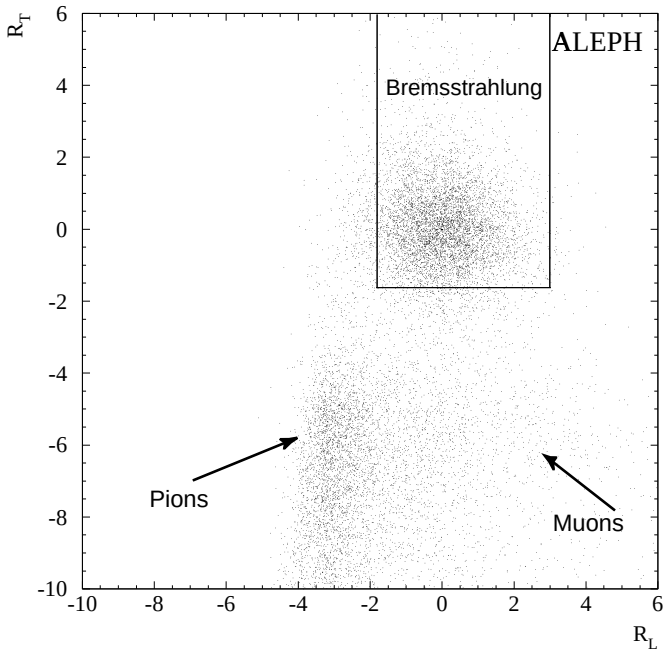

Figure 5.1: The ECAL estimators for a sample of electrons from photon conversions in the data.

mulation of events around $R_{\rm T} = 0$ and $R_{\rm L} = 0$ is due to electrons. The following cuts are optimized to maximize hadron rejection with a limited efficiency loss:

$$\begin{aligned} -1.8 \;&<\; R_{\rm L} < 3.0 \;, \\ -1.6 \;&<\; R_{\rm T} \;. \end{aligned} \tag{5.7}$$

No upper cut is applied on the R_{T} estimator because the four central towers associated to an electron can contain additional energy from a Bremsstrahlung photon, thus increasing the measured value of R_{T}. The region delimited by the cuts in R_{L} and R_{T} is shown in Fig. 5.1.

Electron identification performance

The simultaneous use of the three estimators described above provides an electron identification efficiency in jets independent of the transverse momentum of the electron candidate with respect to its jet. A dependence on the track momentum is observed, the efficiency decreases with momentum. The electron identification efficiency, computed after the momentum and R_{T}, R_{L} cuts, is 98.9% with a purity of 94.4%. The most important contamination to the electron candidate sample comes from pions, which account for $\sim 90\%$ of it.

5.3.2 Muon identification

Muons are identified by the tracking capabilities of the hadron calorimeter, together with the muon chamber information. The TPC $\mathrm{d}E/\mathrm{d}x$ estimator can be used as well to further reject misidentified charged kaons.

To identify muons in the HCAL, a 'road' is opened around the track extrapolated, with a width of three times the estimated extrapolation uncertainty due to multiple scattering in the calorimeter material. The number of hits in each of the 23 HCAL planes of streamer tubes is counted. In each of these planes, a muon is expected to leave one hit if the 'road' intersects it within an active region, the plane is said to have fired if a digital hit lies within the multiple scattering 'road'. The number of expected hits is then compared to the observed number. Tracks are considered for muon identification in hadronic events if their momentum is greater than $2.5\,\mathrm{GeV}/c$, which ensures that all muons exit the muon chambers, and that the identification efficiency does not vary significantly with momentum.

The cuts used to define a muon are

- $N_{\mathrm{fir}}/N_{\mathrm{exp}} \geq 0.4$;
- $N_{\mathrm{exp}} \geq 10$;
- $N_{10} > 4$;

where N_{exp}, N_{fir}, and N_{10} are, respectively, the number of planes expected to be hit by a muon, the number of planes actually hit, and the number of planes hit within the last expected ten. These cuts are well suited for isolated muons. In order to enhance the rejection power against hadrons in jets, the typical features of the digital pattern created by a hadron shower in the HCAL are used. A variable, X_{mult}, is computed by counting all the HCAL digital hits in the last eleven planes within a wider 'road' around the extrapolated track. The result is divided by the number of fired planes so that X_{mult} represents the average hit multiplicity per fired plane. For muon identification it is requested to be smaller than 1.5.

In Fig. 5.2, the distributions of $N_{\mathrm{fir}}/N_{\mathrm{exp}}$, N_{10}, and X_{mult} for muons from $\mathrm{Z} \to \mu^+\mu^-$ events are compared to those of pions produced in τ decays.

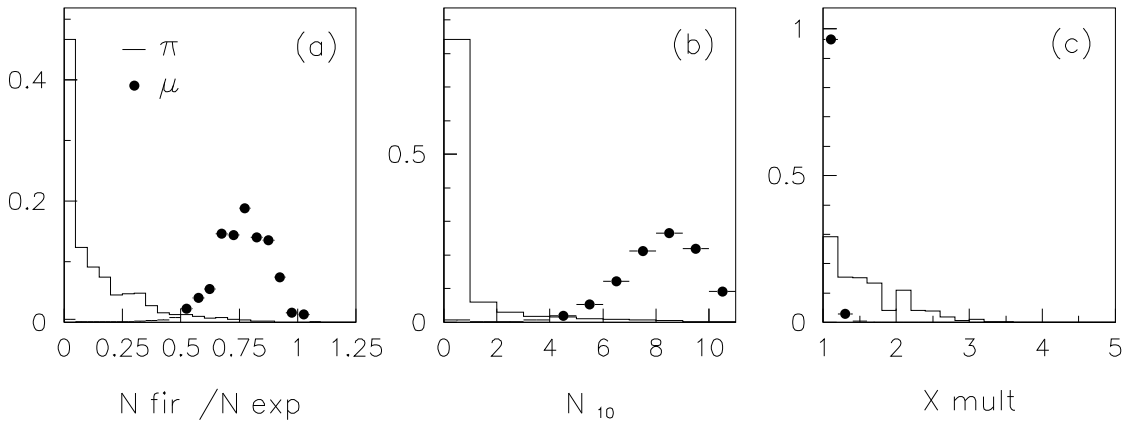

Figure 5.2: Distribution of (a) $N_{\rm fir}/N_{\rm exp}$, (b) N_{10}, and (c) $X_{\rm mult}$, for muons (points) and pions (solid line), with $N_{\rm exp} > 10$. The plots have been normalized to equal areas, and the vertical scale is arbitrary.

A track is defined to have hit a muon chamber if at least one of the two tube planes yields a space point whose distance from the extrapolated track is less than four times the estimated standard deviation from multiple scattering.

The requirement for the association of a muon chamber hit in addition to the HCAL muon criteria is very powerful for background rejection. The simulation predicts that, while about 94% of the muons identified in HCAL are associated to a muon chamber hit, only about 20% of the misidentified hadrons have this association.

A rejection cut against charged kaons which do not decay in the TPC can be applied to the TPC $\mathrm{d}E/\mathrm{d}x$ estimator with the muon-mass hypothesis, $R_{\rm I}(\mu)$,

$$R_{\rm I}(\mu) = \frac{\mathrm{d}E/\mathrm{d}x - \langle \mathrm{d}E/\mathrm{d}x \rangle_\mu}{\sigma_{\mathrm{d}E/\mathrm{d}x}} > -2\,. \tag{5.8}$$

In the ALEPH track reconstruction, kinks are identified as tracks for which a double helix fit with a common kink vertex point is better than a single helix fit. Those muon candidates which present a kink in the TPC are rejected. The purity on prompt muons is increased by 1%, with a negligible loss in efficiency.

The muon identification efficiency above $2.5\,\mathrm{GeV}/c$ is higher than 90% for a purity between 80% and 85%, with a mild dependence on the momentum. The most important contribution to the contamination comes from pions.

5.4 Primary vertex finding

The first step of the reconstruction of an event topology is the determination of the interaction point of the incoming $\mathrm{e}^+\mathrm{e}^-$, the primary vertex. Some analyses, like those described in this thesis, need a three-dimensional position measurement with an accuracy better than $100\,\mu$m.

The performance achieved by the ALEPH tracking can be used to compute the primary vertex event by event using the reconstructed charged particles, and therefore improve on the precision of the LEP beam spot.

It is important that the algorithm chosen for the primary vertex determination has similar performance for all type of events. Some of the particles produced in a Z-boson decay, those which contain c- and b-quarks in particular, travel some distance from the interaction point through the detector volume before decaying in a heavy-hadron decay point. These particles are known as long-lived. In the presence of such particles, not all tracks in the event come from the interaction point, and therefore care has to be taken to avoid a bias in its reconstruction.

The algorithm used in ALEPH [76] removes the lifetime information of tracks by projecting them on the plane perpendicular to the jet direction as illustrated in Fig. 5.3. The

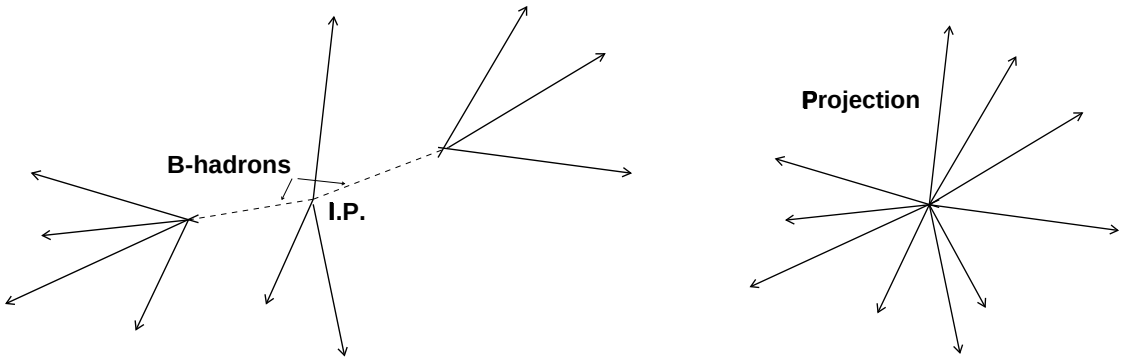

Figure 5.3: Schematic of a $Z \to b\overline{b}$ event on the left and its projection for primary vertex finding on the right

jet direction is chosen because it is a good approximation of the parent hadron direction. All tracks are projected onto the plane perpendicular to the direction of the nearest jet. The primary vertex is found in each of these planes, for each jet. The combination of the different projections of the primary vertex in each jet of an event is used to obtain the three-dimensional position of this vertex. The information from the LEP beam spot in the x-y direction is used as a constraint. A procedure to identify tracks which fit badly with the primary vertex is defined, and these tracks rejected from its determination. A correction to take into account the finite resolution on the jet direction is applied. As a final step, tracks with no lifetime are identified, and their information along the jet direction is used in the final primary vertex fit.

The algorithm is based on a χ^2 minimization, and at each of the step above, the uncertainty estimations are checked to give correct pull distributions. The core of the method relies on the first position estimation (before track rejection and other refinements) with the following χ^2 minimization with respect to V_i

$$\chi^2 = \sum_{i=1}^{3} \frac{(B_i - V_i)^2}{(\sigma_i^B)^2} + \sum_{j=1}^{N_{\text{jet}}} \sum_{k=1}^{N_{\text{track}}^j} \frac{(D_\perp^k)^2}{(\sigma_\perp^k)^2} \ , \tag{5.9}$$

where B_i is the beam spot position (from LEP) in the i- (x, y, or z) direction, σ_i^B is the size

of the beam envelope in the i-direction, $N_{\rm jet}$ is the number of jets in the event, $N^j_{\rm track}$ is the number of tracks in the j-th jet, $D^k_\perp$ is the distance between the vertex V and the track k on a plane perpendicular to the jet, and $\sigma^k_\perp$ is the uncertainty on that distance.

The typical resolution obtained for the primary vertex position depends on the type of events selected. It is of 30 μm, 10 μm, and 30 μm in the x, y, and z direction, respectively for $\mathrm{Z} \to \mathrm{u\overline{u}}$, $\mathrm{Z} \to \mathrm{d\overline{d}}$, or $\mathrm{Z} \to \mathrm{s\overline{s}}$, and 50 μm, 10 μm, and 60 μm for $\mathrm{Z} \to \mathrm{b\overline{b}}$ events.

5.5 Impact parameter

A jet axis is defined from the vector momenta of its particles and from the primary vertex position. Let S be the point of closest approach of a track to this axis. The distance D between the tangent of the track at the point S, and the primary vertex is called the impact parameter of the track with respect to the primary vertex. It is defined to have the same sign as the scalar product of $\vec{VS}$ and the jet direction. A schematic drawing of the impact parameter definition is given in Fig. 5.4. The uncertainty $\sigma_{\rm D}$ on the impact parameter is

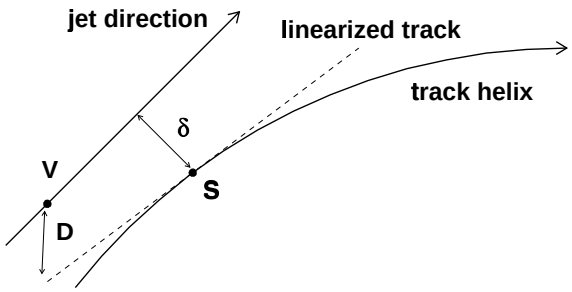

Figure 5.4: Impact parameter definition: V is the primary vertex, S is the point of closest approach of the track helix to the jet direction, δ is the jet-track distance, D is the impact parameter. The linearized track and the jet direction may lie in different planes. In this example, the track has a positive impact parameter.

calculated from the track fit covariance matrix, the primary vertex uncertainty, and their correlation. A non-zero correlation between the primary vertex position and the track fit parameters arises because tracks are used in the interaction point position estimate.

The impact parameter and its sign are used to identify charged particles with or without lifetime, *i.e.*, particles which are likely to come from the decay of a long-lived particle (such as a b-hadron) or not. Tracks with a negative impact parameter are likely to originate from the interaction point, whereas tracks issued from secondary vertices have a positive impact parameter. These properties are used for the selection of $\mathrm{Z} \to \mathrm{b\overline{b}}$ events, as explained in Section 5.7.

5.6 Inclusive secondary vertex finders

Examples of particles which end in a decay point separated from the interaction point are b-hadrons or c-hadrons, with a typical decay path of 3 mm and 2 mm respectively. The expected number of charged particles in the decay point of a b-hadron is five, and about five other charged particles are present on average from the fragmentation process. With as many tracks, the secondary vertex finding technique of forming a vertex from all possible track combinations is impracticable (because of the too large number of combinations) and a different approach must be followed. The aim of inclusive secondary vertex finder algorithms is to obtain an estimate of the decay position of these long-lived particles, without forming a vertex from all combinations of tracks. The algorithm does not provide a track-to-vertex association, instead, for each track, a probability of belonging to the decay vertex is given. Two such algorithms are used in ALEPH. The methods and the performance of both of them are explained in the following Sections.

5.6.1 QVSRCH

The QVSRCH algorithm [77] is based on a vertex search in a predetermined coordinate space, rather than a search among track combinations. For given primary and secondary vertex positions, a $\Delta\chi^2$ is defined as the difference between the χ^2 when all tracks are assigned to the primary vertex, and the sum of the primary and secondary vertex χ^2 values when some tracks are transferred from the primary to the secondary vertex. The maximum $\Delta\chi^2$ for fixed vertex coordinates is obtained by assigning each track to its closest vertex. The central idea in the algorithm is to calculate this $\Delta\chi^2$ for a grid in secondary vertex coordinate space, and determine the secondary vertex as the point of maximum $\Delta\chi^2$.

The event is divided in two jets and a secondary vertex is looked for in each of them. Two cylinders centred along both jet directions, with radius of 500 μm and length of 1 cm are defined. The $\Delta\chi^2$ is computed in every point of a grid in these cylinders (steps of 200 μm along the jet direction and 20 μm in the two orthogonal directions to it). The $\Delta\chi^2$ value between grid points is extrapolated from the values at the grid points with parabolic functions. These same functions are used to estimate the secondary vertex covariance matrix.

The way the algorithm is built guarantees a secondary vertex position to be obtained for most events, and therefore the efficiency on the decay length reconstruction is reasonably uniform as a function of the decay length (as seen later in Fig. 5.6). The impact parameter of each track with respect to the primary and secondary vertices is used to compute a probability of each track to belong to one or the other vertex. The maximum value of the $\Delta\chi^2$ gives an estimate of the vertex quality.

5.6.2 VNFIT

The VNFIT algorithm [78] is based on a likelihood fit which allows an arbitrary number of vertices for a set of tracks to be determined. The likelihood function used is the probability that the tracks are produced at a given number of vertices. A numerical maximization of this likelihood provides an estimate of a predetermined number of vertex positions as well as their uncertainties. The most common configuration for using this algorithm is to set

the number of vertices to three: the primary vertex (which can be given as input) and two secondary vertices for two b- or c-hadron decays (one in each hemisphere). Some additional constraints can be imposed, as the *a priori* knowledge of a track to come from a given vertex, for instance.

The likelihood value obtained after the maximization, assuming Gaussian uncertainties, is used to estimate the probability of the vertex as well as the probability of each track to come from its closest vertex.

Some differences are observed in the performance of the two algorithms described above. The second one, VNFIT, has the advantage to provide, on average, a slightly more accurate estimate of the decay point position. A comparison on the decay length resolution achieved for a common sample of simulated b-hadron decays with both algorithms is presented in Fig. 5.5.

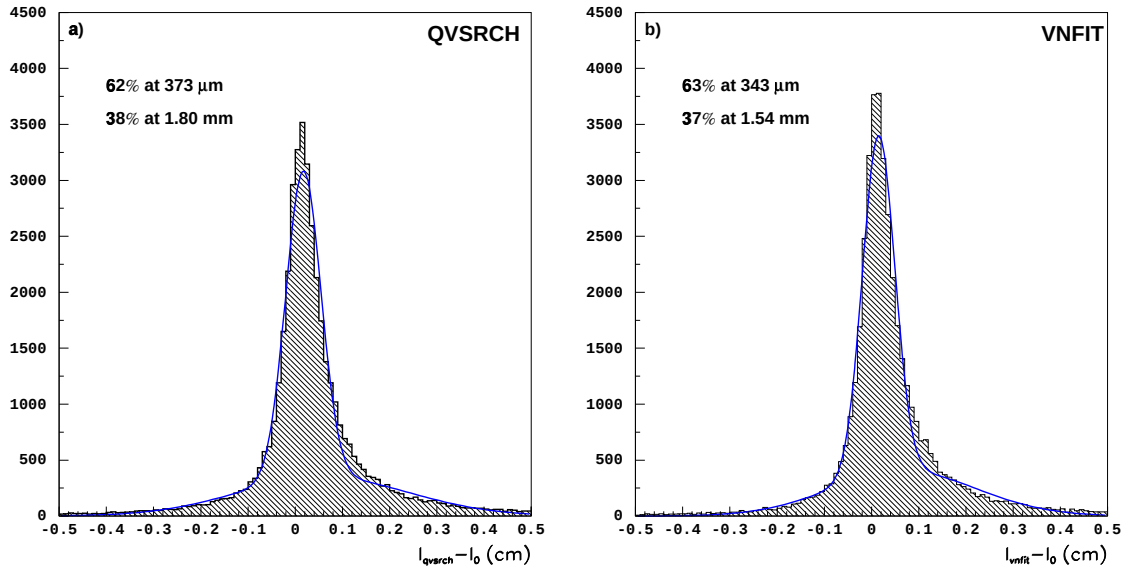

Figure 5.5: Decay length resolution for b-hadrons using a) QVSRCH, and b) VNFIT.

However, the QVSRCH algorithm is more efficient overall, in particular for b-hadron decay points close to the interaction point. The decay length reconstruction efficiency for both algorithms as a function of the decay length is shown in Fig. 5.6.

The two algorithms can therefore be used complementarily to obtain the best possible performance. For those decays for which VNFIT provides a secondary vertex position, this position and its uncertainty estimates are considered, the QVSRCH estimates are taken otherwise.

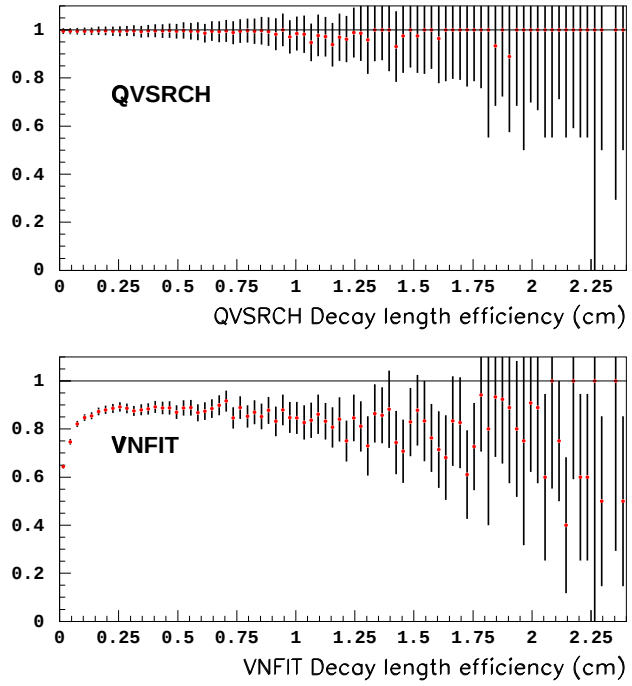

Figure 5.6: Decay length reconstruction efficiency as a function of the decay length for QVSRCH and VNFIT algorithms.

5.7 General b tagging

The aim of "b tagging" is to identify $Z \to b\overline{b}$ decays in a sample of $Z \to q\overline{q}$. The basic characteristics which distinguish $Z \to b\overline{b}$ decays from the other hadronic Z-boson decays are the large lifetime and the large mass of b-hadrons compared to those of other particles produced in $Z \to q\overline{q}$ decays.

Different approaches are possible for "b tagging": from the complete reconstruction of specific b-hadron decays on one side, to fully inclusive tagging where no specific b-hadron, nor decay, are identified, on the other. It is also possible to use the leptonic spectrum properties for the selection of semileptonic b-hadron decays. In this Section only generic selection methods used in ALEPH are described.

5.7.1 QIPBTAG

The QIPBTAG (Ref. [79]) algorithm exploits the non-zero b-hadron lifetime. The signature of a b-hadron is the presence of charged particles in the event which do not originate from the interaction point. In this algorithm the secondary decay vertex is not reconstructed. Instead, the signed impact parameter of all tracks from the Z-boson decay is used (the impact parameter definition, its sign and its uncertainty are given in Section 5.5).

The main idea of the algorithm is to use all tracks with lifetime information (*i.e.*, with significant impact parameter with respect to the interaction point), either coming directly

from the b-hadron decay point or, in the typical case where the b-hadron decays to a c-hadron, coming from the decay point of the c-hadron. A track selection (slightly more restrictive than that described in Section 5.1) based on the quality of the track reconstruction is applied, and the signed impact parameter is computed for all selected tracks. Only tracks with positive impact parameter are considered for the "b-tagging" estimator computation. Tracks with negative impact parameter are used as a calibration sample as they constitute an almost pure set of tracks originated from the interaction point (the contamination of tracks originating from a secondary vertex in that sample is typically about 1%).

The statistical resolution on the impact parameter varies strongly as a function of the charged particle momentum, angle, and number of VDET hits. It is therefore difficult to interpret the value of the impact parameter itself. Instead, the algorithm uses the impact parameter significance, $\mathrm{D}/\sigma_{\mathrm{D}}$. The resolution function for $\mathrm{D}/\sigma_{\mathrm{D}}$ is measured on the calibration sample, and it is used to define the probability, $\mathcal{P}_t$, that a track with positive measured $\mathrm{D}/\sigma_{\mathrm{D}}$ originates from the interaction point.

Given a group of N tracks with positive impact parameter, the combined b-tagging estimator is defined as

$$\mathcal{P}_N \;=\; z \sum_{j=0}^{N-1} \frac{(-\ln z)^j}{j!} \;, \tag{5.10}$$

where $z \equiv \prod_{i=1}^{N} \mathcal{P}_{t_i}$, so as to give a uniform distribution of $\mathcal{P}_N$ between zero and one, for events without lifetime.

Small values of the combined estimator indicate low probability of the tracks to all come from the interaction point, and therefore a high probability that they belong to a $\mathrm{Z} \to \mathrm{b\overline{b}}$ decay. In Fig. 5.7, the value of the estimator computed with all selected tracks in each event is displayed; on Fig. 5.7a for simulated events with a Z-boson which decays into a light-quark pair, $\mathrm{u\bar{u}}$, $\mathrm{d\bar{d}}$, or $\mathrm{s\bar{s}}$, and on Fig. 5.7b for $\mathrm{Z} \to \mathrm{b\overline{b}}$ decays. For a generic cut on the global estimator at 10^{-2}, an efficiency of 68% with a purity of 83% are obtained. The most important contamination comes from $\mathrm{c\overline{c}}$ events due to their lifetime; this contamination accounts for $\sim 14\%$ in the case of the previous cut.

5.7.2 QVSRCH

The inclusive vertex finder algorithm QVSRCH (Section 5.6.1), can be used to select $\mathrm{Z} \to \mathrm{b\overline{b}}$ decays as well. The relevant variable to distinguish $\mathrm{b\overline{b}}$ decays from the rest is the $\Delta\chi^2$ for the two secondary vertices found in the event.

Figure 5.8 shows the distribution of the sum of the two $\Delta\chi^2$ for $\mathrm{u\bar{u}}$, $\mathrm{d\bar{d}}$, and $\mathrm{s\bar{s}}$ in Fig. 5.8a and for and $\mathrm{b\overline{b}}$ simulated events in Fig. 5.8b. The requirement that $\Delta\chi^2$ be larger than 16 gives a $\mathrm{b\overline{b}}$ purity of 82% for an efficiency of 63%. The main background contribution, as in the case of QIPBTAG, comes from $\mathrm{c\overline{c}}$ events which contribute 67% of the total contamination.

5.7.3 Mass tag: QBMTAG

The QBMTAG algorithm exploits the b-hadron mass, larger than that of other hadrons produced in Z-boson decays, and allows the b-purity of the selected sample to be further

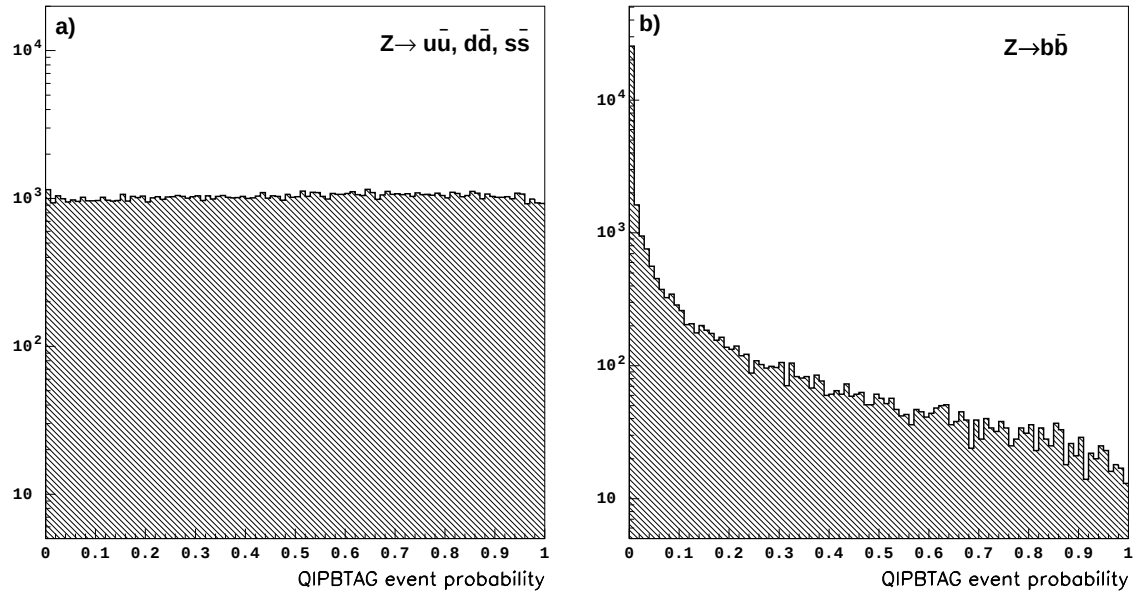

Figure 5.7: QIPBTAG estimator for a) $Z \rightarrow u\bar{u}, d\bar{d}, s\bar{s}$, and b) $Z \rightarrow b\bar{b}$ simulated events.

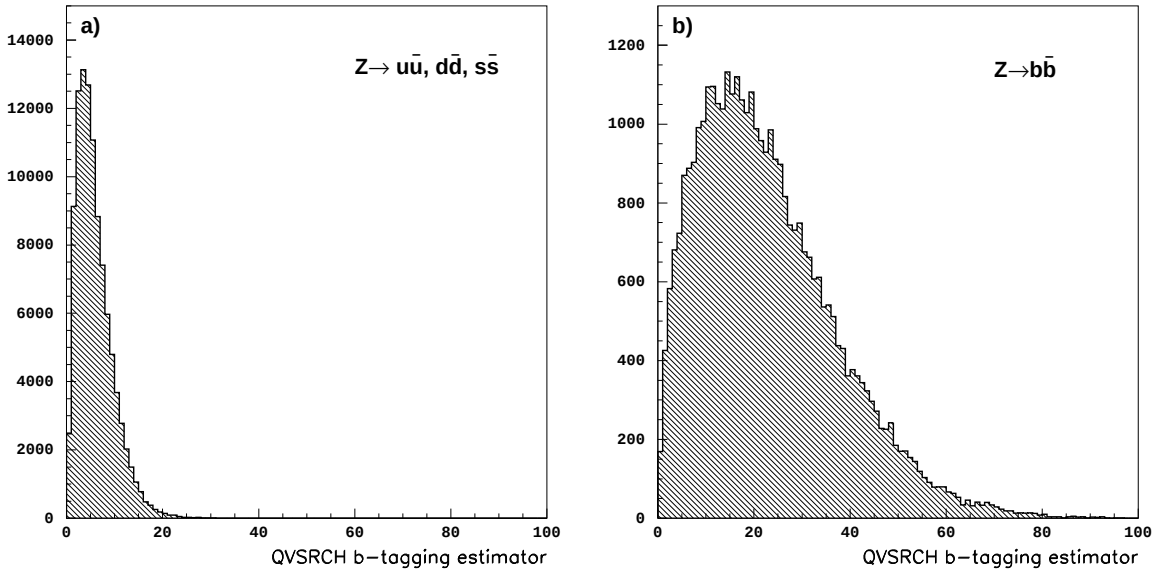

Figure 5.8: QVSRCH b-tagging estimator distribution for a) $Z \rightarrow u\bar{u}, d\bar{d}, s\bar{s}$, and b) $Z \rightarrow b\bar{b}$ simulated events.

increased.

For each hemisphere, tracks are ordered from the least to the most consistent with originating from the interaction point. The first tracks are grouped to form a c-hadron, *i.e.*, a particle with an invariant mass below $1.8\,\mathrm{GeV}/c^2$. The probability of the last track of this group to come from the interaction point, $\mathcal{B}_t$, is taken as b-tag estimator for the hemisphere. The track ordering may follow either the QIPBTAG or the QVSRCH scheme.

In Fig. 5.9a, the distribution of, $-\log_{10}(\mathcal{B}_t)$, is shown for hemispheres in $\mathrm{u\bar{u}}$, $\mathrm{d\bar{d}}$, and $\mathrm{s\bar{s}}$ simulated events, and in Fig. 5.9b for $\mathrm{Z} \to \mathrm{b\bar{b}}$ hemispheres (for illustration the track ordering was taken from the QIPBTAG algorithm).

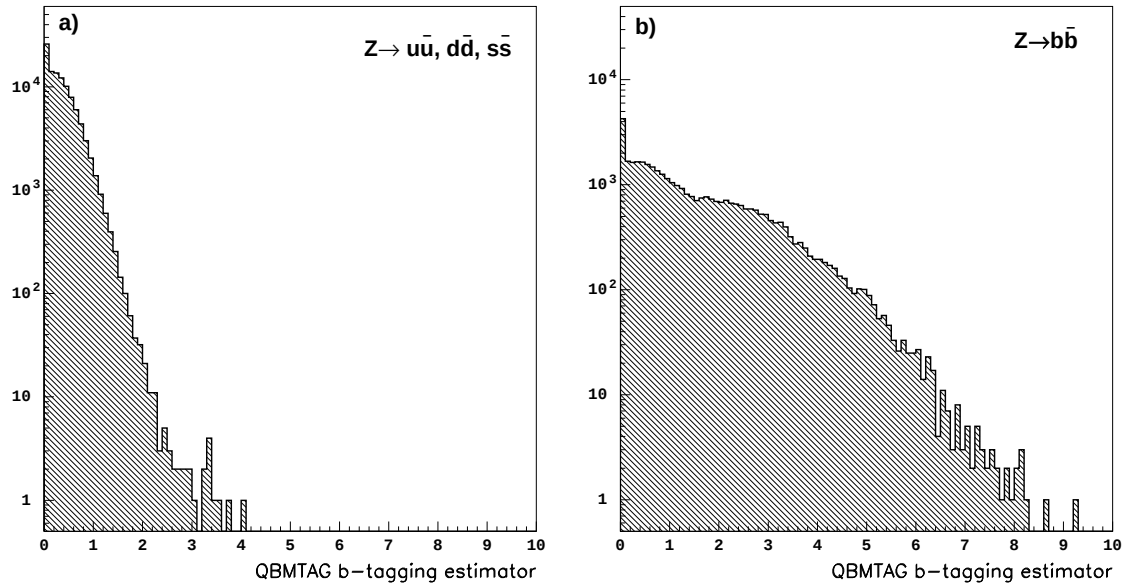

Figure 5.9: QBMTAG estimator distribution for $\mathrm{Z} \to, \mathrm{u\bar{u}}, \mathrm{d\bar{d}}, \mathrm{s\bar{s}}$, and b) $\mathrm{Z} \to \mathrm{b\bar{b}}$ simulated events.

A combination of the QBMTAG estimator on an event with either QIPBTAG or QVSRCH allows a purer $\mathrm{Z} \to \mathrm{b\bar{b}}$ sample to be selected. Typically values of a 98% $\mathrm{b\bar{b}}$ purity and an efficiency of 40% can be achieved.

Chapter 6

Inclusive semileptonic event selection

At LEP, the B_s oscillations analyses with the highest sensitivity are obtained with inclusive semileptonic event samples. The charge of the lepton issued from the b-hadron decay is used to determine the flavour of the reconstructed b-hadron at decay time (b or $\bar{b}$). The advantage of inclusive event selections over partial decay reconstructions like $D_s\,\ell$ pairs comes from statistics. In the first case, over 50 thousand events can be selected at LEP, to be compared with only few hundreds in the second case.

In this Chapter, a new inclusive semileptonic event selection performed on the ALEPH data sample collected around the Z resonance is presented. To perform a B_s oscillation analysis some characteristics of the events need to be determined. These characteristics are mainly, *i*) the proper time of the b-hadron decay; *ii*) the probability of the lepton to originate through a direct $b \to \ell$ decay; *iii*) the probability of the b-hadron to be a B_s; *iv*) the unmixed/mixed label of the decay; and *v*) the probability of this label to be incorrect, called the "mistag probability". To achieve the best possible sensitivity for oscillations, a detailed characterization is needed. Ideally all above properties are evaluated on an event-by-event basis. All the steps of the selection and event characterization are described in this Chapter.

With respect to the previous ALEPH analysis on the subject [10], significant improvements were obtained on both the event selection and the characterization of the sample. The event selection is tailored to identify events with a good b-hadron decay vertex reconstruction. To this end, a new vertex algorithm was designed and the sample was divided in classes based on the event topological properties. Some aspects of the event characterization require the combination of several discriminant variables. This combination is in all cases based on neural networks (Ref. [80]) to obtain an optimal performance and a correct treatment of the correlations among the discriminant variables.

The data sample and the simulated event samples used in the analysis development are briefly described in Section 6.1, and the event pre-selection in Section 6.2. The new procedure to reconstruct the decay position of the charmed particle, and the b-hadron, as well as the criteria followed to reject poorly reconstructed events, are described in Section 6.3. The final

event selection is obtained after the b-hadron enrichment (Section 6.4) and the $b \to \ell$ enrichment (Section 6.5) procedures are applied. The overall selection performance is summarized in Section 6.6.

The first step of the event characterization is the proper time measurement explained in Section 6.7. The presence of a kaon in the decay products of the b-hadron, or among the fragmentation particles, can be used both to enrich the sample in B_s decays, and for the estimate of the $b/\bar{b}$ content of the candidate at production time. The procedure followed to identify charged and neutral kaons is described in Section 6.8. A technique to statistically disentangle B_s decays from other b-hadron decays was developed as explained in Section 6.9. Finally, the estimate of the b-quark content of the b-hadron at production time is described in Section 6.10.

6.1 Data and simulation samples

The data sample analysed corresponds to about four million hadronic Z decays collected by the ALEPH detector from 1991 to 1995. In 1998, these data were reprocessed with a refined version of the reconstruction program. The main improvements concern the track reconstruction and the particle identification [67]. The analysis presented here benefits from all these improvements.

During its development phase, the analysis relied exclusively on samples of simulated events. The generation of these events is done through the following process, *i*) all particles originated from the Z decay are generated according to the theoretical knowledge of the simulated process; *ii*) their interaction with the ALEPH detector is simulated using the GEANT [81] package; and *iii*) the data obtained is reconstructed with the same procedure as the *real* data. The events obtained have the same format as the *real* data, with the addition of the *true* information about the properties of all the particles in each event. This knowledge was used to optimize the selection procedure, estimate the selection efficiency, and extract all parametrizations needed for the complete event characterization.

Hadronic events were generated with JETSET [82]. Different type of events were simulated: 8.5 million unbiased $Z \to q\overline{q}$ decays, 5 million $Z \to b\overline{b}$ decays, 2.3 million $Z \to c\overline{c}$ decays, almost one million $Z \to bB_s$ ($B_s \to \ell$) decays, and about half a million $Z \to bB_s$ ($B_s \to c \to \ell$). All physical parameters used as inputs for the event generation (b-hadron fractions, decay branching ratios, fragmentation function, etc.) were adjusted to the knowledge to date, as summarized in Ref. [13].

6.2 Data sample pre-selection

The first selection criterion applied is the "hadronic selection", explained in Section 5.1, to obtain a quasi-pure sample of $Z \to q\overline{q}$ decays. The events are separated in two hemispheres with respect to a plane perpendicular to the thrust axis. A cut on the polar angle of the thrust axis is applied to reject events with most of the tracks outside the VDET acceptance: events must satisfy $|\cos\theta_{\rm thrust}| \leq 0.85$.

The events are then clustered in jets reconstructed from all energy-flow objects using the JADE algorithm (Section 5.2.1), with a jet-resolution parameter ($y_{\rm cut}$) of 0.0044 (this parameter was optimized for heavy flavour analyses with leptons in ALEPH [83]). In the following, this is the jet definition used to characterize events by their number of jets.

Events are required to have at least one lepton candidate (Section 5.3) reconstructed with a VDET hit. The transverse momentum of the lepton is defined with respect to the axis of its JADE jet (with the definition above) once the lepton itself is removed from the jet. If more than one lepton candidate is found, only that with highest transverse momentum is considered in what follows. A hemisphere is labeled as "same" or "opposite" according to whether it contains the lepton candidate or not.

The lepton requirement enriches the event sample in $Z \to b\overline{b}$ decays, from about 20% to more than 55%.

6.3 Vertex algorithm and selection

The typical topology of a semileptonic B_s decay is shown in Fig. 6.1 together with the decay of the other b-hadron in the event. The B_s decay position is the origin of the lepton track, of

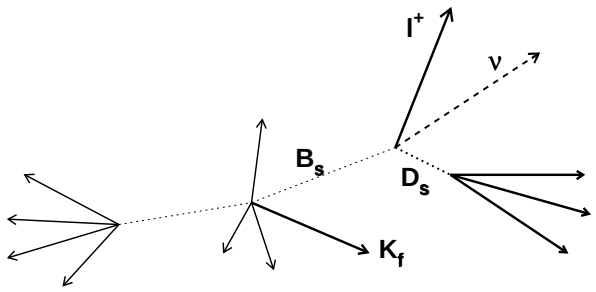

Figure 6.1: Semileptonic B_s decay schematic, tracks from the b-quark fragmentation and the decay of the other b-hadron in the opposite hemisphere of the event, are also shown.

the undetected neutrino, and of a charmed particle. The latter, with a typical decay length of 1 mm, decays at a tertiary vertex. The vertex algorithm is optimized on such topologies and it also provides criteria to select such events.

The vertex algorithm developed for this analysis starts by the inclusive reconstruction of the charmed particle. This particle is extrapolated back to the lepton and a common vertex is looked for. The B_s oscillation study requires a good estimate of the b-hadron decay length and of its uncertainty. The procedure developed for this analysis is explained in the following Sections.

6.3.1 Search for a seed

In a semileptonic B_s decay, the reconstructed decay products are of two types: *i*) the lepton, which is the only reconstructed charged particle issued directly from the meson decay; and *ii*) the particles coming from the decay of the charmed hadron. If the lepton track is ignored, the only remaining tracks with sizeable impact parameter in the B_s hemisphere are the D_s^- (charmed particle) decay products. The "search for a seed", *i.e.*, a point in space, used as a first approximation of the tertiary vertex, is based on this observation.

As explained in Section 5.6, two inclusive secondary vertex finders, VNFIT and QVSRCH, are used. The lepton candidate track is ignored and a first estimate of the tertiary vertex position is searched for with VNFIT. If the search is successful, the vertex position and uncertainties are kept as the seed for the charmed particle. When VNFIT fails, QVSRCH is used instead.

Both algorithms provide, for each track, a variable which estimates its compatibility with originating from the tertiary vertex (Sections 5.6.1 and 5.6.2). However, other track properties can be used to disentangle fragmentation particles from b-hadron decay products, as explained in the next Section.

6.3.2 The c track selection

The first step of the c-track selection consists in separating charged particles from the b-quark fragmentation from those from the b-hadron decay with the use of some of their kinematical properties. All the relevant variables are combined using a neural network. The list of discriminant variables used is given, followed by an explanation of their relevance.

- The compatibility of the track with the inclusive decay vertex (the track probability in the case of a VNFIT vertex and the impact parameter significance in the case of QVSRCH).
- The impact parameter significance of the track with respect to the primary vertex, signed using the b-jet axis.
- The particle momentum: particles from the b-hadron decay are expected to be, on average, of higher momentum than those originated in the b-quark fragmentation.
- The angle of the particle with respect to the estimated b-hadron decay direction. Three estimates of this direction are considered:
  - the thrust axis,
  - the jet axis, and
  - the direction defined by the primary and the inclusive tertiary vertices.

  The angle of the decay particles with respect to the true b-hadron decay direction is expected to be smaller than that of fragmentation particles. The quality of the direction estimates presented above depends on both the number of jets in the event and the decay length of the b-hadron. The b-hadron jet used here is a cone-based jet built around the lepton candidate. The algorithm is chosen for its performance on the direction

estimate. When the tertiary vertex is far apart from the primary vertex, the direction defined by the two vertices is the best estimate of the b-hadron flight direction. When the distance between the two vertices is small, a better estimate can be provided by the thrust axis or the jet axis. The thrust axis is found to be more accurate in two jet events, while for events with more than two jets the b-jet is a better estimator (because of hard gluon emission).

When building a combined discriminant variable with a neural network, it can happen that the best performance is obtained when the combination includes variables that have discriminating power and variables which are not discriminant by themselves, but are correlated with the relative discriminant power of the others. These auxiliary variables, called *control variables* in the following, allow an optimal combination to be obtained in the whole phase space.

In the present case, the control variables included in the neural network are, the number of jets in the event, the χ^2 of the tertiary vertex fit, the measured decay length, and the measured jet momentum.

The distributions of the most relevant variables for all tracks in hemispheres with an identified lepton are shown in Figs. 6.2 to 6.4. In each plot, the dark area corresponds to tracks originated in secondary or tertiary vertices in $Z \to b\overline{b}$ simulated events. Charged particles from $Z \to u\overline{u}$, $d\overline{d}$, $s\overline{s}$, and $c\overline{c}$ decays are displayed as "udsc". The agreement between the data and the simulation is reasonable for most of the variables. In the case of the angular variables a deviation is seen. More tracks are found at large angle in the data than in the simulation. This disagreement is probably due to the lack of accuracy of the quark fragmentation simulation. However, as the track separation combined variable is only used to select charged particles to form the charmed object, the observed disagreement is not considered to be a problem. The peak at $P_{VNFIT} \sim 0.5$ in Fig. 6.4b is mostly due to charged particles with low momentum ($p < 500\,\mathrm{MeV}/c$), and poorly reconstructed tracks.

The neural network training was carried out on simulated $B_s \to \ell$ events, all fragmentation and decay charged particles with momentum above $200\,\mathrm{MeV}/c$, with the exclusion of the lepton candidate, were used for the training. Two different trainings were performed, one for events with a VNFIT vertex and another for the remaining events.

The distribution of the combined track separation variable for events with a VNFIT vertex is shown in Fig. 6.5. The performance obtained with the other training are not as good. Those events for which the VNFIT algorithm does not provide a vertex are topologically more complicated than the average, thus with a more difficult track separation.

Two different c track selections are defined depending whether the VNFIT algorithm was successful in the event or not. In both cases, all tracks except the lepton candidate are ordered according to their compatibility with the c-hadron decay vertex, P_{track}. This compatibility is defined as the output of the track separation neural networks described above.

In the case of events with a VNFIT vertex, the track most compatible with the c-hadron decay vertex is taken. Tracks with a momentum $p > 0.5\,\mathrm{GeV}/c$, a VNFIT track probability $P_{VNFIT} > 0.7$, and $P_{track} > 0$ in decreasing value of P_{track} are added one by one till a mass of $1.8\,\mathrm{GeV}/c^2$ for the charmed candidate is reached.

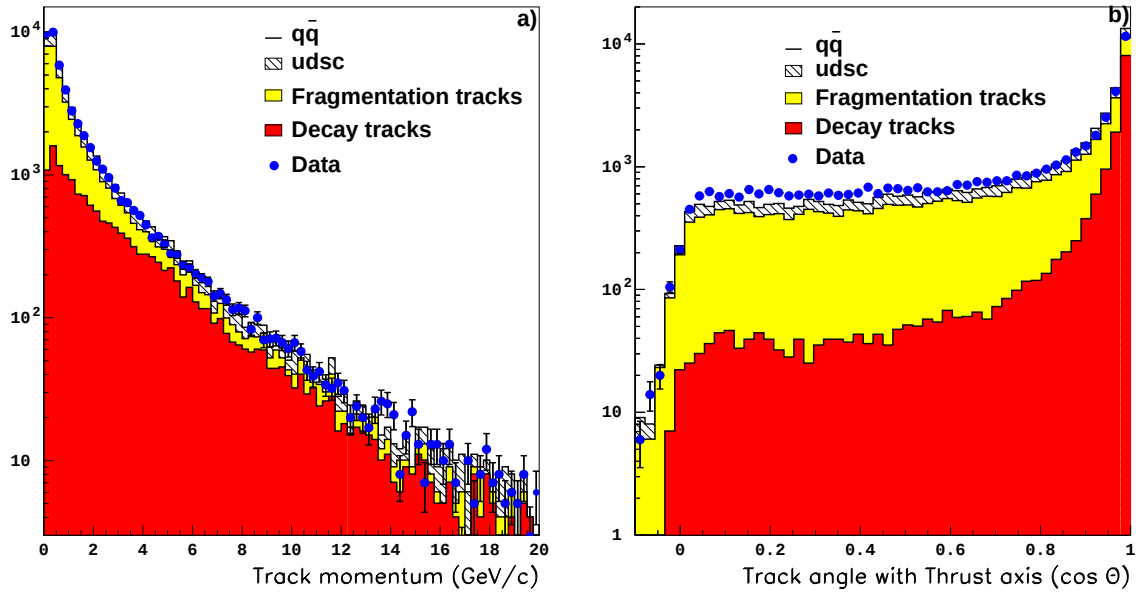

Figure 6.2: Distribution of the discriminant variables for the track separation: a) track momentum, and b) track angle with the thrust axis $\cos(\theta_{\mathrm{trhust}})$.

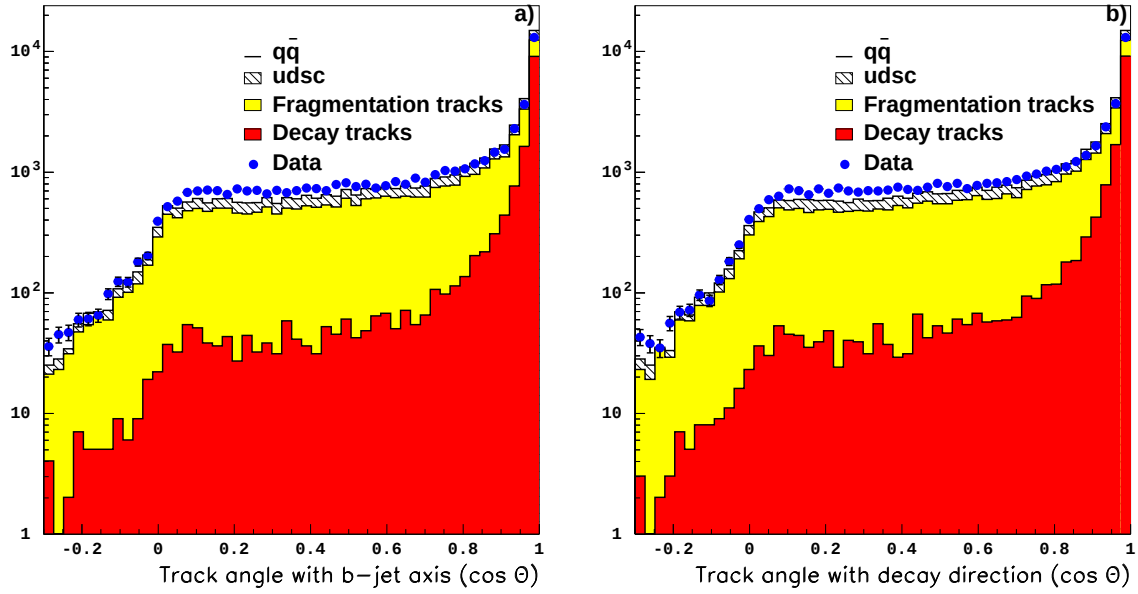

Figure 6.3: Distribution of the discriminant variables for the track separation. Track angle with a) the jet direction $\cos(\theta_{\mathrm{jet}})$, and b) the decay direction $\cos(\theta_{\mathrm{decay}})$.

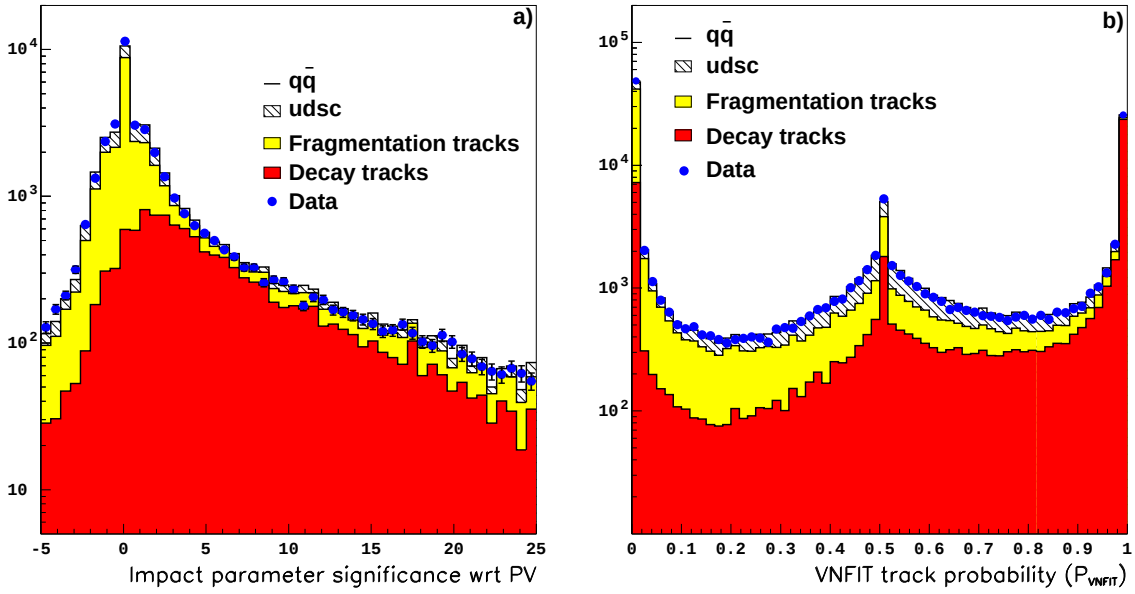

Figure 6.4: Distribution of the discriminant variables for the track separation. a) Impact parameter significance with respect to the primary vertex, and b) VNFIT track probability (compatibility with the decay vertex).

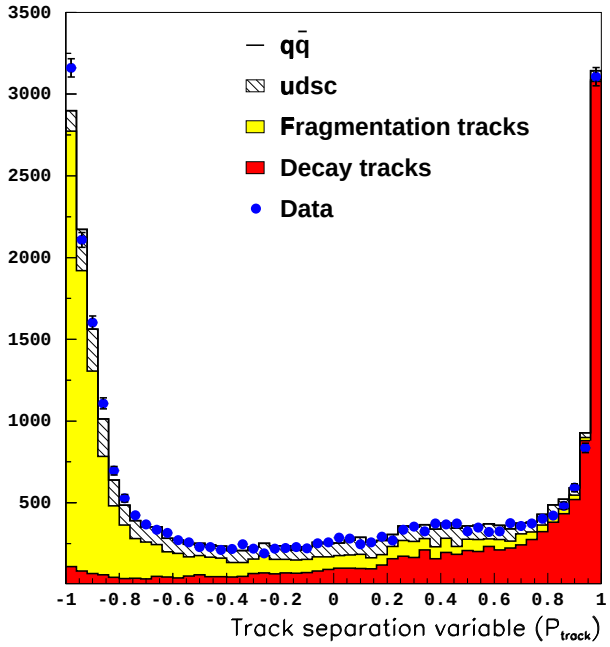

Figure 6.5: Track separation neural network output in events with a vertex reconstructed by VNFIT.

When a VNFIT vertex is not available, all tracks with an impact parameter significance with respect to the QVSRCH vertex smaller than 1.4, a momentum higher than $1.5\,\mathrm{GeV}/c$, and a c-hadron decay vertex neural network probability higher than $\mathrm{P_{track}} > 0$ are taken to form the charmed particle.

Events where no track is selected to form a charmed particle are rejected. In 28% of the events only a single track is selected to form the charmed object. If more than one charged particle is selected, the set of tracks are fit together to obtain the charm vertex.

The purity of the c-track selection is estimated to be 93% on $\mathrm{B_s} \to \ell$ simulated events, with an efficiency of 63%.

6.3.3 Photon recovery

In most of the cases, the charmed hadron decays in charged and neutral particles. In the previous Section, a procedure to select the charged particles was presented. However, if the neutral energy component is ignored, the reconstructed charm direction obtained from the charged particle fit has a larger uncertainty than that estimated from the track fit itself. This effect is translated on the b-hadron decay vertex position obtained from the fit of this charged charmed particle and the lepton. In this Section, a strategy to recover as much as possible of the information from the neutral particles is explained. An improvement on the charm decay direction is only possible when more than one charged particle is found. Indeed, neutral particles cannot help improving the charmed vertex position determination. Only the momentum direction is therefore modified by the inclusion of neutral particles.

The most energetic charged particle included in the charmed object described above is taken as a seed for a cone-based jet. This algorithm is chosen because it provides the best estimate of the decay direction. An angular aperture of 13° between two consecutive particles to be merged is taken. It is verified that all the charged particles included in the charmed object are also included in the cone-based charmed jet. Photons are looked for inside this jet. If more than one photon with energy higher than 1 GeV is found, the most energetic photon in a 16° cone around the jet direction, making an invariant mass of less than $1.8\,\mathrm{GeV}/c^2$ with the charged charmed object is taken.

By default, photons have their momentum calculated as if they were coming from the primary vertex. Here, the momentum is recomputed at the charm vertex to obtain an improvement on the decay length resolution. The position of the electromagnetic shower is defined as the energy-weighted barycentre of the relevant ECAL towers positions. The direction between the charm vertex and the shower centre defines the new photon momentum direction.

The improved charmed object momentum is obtained as the sum of the charged and photon components. Finally the covariance matrix for the new charmed object and the correlation matrix between the new charmed object momentum and the charm vertex are

calculated. The covariance matrix for the charmed object momentum reads

$$\begin{aligned} \mathrm{Cov}_{ij}^{\mathrm{new-mom}} \;=\; & \sum_{kl} \frac{\partial p_i}{\partial p_k}\, \mathrm{Cov}^{\mathrm{mom}}(p_k, p_l)\, \frac{\partial p_j}{\partial p_l} \;+ \\ & \sum_{kl} \frac{\partial p_i}{\partial p_k}\, \mathrm{Cov}^{\mathrm{mom-vertex}}(p_k, v_l)\, \frac{\partial p_j}{\partial v_l} \;+ \\ & \sum_{kl} \frac{\partial p_i}{\partial v_k}\, \mathrm{Cov}^{\mathrm{vertex}}(v_k, v_l)\, \frac{\partial p_j}{\partial v_l} \;+ \\ & \sum_{k} \frac{\partial p_i}{\partial p_k^{\mathrm{phot}}}\, \sigma_k^{\mathrm{phot}}\, \sigma_k^{\mathrm{phot}}\, \frac{\partial p_j}{\partial p_k^{\mathrm{phot}}} \;, \end{aligned} \tag{6.1}$$

where $\mathrm{Cov}^{\mathrm{mom}}$ is the covariance matrix for the charmed object momentum obtained from the fit of the selected c tracks, $\mathrm{Cov}^{\mathrm{mom-vertex}}$ is the correlation matrix between the charged charmed object momentum and the charm vertex position, $\mathrm{Cov}^{\mathrm{vertex}}$ is the covariance matrix of the charm vertex, and σ_k^{phot} is the photon momentum uncertainty in the k-th direction. The correlation matrix between the new charmed object and the charm vertex position reads

$$\mathrm{Cov}_{ij}^{\mathrm{new-vertex-mom}} \;=\; \mathrm{Cov}^{\mathrm{mom-vertex}}(p_i, v_j) \;+\; \sum_{kl} \frac{\partial p_i}{\partial v_k}\, \mathrm{Cov}^{\mathrm{vertex}}(v_k, v_l)\, \frac{\partial v_j}{\partial v_l} \;. \tag{6.2}$$

The addition of a photon to the charged charmed object is found to improve significantly the final b-hadron decay position only in some particular event topologies. In Section 6.3.7 vertex classes are defined and the class of events where a photon is added is identified.

6.3.4 The "B track"

The b-hadron vertex in an inclusive semileptonic selection is generally obtained [10] as the common vertex between the lepton and the charmed object. Here instead, information on the b-hadron flight direction is used to constrain and improve the final estimate of the secondary vertex position. The "B track" is defined as a mathematical object which aims at having the direction and momentum of the b-hadron and is used as a constraint in the global b vertex fit (see Section 6.3.5 below).

The quadri-momentum vector of the "B track" is obtained either from the estimate of the b-hadron flight direction and momentum in the case of multijet events, or from the thrust direction and the b-hadron momentum in the case of two jet events. The "B track" is imposed to originate from the reconstructed primary interaction point.

The estimate of the b-hadron flight direction is obtained from that of a cone-based b-jet formed around the lepton candidate. An angular aperture of 23°, higher than in the case of the charmed jet, is used. The energy of the jet defines the "B track" energy. In the case of two jet events, the direction estimator is chosen to be the thrust axis, which turns out to be more accurate than the jet axis.

To be used as a pseudo-track in the b-vertex fit, the "B track" tracks needs to be complemented with an estimate of its covariance matrix. The covariance matrix is built from the estimated uncertainty on the momentum p of the b-jet, and the uncertainty on the polar and azimuthal angles θ_{B}, φ_{B}, of the direction estimator. The momentum resolution is parametrized as a function of the momentum itself. Three different parametrizations are

taken, one for events with two jets (using the thrust direction and the b-jet momentum), another for events with more than two jets but a single jet in the candidate hemisphere, and a third parametrization for the remaining events (these last two using the cone-based b-jet direction). The polar and azimuthal angles resolutions are also parametrized as a function of the momentum for the same three categories of events. All the parametrizations are obtained from simulated events and optimized for b-jets in semileptonic b-decays. The estimate of the momentum and angles are considered uncorrelated. With all this information the covariance matrix in an (p_x, p_y, p_z, E) basis reads

$$\mathrm{Cov}_{ij} = \sum_{\alpha} \frac{\partial p_i}{\partial p_\alpha} \sigma_{p_\alpha} \sigma_{p_\alpha} \frac{\partial p_j}{\partial p_\alpha} \ , \tag{6.3}$$

where $p_i = p_x, p_y, p_z, E$, and $p_\alpha = p, \theta_{\mathrm{B}}, \varphi_{\mathrm{B}}$. The symbol σ_{p_α} stands for the parametrized resolutions described above.

6.3.5 Global b-vertex fit

The b-hadron decay vertex is obtained from a common fit of the lepton, the charmed object, and the "B track".

With respect to the previous analysis [10], three basic new features are introduced in the b-vertex finding procedure to improve its performance, *i*) the VNFIT inclusive vertex, used here to increase the c track selection efficiency; *ii*) the addition of photons in the charmed particle reconstruction; and *iii*) the use of the "B track" in the final secondary vertex fit.

Most of the improvement in the final estimate of the b-hadron decay position is due to the inclusion of the "B track" in the vertex fit. It introduces information on the decay direction which constrains significantly the vertex fit.

6.3.6 The decay length

The b-hadron decay length is obtained as the three-dimensional distance between the b vertex (Section 6.3.5) and the primary vertex (Section 5.4) positions. In the previous analysis [10], the decay length was estimated as the distance between the primary and secondary vertices projected on the b-jet direction. The projection was found to improve the resolution on the decay length. Here, as the information on the b-hadron flight direction is already included in the secondary vertex fit, a projection does not improve the resolution.

6.3.7 Vertex selection and class definition

A critical study of the previous analysis pointed out that the event selection and the event-by-event estimate of the decay length uncertainty were key issues to achieve an improved performance at high oscillation frequency. In particular, it was clear that defining resolution classes based on the topological properties of the events would be highly beneficial in that sense.

Some of the topological characteristics of the events which serve as basis for the event selection and for the definition of resolution classes are as follows.

- The presence of a VNFIT vertex. Events for which the VNFIT algorithm does not provide a secondary vertex have a less accurate inclusive vertex position estimate to be used as a seed.

- The χ^2 of the reconstructed b-hadron vertex and the χ^2 of the reconstructed charm vertex (if any).

- The angle between the reconstructed charmed particle and the lepton. On one hand, the bigger is the angle between the charmed particle and the lepton, the bigger is the probability of including fragmentation particles in the charm track selection; on the other hand, if the charmed particle and the lepton are close to be parallel, the reconstructed b-hadron vertex has a large uncertainty. Therefore this angle is limited from below and from above.

- The angle between the reconstructed charmed particle and the b-jet. Similar arguments as for the angle with the lepton justifies a cut on the angle of the charmed particle with the b-jet.

- The mass of the charmed particle. In signal events, the charmed particle is expected to be a D_s with a mass $m_{D_s} = 1968.6 \pm 0.6\,\mathrm{MeV}/c^2$ [13]. Therefore the reconstructed charmed particle should not be much heavier than m_{D_s}.

A detailed description of the classes is presented in Table 6.1. These classes are exclusive: if an event fulfils the requirements of more than one class, it is assigned to the first in the list. For classes 3 and 4, the b-hadron decay length is improved when photons are added to the charmed particle as explained in Section 6.3.3. The decay length resolution achieved for two of the vertex classes (the one with the best resolution and the one with the worst resolution) is shown in Fig. 6.6.

The selection criteria described above are tailored to select semileptonic B_s decays, and therefore result in a rejection of light-quark (udsc) hadron decays.

6.3.8 The bias and pull correction

The decay length evaluated as in the previous Sections is found to be slightly biased. This bias, measured in simulated events, shows a dependence on the reconstructed decay length itself. The typical dependence observed is shown in Fig. 6.7 for $B_s \to \ell$ simulated events in the first resolution class. The decay length bias is in all cases smaller than the estimated uncertainty on the decay length.

A correction for this bias is applied to all events in every vertex class. The effect of the bias correction on the decay length resolution in the overall selected sample is shown in Fig. 6.8. The resolution is slightly improved (more events are found in the core) and the bias is reduced on average from $120\,\mu$m to $28\,\mu$m.

An estimate of the decay length uncertainty is obtained, event by event, from the covariance matrices of the primary and secondary vertices, and it is dominated by the uncertainty on the secondary vertex position. The uncertainty on the secondary vertex is estimated from the uncertainty on the helix parameters of the tracks involved, but does not include any contribution from misassigned and missing particles in the charm candidate. As a consequence,

Class number	Vertex properties	Topological properties
Class 1	No VNFIT vertex $\chi^2_{\rm b} < 7.$	$N_{\rm c} = 1$ $5.7^\circ < \widehat{\rm c\,\ell} < 30^\circ$ $\widehat{\rm c\,jet} < 20^\circ$ $0 \leq m_{\rm b\ reco} \leq 8\,{\rm GeV}/c^2$
Class 2	No VNFIT vertex $\chi^2_{\rm b} < 7.$	$N_{\rm c} > 1$ $m_{\rm charm} < 3\,{\rm GeV}/c^2$ $5.7^\circ < \widehat{\rm c\,\ell} < 30^\circ$ $\widehat{\rm c\,jet} < 20^\circ$ $0 \leq m_{\rm b\ reco} \leq 8\,{\rm GeV}/c^2$
Class 3	VNFIT vertex $\chi^2_{\rm b} < 7.$	$N_{\rm c} = 2$ $5.7^\circ < \widehat{\rm c\,\ell} < 37^\circ$ $\widehat{\rm c\,jet} < 20^\circ$
Class 4	VNFIT vertex $\chi^2_{\rm b} < 7.$ $\chi^2_{\rm c} < 5.$	$N_{\rm c} > 2$ $m_{\rm charm} < 3\,{\rm GeV}/c^2$ $5.7^\circ < \widehat{\rm c\,\ell} < 37^\circ$ $\widehat{\rm c\,jet} < 20^\circ$
Class 5	VNFIT vertex $\chi^2_{\rm b} < 25.$ $\chi^2_{\rm c} < 10.$	$N_{\rm c} \geq 2$ $m_{\rm charm} < 3\,{\rm GeV}/c^2$ $\widehat{\rm c\,\ell} < 37^\circ$ $\widehat{\rm c\,jet} < 20^\circ$
Class 6	VNFIT vertex $\chi^2_{\rm b} < 7.$ $\chi^2_{\rm c} < 5.$	$N_{\rm c} = 2$ $5.7^\circ < \widehat{\rm c\,\ell} < 37^\circ$
Class 7	VNFIT vertex $\chi^2_{\rm b} < 7.$ $\chi^2_{\rm c} < 5.$	$N_{\rm c} > 2$ $5.7^\circ < \widehat{\rm c\,\ell} < 37^\circ$
Class 8	No VNFIT vertex $\chi^2_{\rm b} < 7.$ $\chi^2_{\rm c} < 5.$	$N_{\rm c} \geq 2$ $m_{\rm charm} < 3\,{\rm GeV}/c^2$ $5.7^\circ < \widehat{\rm c\,\ell} < 37^\circ$
Class 9	VNFIT vertex $\chi^2_{\rm b} < 20.$ $\chi^2_{\rm c} < 5.$	$N_{\rm c} \geq 2$ $5.^\circ < \widehat{\rm c\,\ell} < 15^\circ$
Class 10	VNFIT vertex $\chi^2_{\rm b} < 2.$ $\chi^2_{\rm c} < 1.$	$N_{\rm c} \geq 1$ $11.^\circ < \widehat{\rm c\,\ell} < 41.^\circ$

Table 6.1: Definition of the vertex classes. The number of tracks in the charged charmed object is indicated as $N_{\rm c}$; the mass of the charmed particle as $m_{\rm charm}$; the χ^2 of the b-hadron vertex as $\chi^2_{\rm b}$; the χ^2 of the charm vertex as $\chi^2_{\rm c}$; the angle between the charmed particle and the lepton as $\widehat{\rm c\,\ell}$; the angle between the charmed particle and the b-jet as $\widehat{\rm c\,jet}$; and the reconstructed mass of the b-hadron as $m_{\rm b\ reco}$.

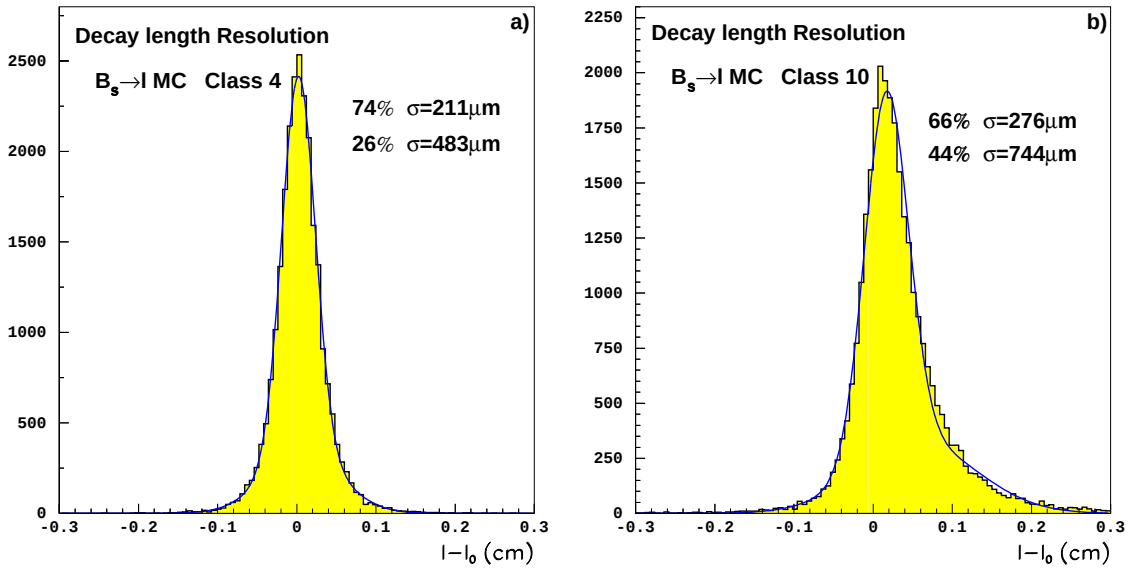

Figure 6.6: Decay length resolution on $B_s \to \ell$ simulated events in two vertex classes, a) the class with the best resolution, and b) the class with the worst resolution. The distributions are fit with the sum of two Gaussians, their fractions and widths are indicated on the plot.

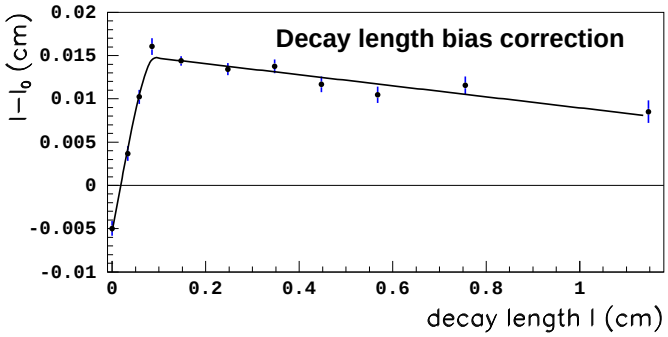

Figure 6.7: Decay length bias as a function of the reconstructed decay length.

the decay length uncertainty obtained from the track fit, is in general, underestimated and a correction based on simulated events needs to be applied. This correction is called *pull correction* in the following, because it can be illustrated with the distribution of $(l - l_0)/\sigma_l$, where l is the reconstructed decay length, l_0 is the true decay length, and σ_l is the estimated uncertainty on the decay length. If the decay length uncertainty were properly estimated for all events, this distribution would be a normal Gaussian.

In the previous analysis a global correction for the selected sample was computed and applied. Here, instead, the problem is addressed class by class to obtain a more accurate description. For the vertex classes in which the charmed particle is made of a single charged particle, the pull correction is parametrized as a function of the reconstructed decay length.

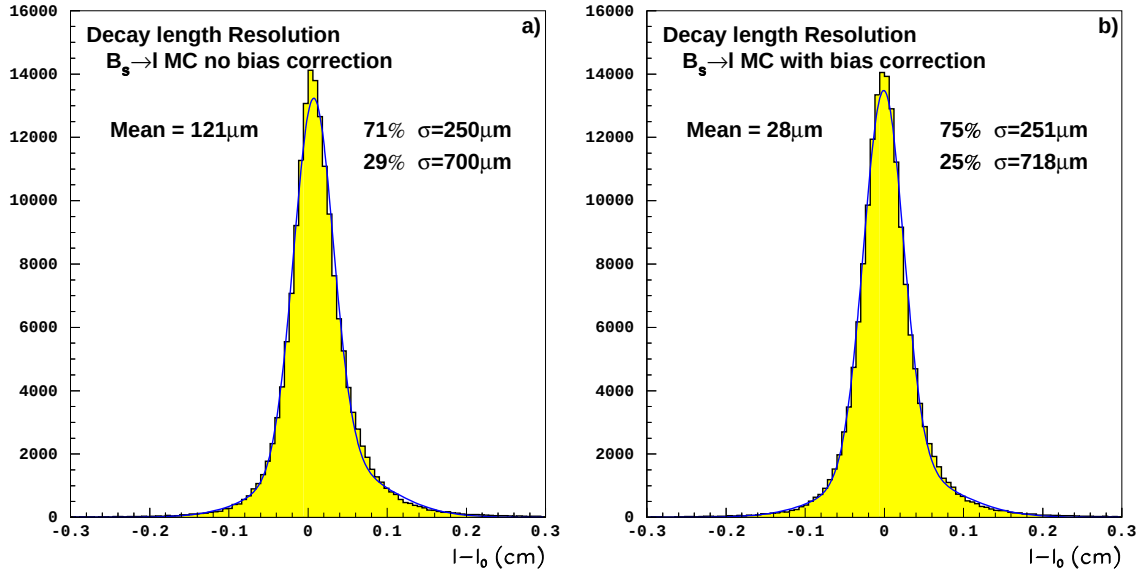

Figure 6.8: Decay length distribution for all selected events, a) before the bias correction, b) after the bias correction is applied.

For the other classes, a double parametrization is performed. The decay length uncertainty is first corrected as a function of the reconstructed charmed mass and then as a function of the reconstructed decay length. In general, as intuitively expected, the correction increases with the decay length and decreases with the reconstructed mass of the charmed particle. The pull distribution for the selected sample is shown in Fig. 6.9, both before and after the parametrized correction is applied.

The impact of the parametrizations described in this Section on the analysis sensitivity is explained in Section 8.4.2.

6.4 The b-hadron enrichment

The requirement of a lepton candidate, and the vertex selection described above, act both as rejection criteria for Z decays in light-quarks. Therefore, the sample obtained at this stage of the selection procedure is enriched in $Z \to b\overline{b}$ decays. Before any further selection is applied, the sample composition estimated from $Z \to q\overline{q}$ simulated events is shown in Table 6.2.

$(82.6 \pm 0.3)\%$ $b\overline{b}$	$(11.00 \pm 0.07)\%$ $c\overline{c}$	$(6.40 \pm 0.05)\%$ $u\overline{u}$, $d\overline{d}$, $s\overline{s}$

Table 6.2: Sample composition after the lepton and vertex selections are applied.

However, a B_s oscillation analysis requires a $b\overline{b}$ purity higher than the $\sim 83\%$ obtained. Therefore, a specific b-tagging variable is designed to increase the $b\overline{b}$ content in the final

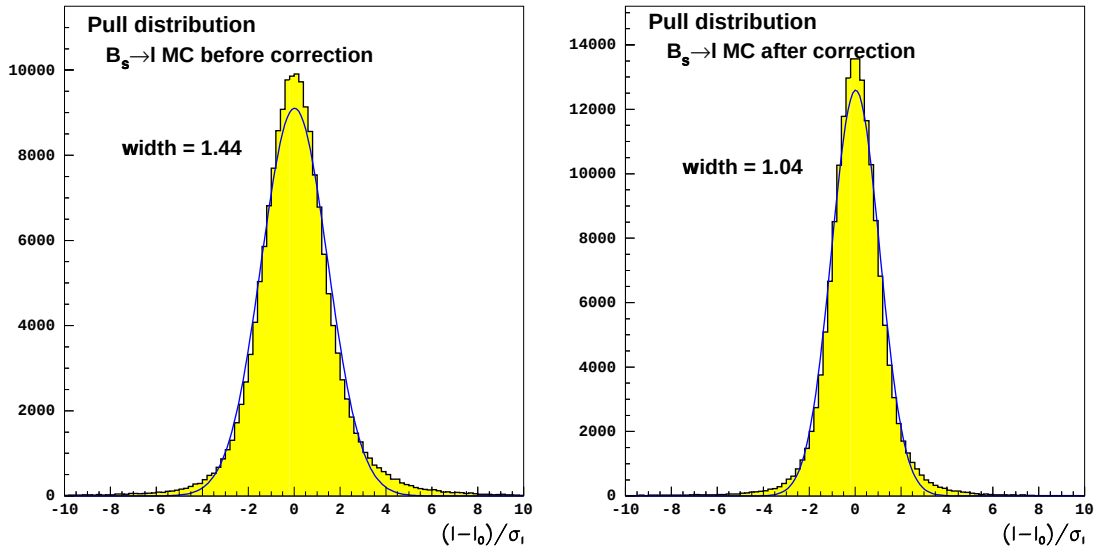

Figure 6.9: Distribution of the decay length pulls, a) before the correction, and b) after the correction is applied.

selected sample.

Several variables which distinguish between $Z \to b\overline{b}$ and other Z decays are combined with a neural network. The hard fragmentation of b quarks into b-hadrons and the high mass of b-hadrons result respectively, in a large momentum and large transverse momentum of the primary lepton compared to that of leptons produced in the decay of light-quark hadrons. The variables described in Section 5.7 for b-tagging are sensitive to both the lifetime and the mass of the hadrons produced in $Z \to q\overline{q}$. All these properties are used to define the set of variables for the b-tagging developed here. The variables are as follows. *i*) The lepton momentum; *ii*) the lepton transverse momentum; *iii*) the same side QIPBTAG probability (Section 5.7.1); *iv*) the opposite side QIPBTAG probability; *v*) the same side QBMTAG estimator (Section 5.7.3); *vi*) the opposite side QBMTAG estimator; *vii*) the same side QVSRCH estimator (Section 5.7.2); and *viii*) the opposite side QVSRCH estimator. The distributions of all the variables which are to be combined in the b-tagging neural network are shown, for events with a candidate lepton and a reconstructed vertex, in Figs. 6.10 to 6.13. Simulated events with a Z decay into light-quarks are labeled as "uds". The flavour discrimination power is shown as well as the agreement between the data and the simulated $Z \to q\overline{q}$ events.

The neural network training is performed with only events which pass both the lepton and the vertex selections. The performance of the combined b-tagging variable is shown in Fig. 6.14.

The performance of the b-tagging variable developed here can be compared with that of a more standard approach (*i.e.*, the sole QIPBTAG event probability) which is normally used to select unbiased $b\overline{b}$ samples (not semileptonic decays specifically). In Fig. 6.15, the selection efficiency on $b\overline{b}$ events versus the $b\overline{b}$ purity curves are displayed for the b-tagging presented here, N_{btag}, and for the QIPBTAG variable defined in the whole event, both applied to the

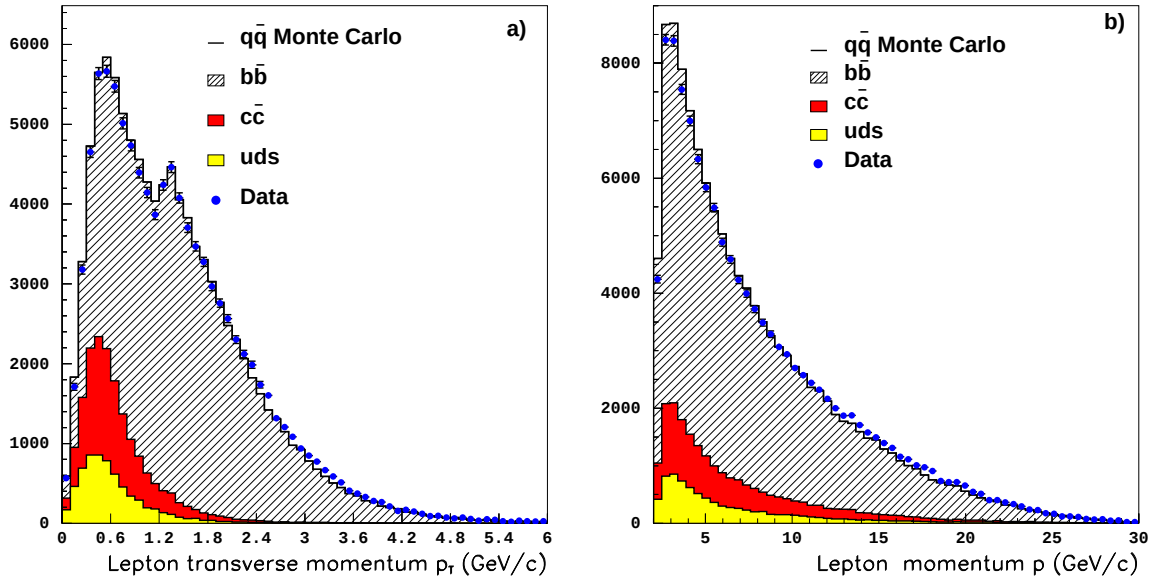

Figure 6.10: Lepton variables used for b-tagging. a) Lepton transverse momentum, b) lepton momentum

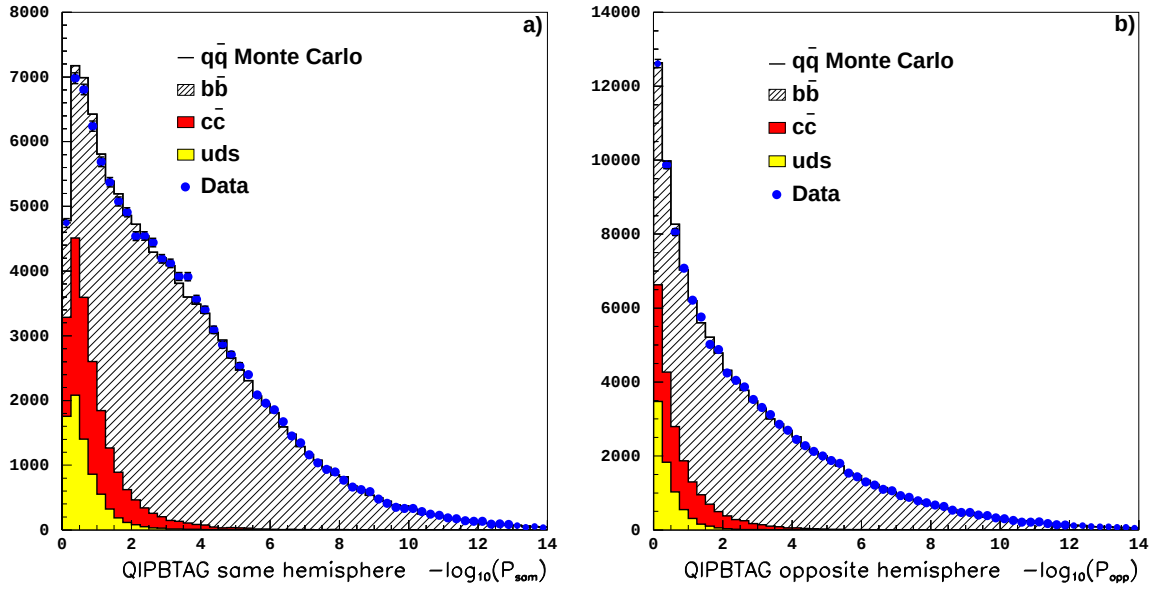

Figure 6.11: QIPBTAG variables used for b-tagging. a) Same side hemisphere probability, b) opposite side hemisphere probability, $-\log_{10}$ of the probability is plotted.

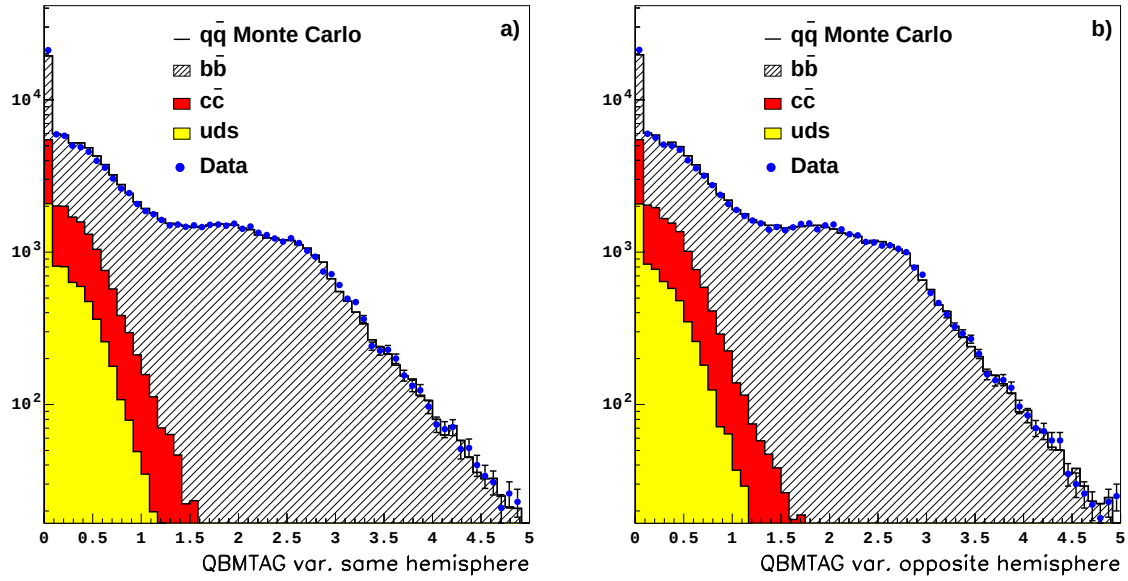

Figure 6.12: QBMTAG variables used for b-tagging. a) Same side hemisphere variable, b) opposite side hemisphere variable.

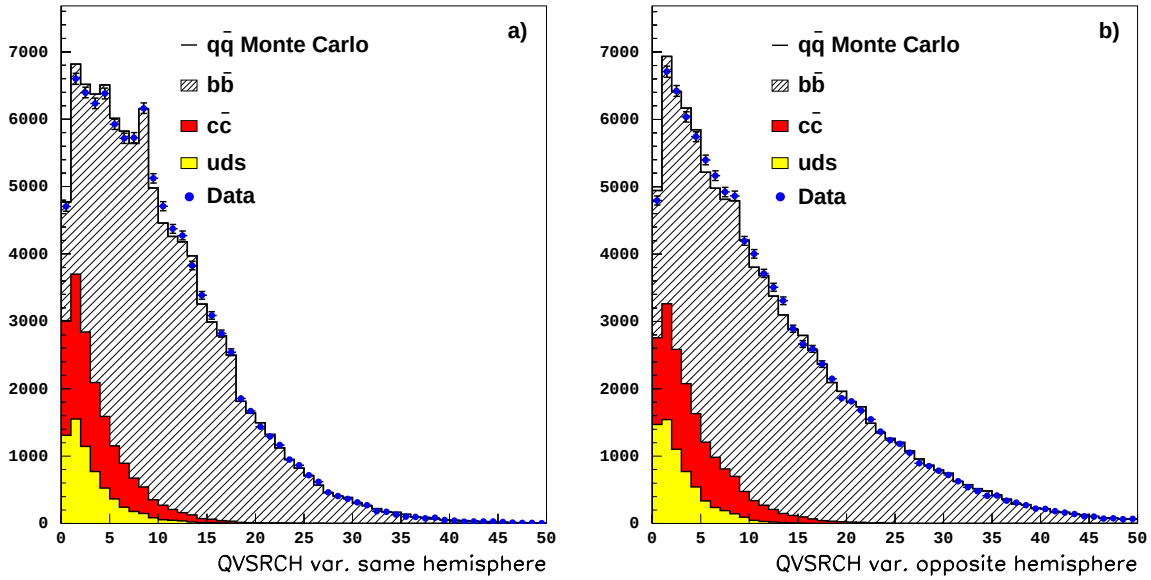

Figure 6.13: QVSRCH variables used for b-tagging. a) Same side hemisphere variable, b) opposite side hemisphere variable.

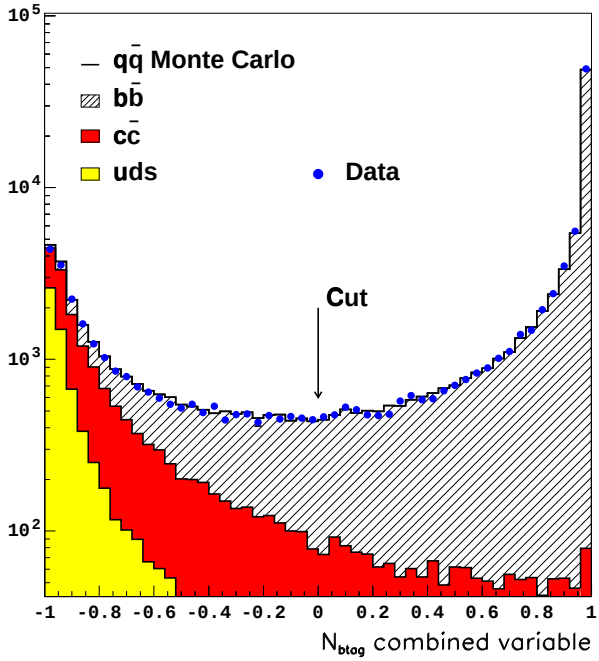

Figure 6.14: Combined b-tagging variable. The distribution of all the sample components is shown, as well as the agreement with the data.

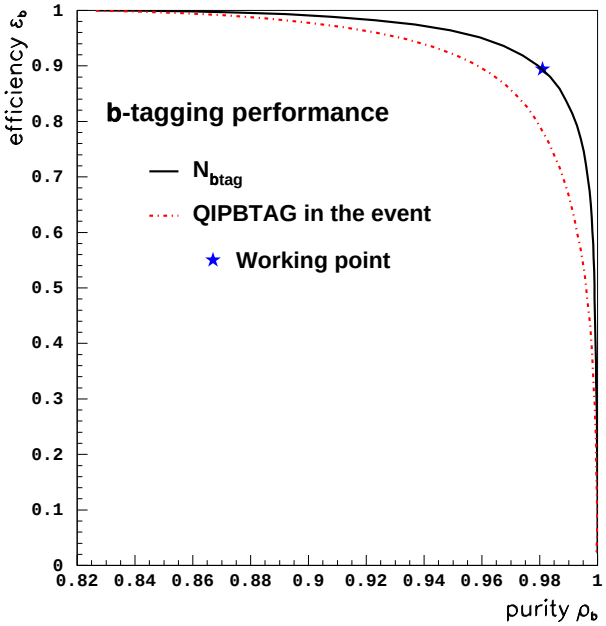

Figure 6.15: Comparison of b-tagging performance, the purity *versus* efficiency curve is plotted for the event QIPBTAG variable, and for that developed here. The working point chosen for the analysis is also indicated.

lepton plus vertex selected sample.

The improvement obtained upon the event QIPBTAG variable is mainly due to the fact that the b-tagging optimization, both for what concerns the selection of variables and what concerns the training, has been done on the specific events it is meant to be used on.

The probability of an event to originate from a $Z \to b\overline{b}$ decay is defined by the ratio of the sample components in Fig. 6.15 (the same probability definition is taken in the following). Events with very low probability of being a $Z \to b\overline{b}$ decay are rejected, an optimal cut is found at $N_{\rm btag} > 0$. The sample selected with this cut has the composition presented in Table 6.3. The selection efficiency for $Z \to b\overline{b}$ events, in a sample in which the lepton and vertex selections are applied, is estimated to be $\varepsilon_{\rm b} = 89.4 \pm 0.3\%$.

$(98.1 \pm 0.3)\%$ $b\overline{b}$	$(1.64 \pm 0.03)\%$ $c\overline{c}$	$(0.26 \pm 0.01)\%$ $u\overline{u}$, $d\overline{d}$, $s\overline{s}$

Table 6.3: Sample composition after the lepton, the vertex, and the b-tagging selections are applied.

The sample composition in $Z \to b\overline{b}$ or $Z \to u\overline{u}$, $d\overline{d}$, $s\overline{s}$, $c\overline{c}$ is parametrized as a function of the combined b-tagging variable $N_{\rm btag}$ on simulated events. An event-by-event signal purity is defined from this parametrization and then used in the oscillation fit (Section 8.1.1).

6.5 The b → ℓ enrichment

The sign of the electric charge of the lepton candidate is used to identify the b ($\bar{\rm b}$) content of the b-hadron at decay time. As explained in Section 3.1.2 (and Figs. 3.3, 3.4, and 3.5), the presence of a negatively (positively) charged lepton is taken as a tag for a b-hadron ($\bar{\rm b}$-hadron). However, a lepton is not guaranteed to be produced via a $b \to \ell$ quark decay. Cascade decays such as $b \to c \to \ell$ and $b \to \overline{c} \to \ell$ are also possible. In the case of $b \to c \to \ell$, the charge correlation between the lepton and the b-quark is opposite to that of $b \to \ell$ decays. The process $b \to \overline{c} \to \ell$ has the same charge correlation as the $b \to \ell$. In this Section, a method to statistically separate $b \to \ell$ decays from the other decays is presented.

Several topological differences distinguish direct semileptonic decays from the rest. Eight discriminant variables are identified here and combined with a neural network to obtain an optimal separation between $b \to \ell$ and $b \to c \to \ell$ decays.

The lepton momentum and transverse momentum spectra are expected to be harder for direct $b \to \ell$ decays than for other decays. The energy of the neutrino determined as in Ref. [84] (Section 6.7.1), is harder for direct semileptonic decays as well.

Some properties of the b-hadron jet which contains the lepton candidate are also considered. The decays of the type $b \to \ell$ ($B \to \ell\nu D$) and $b \to c \to \ell$ ($B \to WD(D \to \ell\nu X)$), where D refers to either D^0, D^+, D^*, or D^{**}, lead to significantly different jet topologies which are exploited here. The boost of the b-hadron, however, tends to dilute some of these differences. It is therefore more appropriate to study the separation in the rest frame of the lepton and the charmed particle, referred as the (ℓD) rest frame in the following. All the charged and neutral particles in the b-hadron jet (defined with the cone-based algorithm)

are boosted back in the (ℓD) rest frame. Two hemispheres are defined with respect to the plane orthogonal to the lepton direction. In Fig. 6.16, a schematic of the effect of the boost on typical topologies of b $\to \ell$ and b $\to$ c $\to \ell$ events is shown.

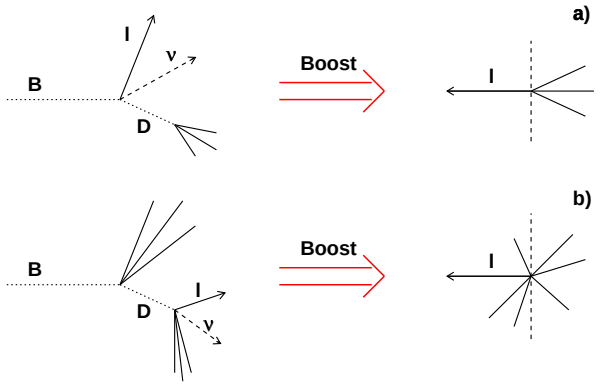

Figure 6.16: A schematic of b $\to \ell$ (a) and b $\to$ c $\to \ell$ (b) events topologies before and after applying a boost to the (ℓD) rest frame is shown. The plane used to define hemispheres in the (ℓD) rest frame is also displayed.

If no fragmentation particle were included in the b-hadron jet used to define the (ℓD) rest frame, the D and the lepton from the B $\to \ell\nu$D decay would be, by construction, produced back to back in this frame. In contrast, a lepton originating from the cascade decay B $\to$ WD(D $\to \ell\nu$X) is expected to be found close in phase space to the X system. The variable $E_{\rm CM}$ is defined as the sum of energies of the particles in the lepton hemisphere (except the lepton itself). A low value of $E_{\rm CM}$ is characteristic of a b $\to \ell$ decay.

A second shape variable, $P_{\ell \rm t}$, is designed to be sensitive to the two "directions of hadronization" present in most of the b $\to$ c $\to \ell$ decays (one direction for the c-quark decay and the other for the other b-hadron decay products), and not in b $\to \ell$ decays. This variable is defined as, $P_{\ell \rm t} = \frac{|P_+ - P_-|}{|P_+ + P_-|}$, where P_+ is the sum of the parallel momenta of the particles in the lepton hemisphere (with the (ℓD) rest frame definition) with respect to the lepton direction, and P_- is the equivalent quantity for particles in the hemisphere opposite to the lepton.

The total charged energy of the b-jet, and the lepton energy, both computed in the (ℓD) rest frame, also discriminate between b $\to \ell$ and b $\to$ c $\to \ell$ decays.

Finally, the signed impact parameter significance (see Section 5.5) of the lepton with respect to the charm vertex is computed, with the sign given by the charmed particle flight direction (Section 5.5). For events with two or more charged particles in the reconstructed charmed meson, the charm vertex (Section 6.3.2) is used. The inclusive tertiary vertex is used otherwise. The pull distributions for the two cases were checked, and the impact parameter uncertainty was corrected for the observed deviations. In the first case, the correction factor was found to be 1.19 and in the second 1.75. The lepton is expected to come prior to the charm vertex for b $\to \ell$ decays, as seen in Fig. 6.17, and therefore to have a negative impact parameter significance.

The distribution of all the variables described above are shown in Figs. 6.18 to 6.21 for

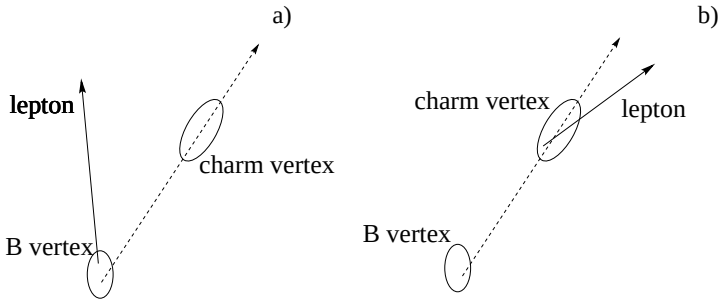

Figure 6.17: Schematic of the lepton impact parameter with respect to the charm vertex for a) a typical b → ℓ decay configuration with negative impact parameter with respect to the charm vertex, and b) a typical b → c → ℓ decay with, on average, zero impact parameter with respect to the same vertex.

events with a lepton candidate and a reconstructed secondary vertex. The simulated events are divided in four categories: *i*) events in which the lepton originates from light-quark decays or events in which a hadron is misidentified as a lepton, called "udsc and fake ℓ"; *ii*) events in which the lepton originates from a b → $\overline{c}$ → ℓ; *iii*) events in which the lepton originates from a cascade decay b → c → ℓ; and *iv*) the direct b → ℓ semileptonic decays. Agreement between data and simulation is observed for all variables.

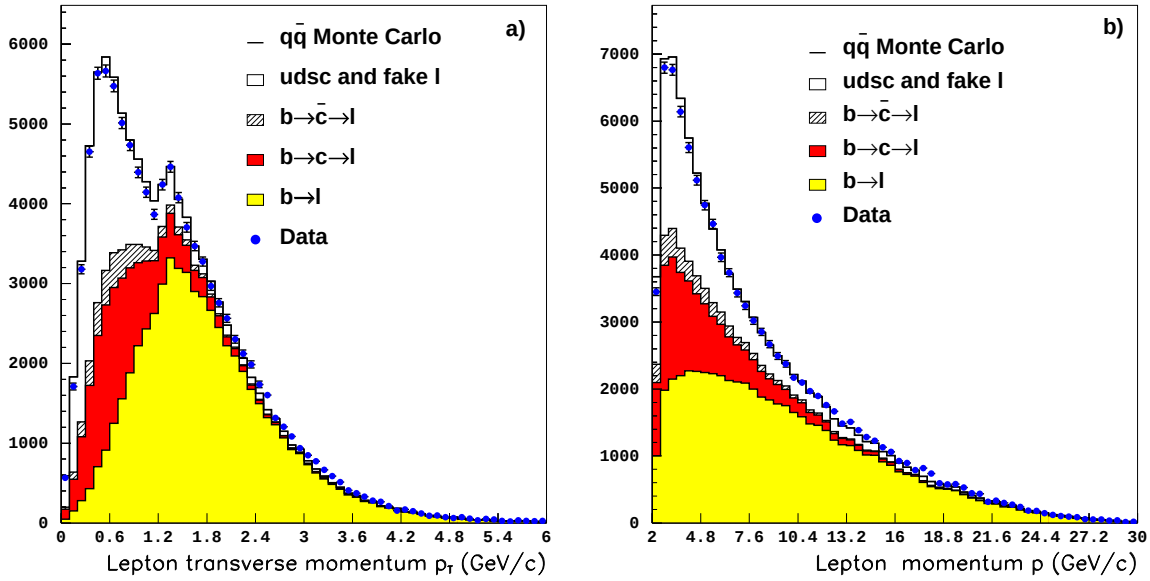

Figure 6.18: Lepton variables used for b → ℓ enrichment. a) Lepton transverse momentum, b) lepton momentum.

The b → ℓ enrichment variable combination was done with simulated B_s → ℓ events as signal and B_s → c → ℓ events as background, to obtain the best separation performance for the B_s oscillations analysis. The neural network was trained with a simulated sample in

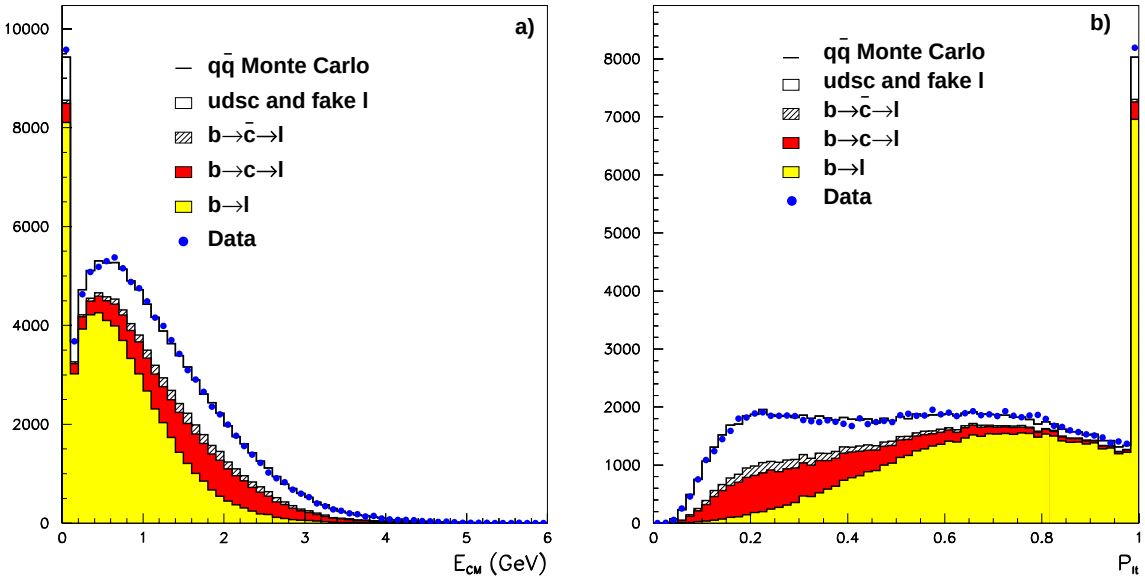

Figure 6.19: Shape variables used for b $\rightarrow \ell$ enrichment. a) Energy of the lepton hemisphere, with the lepton excluded, b) normalized sum of the track momenta parallel to the lepton.

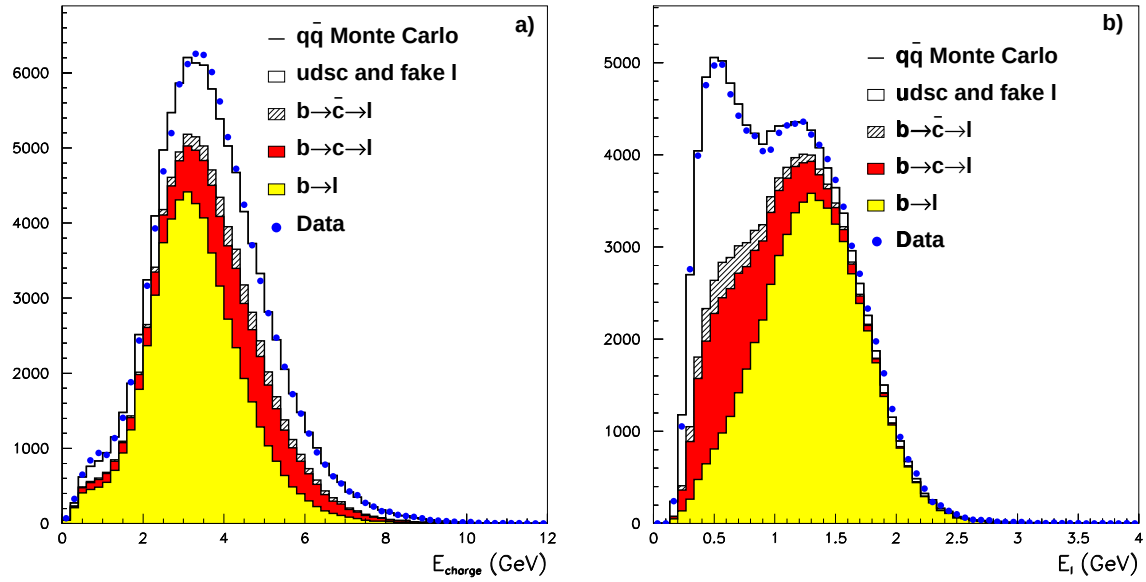

Figure 6.20: a) Charged energy of the b-jet, and b) lepton energy, both computed in the (ℓD) rest frame.

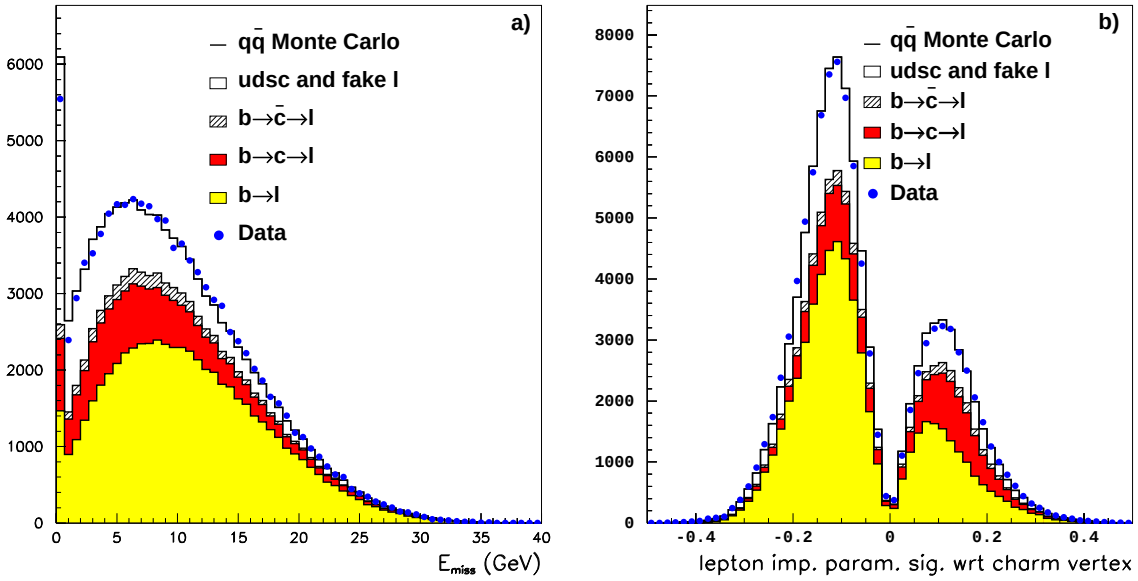

Figure 6.21: a) Neutrino energy, $E_{\rm miss}$, and b) lepton signed impact parameter significance. For the sake of visual clarity, the variable actually displayed is a function of the impact parameter significance $\mathrm{D}/\sigma_{\mathrm{D}}$ as: $\log\left[1+\left|\left(\log(1+|\mathrm{D}/\sigma_{\mathrm{D}}|)\right)\times\mathrm{sign(D)}\right|\right]\times\mathrm{sign(D)}$.

which only the lepton and the vertex selection criteria were applied.

The distribution of $N_{\rm bltag}$, the combined b → ℓ enrichment variable, is shown in Fig. 6.22, before and after the cut on $N_{\rm btag}$. In both cases a reasonable agreement between the data and the simulation is found.

A selection cut is applied on the $N_{\rm bltag}$ variable, optimized to obtain the highest statistical sensitivity for $\mathrm{B_s}$ oscillations. All events for which the probability for the lepton charge and the final state b-quark flavour to have the same correlation to that expected in b → ℓ decays is 50% or lower are rejected. The optimization is performed on simulated $\mathrm{B_s}$ decays and the cut is set at $N_{\rm bltag} > -0.5$.

The performance of the b → ℓ enrichment on $\mathrm{Z} \to \mathrm{b\bar{b}}$ simulated events are shown in Table 6.4. The sample composition and b → ℓ purity is given for the lepton plus vertex selected sample, after the b-tagging selection, and after both the b-tagging and the b → ℓ enrichment selections.

The performance of $N_{\rm bltag}$ can be compared to that of a simple cut in the lepton transverse momentum, as it was done in the previous analysis. The efficiency versus purity curve of a b → ℓ decays selection is displayed in Fig. 6.23, for a sample where the lepton and the vertex selection criteria are applied. The use of more information than the lepton transverse momentum brings a sizeable improvement on the b → ℓ enrichment performance.

The mistag probability due to b → c → ℓ decays is parametrized as a function of $N_{\rm bltag}$, and used event by event in the oscillation fit (Section 8.1.1) to increase the statistical power of the analysis. This parametrization is done independently for the four types of b-hadrons ($\mathrm{B_s}$, $\mathrm{B_d}$, $\mathrm{B^+}$, and b-baryons) because the b → c → ℓ rejection performance is not the same

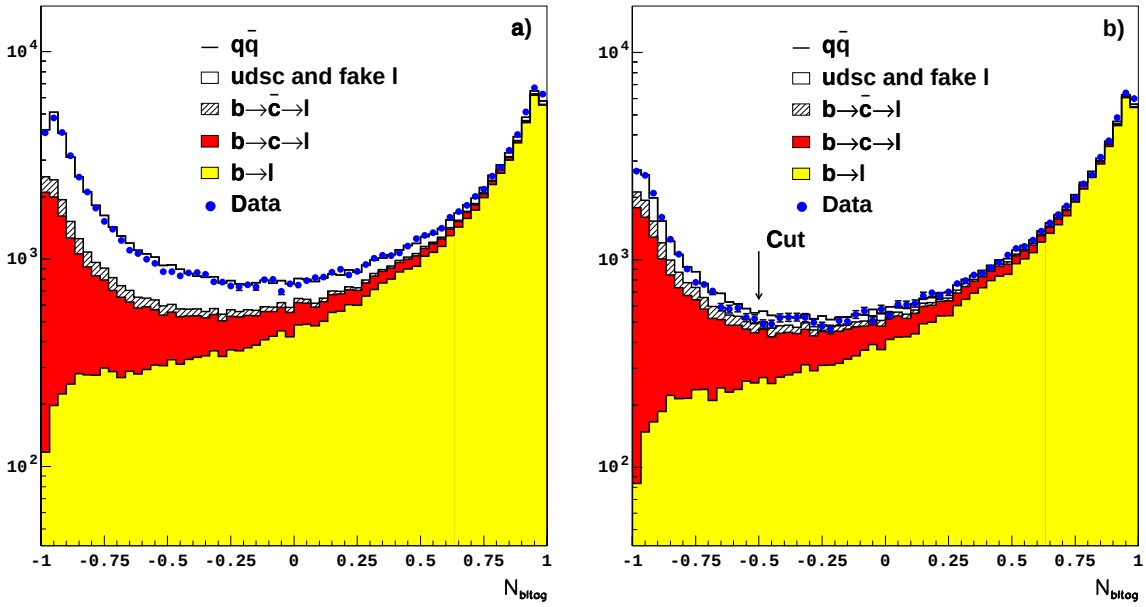

Figure 6.22: Combined b $\to \ell$ enrichment variable distribution, a) before the b-tagging cut, and b) after the b-tagging cut is applied.

	lepton+vertex	+ b-tagging	+ b $\to \ell$ enrichment
b $\to \ell$ purity on b$\overline{\text{b}}$	$(67.5 \pm 0.3)\%$	$(69.6 \pm 0.3)\%$	$(86.0 \pm 0.4)\%$
b $\to$ c $\to \ell$ fraction on b$\overline{\text{b}}$	$(22.6 \pm 0.1)\%$	$(21.1 \pm 0.1)\%$	$(9.7 \pm 0.1)\%$
b $\to \overline{\text{c}} \to \ell$ fraction on b$\overline{\text{b}}$	$(3.96 \pm 0.05)\%$	$(3.71 \pm 0.05)\%$	$(1.70 \pm 0.04)\%$
other lepton candidates	$(5.94 \pm 0.06)\%$	$(5.59 \pm 0.06)\%$	$(2.6 \pm 0.04)\%$
"correct" lepton sign fraction	$(71.5 \pm 0.2)\%$	$(73.3 \pm 0.3)\%$	$(87.7 \pm 0.3)\%$

Table 6.4: Performance of the b $\to \ell$ enrichment on Z $\to$ b$\overline{\text{b}}$. In the last row, the fraction of events in which the lepton charge has the "correct" correlation with the b-hadron flavour is indicated.

for all of them. Differences are also seen upon the vertex classes defined in Section 6.3.7, and are taken into account. For illustration, the b $\to \ell$ purity is shown in Fig. 6.24 for B_s simulated events in different vertex classes.

6.6 Final selection

The b-tagging and the b $\to \ell$ enrichment variables (N_{btag}, N_{bltag}) are treated independently in the previous two Sections, both for the cut definitions and for the parametrization of the b and b $\to \ell$ event-by-event purities. Although these variables are actually not independent, a combined treatment was found not to bring a significant improvement, and it was dropped for simplicity.

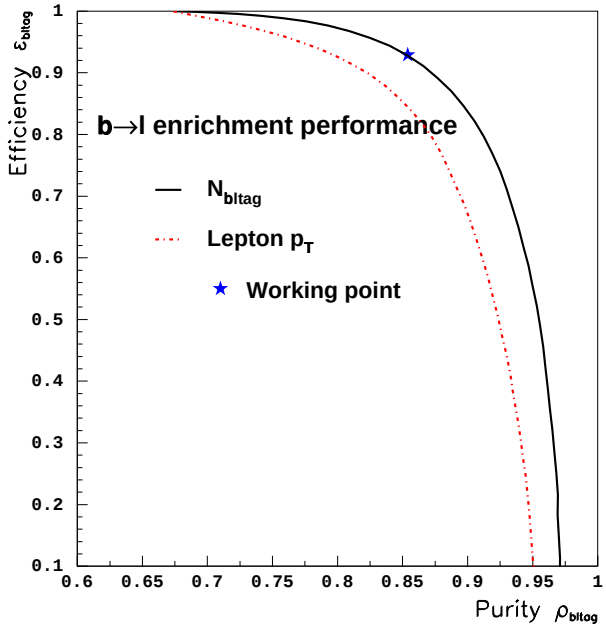

Figure 6.23: Comparison of b $\rightarrow \ell$ enrichment performance. The purity *vs* efficiency curve is displayed for the lepton p_T cut (dashed curve), and for a cut on the N_{bltag} variable (full curve).

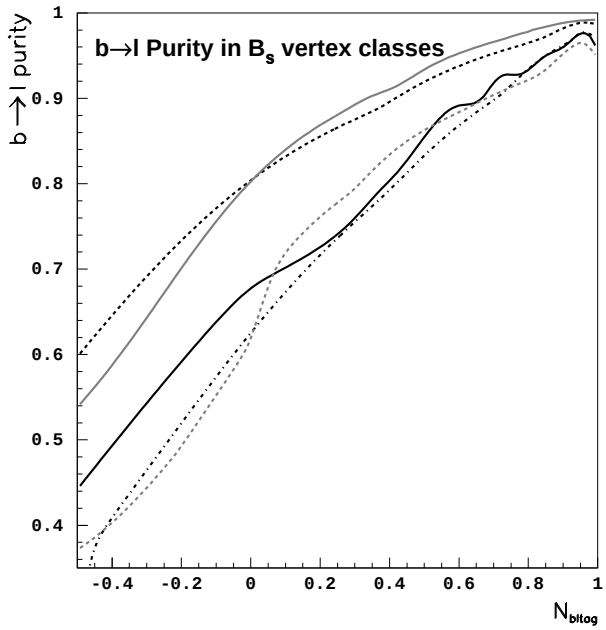

Figure 6.24: b $\rightarrow \ell$ purity as a function of N_{bltag} for $\mathrm{B_s}$ events in different vertex classes.

The final event sample is defined by the lepton selection, the vertex selection, the b-hadron selection, and the $b \to \ell$ decays selection. After all these criteria are applied, 74026 events are selected in the data. The sample composition is evaluated from simulated $Z \to q\overline{q}$ events as in Table 6.5.

$(98.6 \pm 0.4)\%$ $b\overline{b}$	$(1.15 \pm 0.03)\%$ $c\overline{c}$	$(0.25 \pm 0.02)\%$ $u\overline{u}$, $d\overline{d}$, $s\overline{s}$

Table 6.5: Final Sample composition.

The selected b-hadron hemispheres are divided in four different classes according to the process which originates the lepton candidate, as shown in Table 6.4.

The number of events selected in the data sample can be compared to the 79867 events expected from the $Z \to q\overline{q}$ simulated samples. An excess of 8% is found in the simulation. No significant difference is observed between the electron and muon selections. Some of the branching ratios for the production of b-hadrons have measured uncertainties of the order of a few percent. The corresponding systematic uncertainties are discussed in Section 8.2. The overall selection efficiency also depends on the average b-hadron lifetime, but the analysis is to first order insensitive to that. From this point onwards in this Chapter, all data/simulation comparisons are made with the number of events in the simulation normalized to that found in the data.

The vertex class population is shown in Fig. 6.25; the data distribution is compared to the expectation from the simulation. The fraction of B_s decays expected for each class is shown as well (Fig. 6.25c). A significant difference on the B_s fraction is observed between the vertex classes, in particular, the 3 classes with an odd number of charged particles in the charmed object have a B_s fraction close to 13%.

6.7 Proper time measurement

The proper time of the b-hadron at decay is computed as

$$t \;=\; l\, m_{\rm B} \,/\, p_{\rm B}\, c \;, \tag{6.4}$$

where l is the decay length estimated as explained in Section 6.3.6, $m_{\rm B}$ is the mass of the b-hadron, (taken to be $5.3\,{\rm GeV}/c^2$) and $p_{\rm B}$ is the estimate of the b-hadron momentum performed as described in the next Section.

6.7.1 Momentum measurement

The b-hadron momentum is estimated from three ingredients, *i*) the energy of the lepton candidate; *ii*) the energy of the charmed particle; and *iii*) the estimate of the neutrino energy.

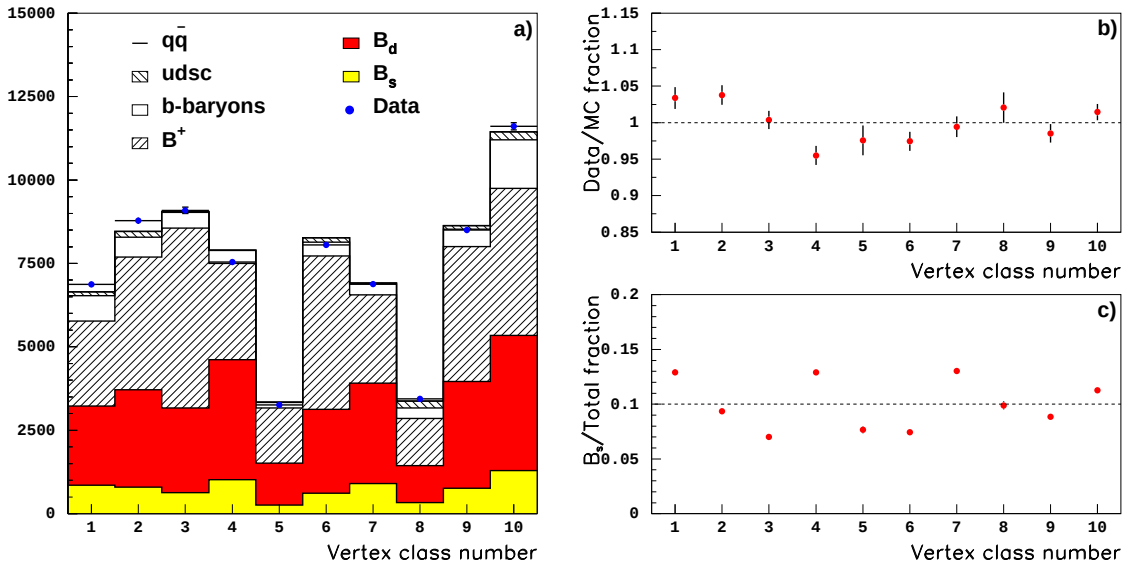

Figure 6.25: a) Distribution of the events in the selected sample in the vertex classes. The simulation histogram is normalized to the number of events selected in data.
b) Fraction of data selected events with respect to the simulation expectations for each class; some class-to-class differences are seen, but are compatible with being due to statistical fluctuations.
c) Fraction of B_s events in each class.

Charmed particle energy

The energy of the charmed particle, E_c, is estimated from a nucleated jet (Section 5.2.2) clustered around the charged particles at the charm vertex until a mass of $2.7\,\mathrm{GeV}/c^2$ is reached. Particles with momentum smaller than $0.5\,\mathrm{GeV}/c$ are excluded because they are more likely to come from fragmentation than from the b-hadron decay. This method was already used in Ref. [10].

Neutrino energy

The energy of the neutrino, E_{miss}, is estimated from the total energy and momentum conservation. If a neutrino in the lepton hemisphere is considered to be the only contribution to the unmeasured momentum in the event, the following equation holds,

$$E_{\mathrm{miss}} = \frac{\sqrt{s}}{2} + \frac{m_1^2 - m_2^2}{2\sqrt{s}} - E_1 \ . \tag{6.5}$$

The hemisphere masses m_i are determined from the measured energies E_i and momenta p_i ($m_i^2 = E_i^2 - \vec{p_i}^{\,2}$), where $i = 1$ denotes the lepton hemisphere, $i = 2$ the opposite hemisphere, and $\sqrt{s}$ is the centre-of-mass energy. This method was first used in ALEPH in Ref. [84].

The neutrino energy resolution obtained on $b \to \ell$ simulated events is shown in Fig 6.26.

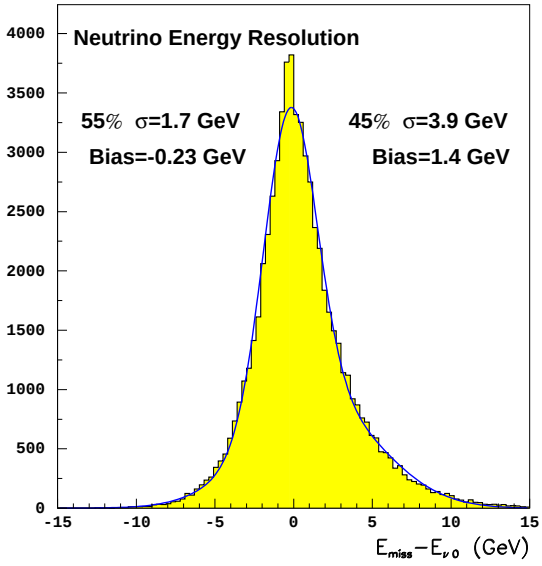

Figure 6.26: Neutrino energy bias on b $\to \ell$ simulated events. The distribution is fit with the sum of two Gaussians, their fractions and widths are indicated on the plot.

The b-hadron momentum

Once the energy of the charmed particle and the neutrino energy are determined, the b-hadron momentum is obtained as

$$p_{\mathrm{B}} \;=\; \sqrt{(E_{\mathrm{c}} + E_{\mathrm{miss}} + E_{\ell})^2 \;-\; {m_{\mathrm{B}}}^2} \;, \tag{6.6}$$

where E_ℓ is the energy of the lepton.

The b-hadron momentum obtained with this method is found to have a bias with respect to the true momentum, increasing with the measured neutrino energy. The estimate of the charmed particle energy was optimized to obtain the best b-hadron momentum resolution possible. However this optimization was performed independently of the neutrino energy, this fact explains the bias observed. A correction is applied to account for this effect. Sizeable differences on the bias are observed upon the vertex classes. A specific correction is therefore applied to each of them. The typical dependence of the momentum bias on the neutrino energy is shown in Fig. 6.27.

The improvement brought by the bias correction is shown in Fig. 6.28. The correction reduces the global bias of the distribution and also improves the resolution.

The average relative momentum resolution for all selected events in a sample of simulated $\mathrm{B_s} \to \ell$ events is shown in Fig. 6.28b: 60% of the events are found in a core with 6.4% relative momentum resolution, and 40% in a tail with 20% resolution. However, these values are not used for the oscillation fit. The momentum resolution cannot be measured event by event and a parametrization is needed. The vertex classes are also used here to divide the sample in phase space regions which contain events with similar properties. For each class, the sample is divided in bins of lepton momentum. A double Gaussian fit is performed to the relative

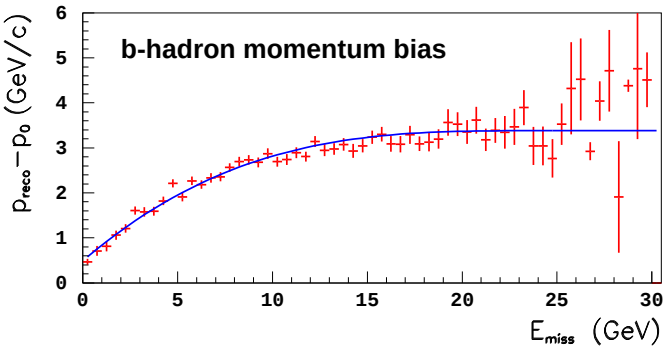

Figure 6.27: Average bias on the b-hadron momentum as a function of the neutrino energy.

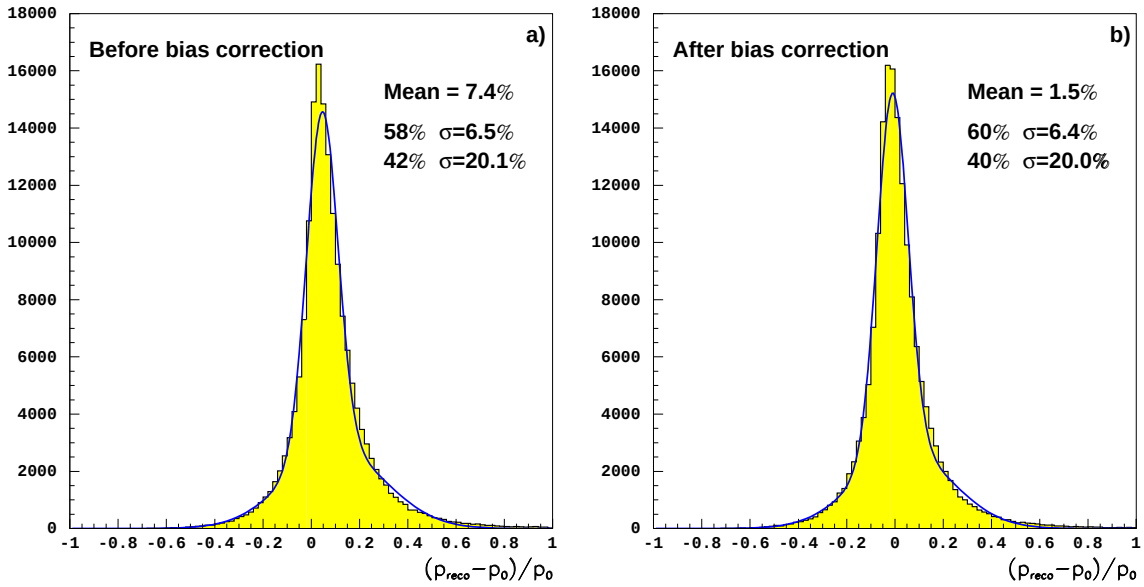

Figure 6.28: Relative momentum bias, a) before, and b) after average-bias correction. The distributions are fit with the sum of two Gaussians. Their fractions and widths are indicated on the plot along with the mean of the distributions.

momentum resolution distribution in each of these bins. The dependence of the widths of these Gaussians on the lepton momentum is parametrized with a straight line.

The relative momentum resolution distributions for two classes are given in Fig. 6.29 for illustration. Figure 6.29a contains events from a class with very good resolution (60% core at 5% and tails at 13%) and Fig. 6.29b from another class with poor resolution (53% core at 7% and tails at 22%). With the parametrization explained above, a more refined treatment of the events is achieved and the best 10% events in the first class are found to have a core resolution of 4% and tail resolution of 8% (with core and tails in the same proportion of 60% to 40%).

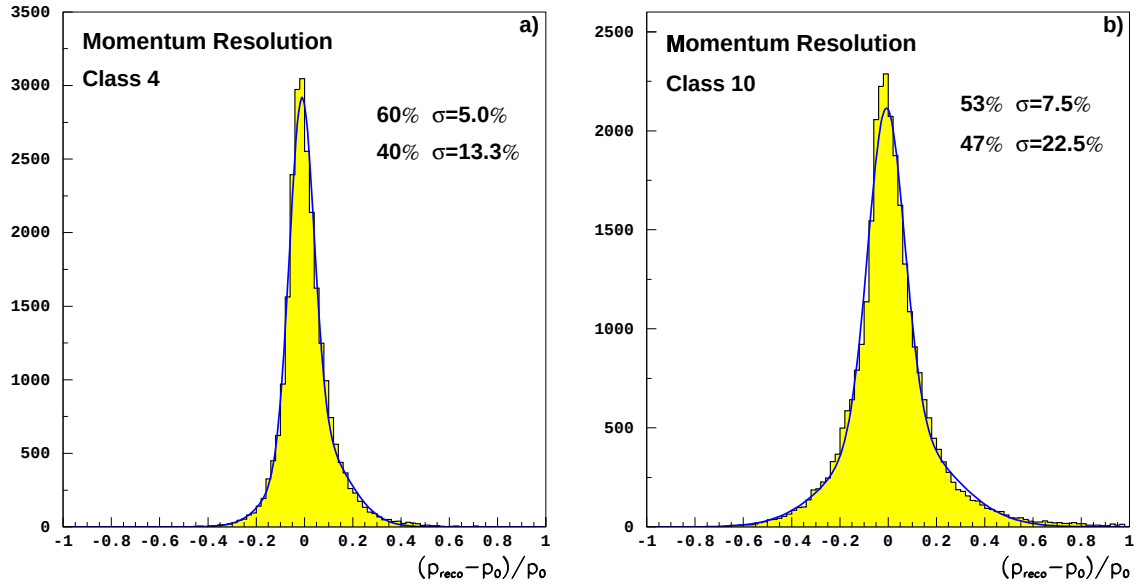

Figure 6.29: Relative momentum resolution, a) vertex class 4, and b) vertex class 10. The distributions are fit with the sum of two Gaussians. Their fractions and widths are indicated on the plot.

All corrections and parametrizations, both for the momentum measurement and for the estimation of its uncertainty, were performed on $B_s \to \ell$ simulated events.

6.7.2 Proper time

The estimate of the proper time for every selected event is obtained with Eq. 6.4 and the estimates of the reconstructed decay length and momentum. The proper time distribution for all selected events is shown in Fig 6.30, the simulation distribution is in agreement with that of the data.

The estimate of the proper time uncertainty is obtained from the convolution of the decay length uncertainty and the relative momentum uncertainty (Sections 6.3.7 and 6.7.1) as:

$$\sigma_t = \sqrt{\left(\frac{m_B}{p\,c}\,\sigma_l\right)^2 + \left(t\,\frac{\sigma_p}{p}\right)^2}\,. \tag{6.7}$$

The distribution of the estimated proper time uncertainty, σ_t, for the selected data sample is shown in Fig 6.31.

6.7.3 Proper time efficiency

The vertex selection efficiency (Section 6.3) is a function of the decay length of the b-hadron candidate. This dependence results in a non-flat proper time efficiency as a function of the

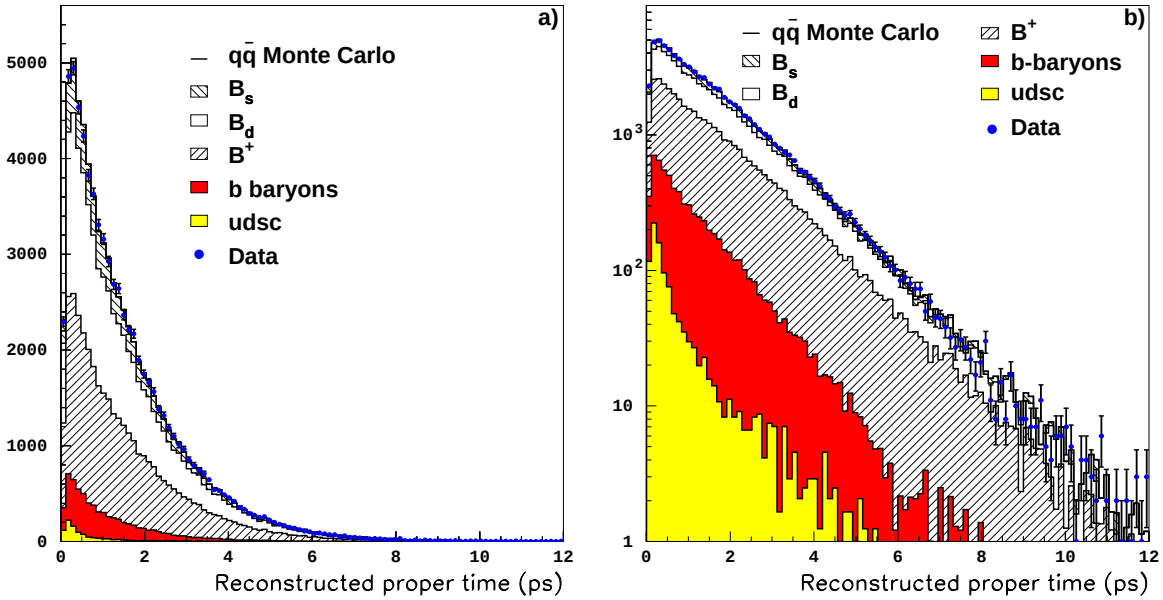

Figure 6.30: Proper time distribution in the selected sample.

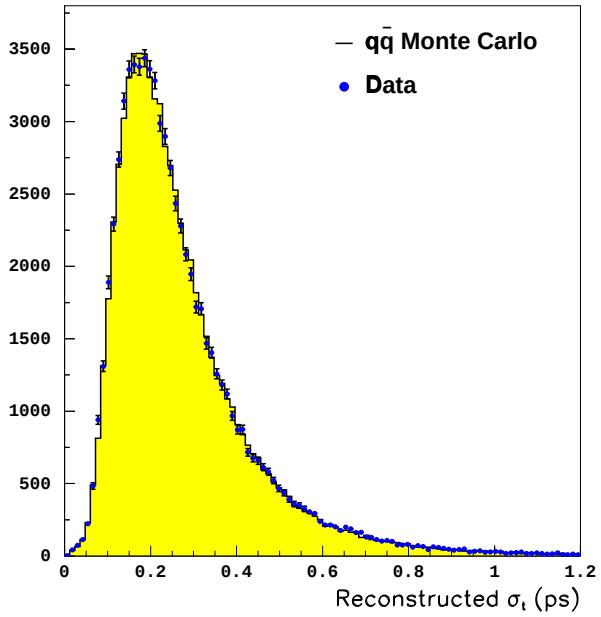

Figure 6.31: Distribution of the uncertainty on the proper time in the selected sample.

true proper time. The shape of the proper time efficiency needs to be taken into account in the oscillation fit. To this end, it has been parametrized for each vertex class and each b-hadron species.

As an illustration, the global (all vertex classes together) proper time efficiency as a function of the true proper time is shown in Fig. 6.32 for $B_s \to \ell$ signal Monte Carlo events.

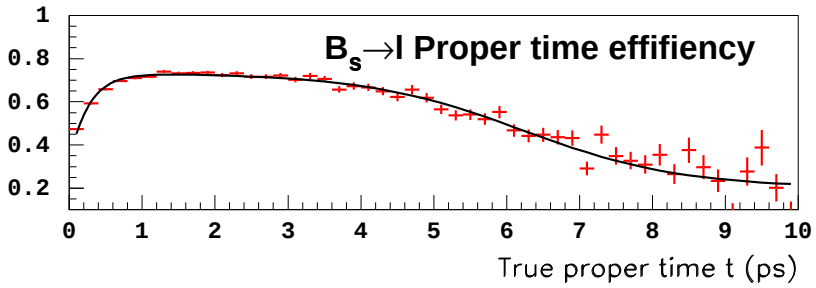

Figure 6.32: Proper time efficiency for $B_s \to \ell$ simulated events, with the parametrization superimposed.

6.8 Kaon identification

The presence of kaon candidates is used in this analysis for the B_s enrichment (Section 6.9) and for the initial state tag (Section 6.10). Charged and neutral kaons are used in the first case, and only charged kaons in the second. In this Section, the procedure used for kaon identification is explained. A jet clusterization with the JADE algorithm and $y_{\rm cut} = 0.02$ is performed to obtain jets which contain all the b-hadron decay products and the fragmentation particles closest in phase space with the b-hadron. In the B_s decay hemisphere, all kaons are searched for inside the jet which contains the lepton and the charmed particle. This jet is called "big" jet in what follows.

6.8.1 Charged kaons

Charged kaons are searched for in the two event hemispheres. In the B_s candidate hemisphere, kaons issued from both the B_s fragmentation process and from its decay can be exploited for flavour tagging and/or B_s enrichment. In the opposite hemisphere, only decay kaons are relevant for flavour tagging.

Charged kaons are distinguished from other charged particles by their dE/dx measurement in the TPC. The separation between fragmentation and decay kaons is done with the same variables used for the general track separation (Section 6.3.2). However, for those events where a charm vertex is reconstructed, the impact parameter significance with respect to this vertex (rather than the tertiary vertex seed) along the charmed particle flight direction is taken as a variable to measure its compatibility with originating from the b-hadron decay.

A neural network is used to combine the dE/dx estimator with the track separation variables. A different treatment is introduced for charged particles included in the charm track selection and for the others. In the first sample, only decay kaons are expected, because the contamination from fragmentation particles can be neglected. A special network is trained on this track sample. In the case of tracks not included in the charm selection, the network is trained to distinguish three track classes: charged kaons from fragmentation, charged kaons from the B_s decay, and non-kaon particles. Both neural network trainings are performed using tracks included in the "big" jet of $B_s \to \ell$ selected simulated events. The combined estimator distributions for fragmentation kaons, and for decay kaons in the charm track selection are shown in Fig 6.33.

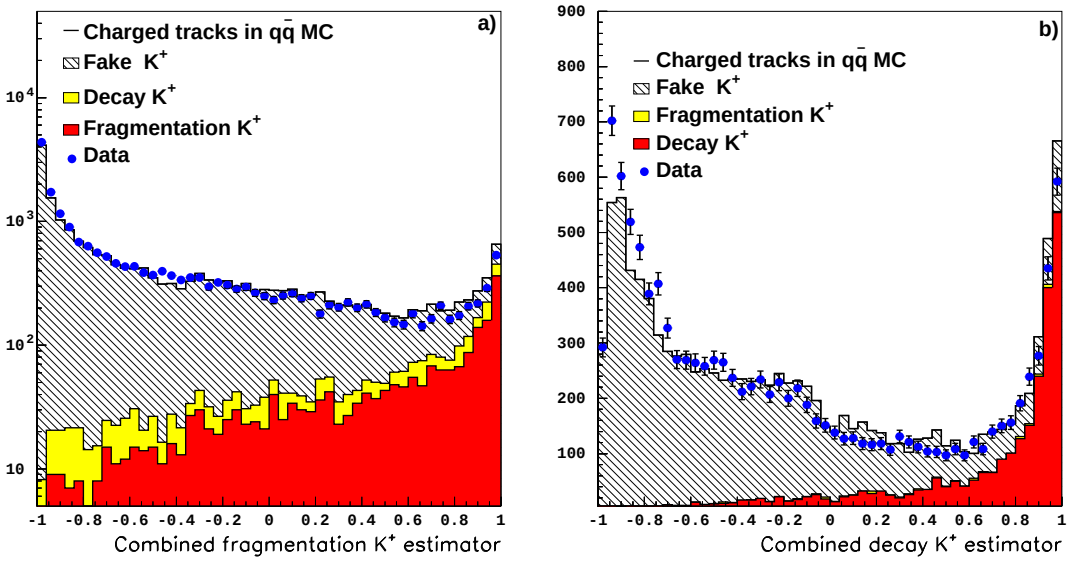

Figure 6.33: Charged kaon estimators, a) for a fragmentation kaon, and b) for a kaon from the b-hadron decay in the charm track selection.

The two estimators for the decay kaon are parametrized to obtain a probability for each track to be a kaon from the c-hadron decay. This probability is the estimator used in the B_s enrichment (Section 6.9).

For the opposite side hemisphere, a new estimator is built to identify kaons from the decay of any b-hadron. The same set of discriminant variables is used, and the training is done on unbiased b hemispheres.

6.8.2 Neutral kaons

Neutral kaons, K_S, both issued from the b quark fragmentation process and from the meson decay, are searched for in the B_s candidate hemisphere. Only the reconstruction of K_S into two charged pions is attempted.

In the event reconstruction, long-lived neutral particle candidates (V_0) are identified as two oppositely charged particles, each of them incompatible with originating from the primary

vertex, but compatible with being the decay products of a known neutral long-lived particle originating from the primary vertex [66]. In each event, K_S candidates are looked for in the V_0 selection. Several variables distinguish a true from a fake K_S, their list follows. *i*) The mass of the reconstructed neutral particle; *ii*) its distance of flight from the primary vertex; *iii*) the χ^2 of the fit of each pion with the primary vertex; *iv*) the χ^2 of the K_S mass constrained fit; *v*) the reconstructed K_S momentum; *vi*) the distance of closest approach between the helices of the two charged pions; and *vii*) the angle between the reconstructed K_S momentum and its flight direction. The distributions of all these variables, for V_0 candidates in data and simulation are shown in Figs. 6.34 to 6.37.

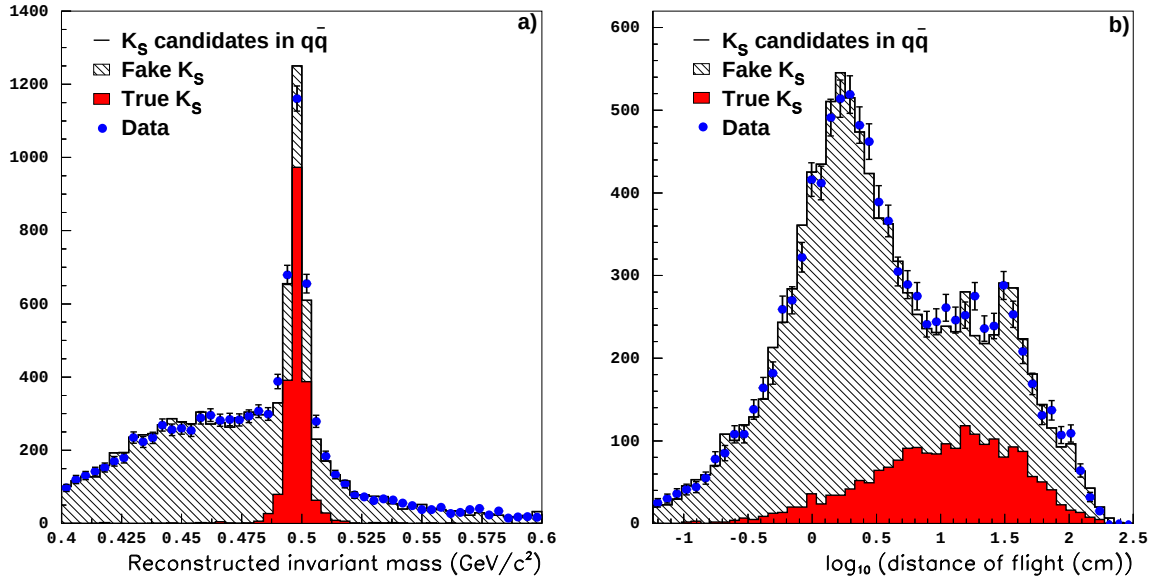

Figure 6.34: K_S estimators a) K_S reconstructed mass, and b) K_S distance of flight (the logarithm of the distance is plotted).

Only K_S candidates with a reconstructed mass within $30\,\mathrm{MeV}/c^2$ from the nominal value ($m_{K_S} = 497.672 \pm 0.031\,\mathrm{MeV}/c^2$ [13]) are considered. The discriminant variables listed above are combined with kinematical variables (estimators of the angle with the b-jet, and compatibility with the primary and tertiary vertices, as for the track separation described in Section 6.3.2) to obtain an estimator for both fragmentation and decay neutral kaons. The combination is performed with a neural network, the training of which is performed on K_S candidates in $B_s \to \ell$ selected simulation events. The combined neutral kaon estimator distributions for kaons issued from the b fragmentation process and from the b-hadron decay are displayed in Fig. 6.38.

With a cut at zero on each estimator, 90% of the pion pairs selected are found to originate from K_S's. Out of those, 70% originate from the b-hadron decay in the case of the decay K_S estimator, and 90% come from the b quark fragmentation in the case of the fragmentation K_S estimator.

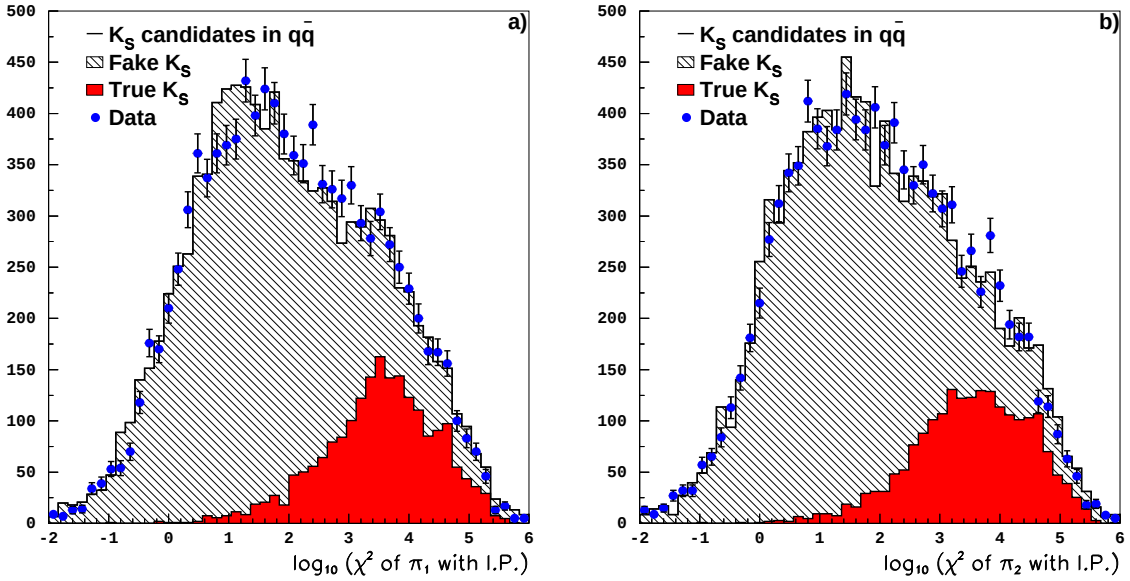

Figure 6.35: The χ^2 of the two decay pions with the primary vertex (the logarithm of the χ^2 is plotted).

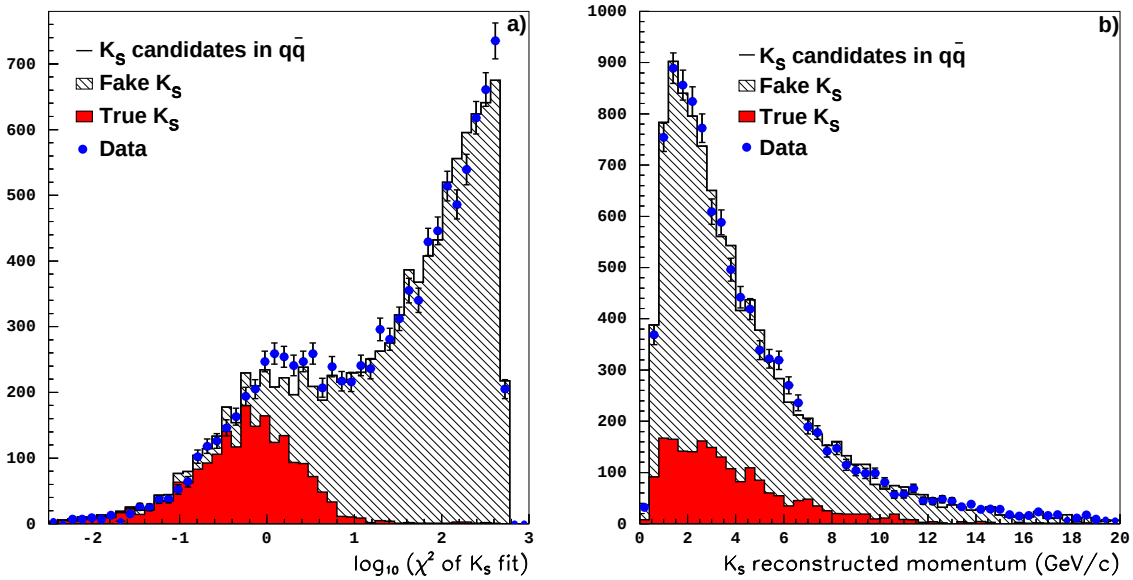

Figure 6.36: K_S estimators a) χ^2 of the K_S mass constraint fit (the logarithm is plotted), and b) K_S reconstructed momentum.

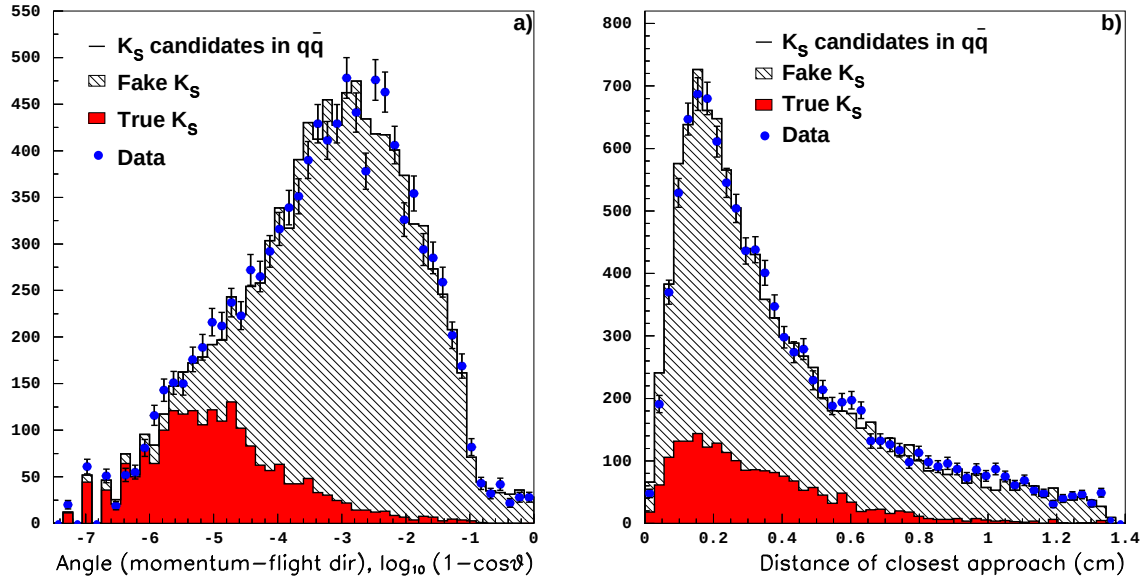

Figure 6.37: K_S estimators a) angle between the momentum and the flight direction of the K_S candidate, and b) distance of closest approach between the two pion helices.

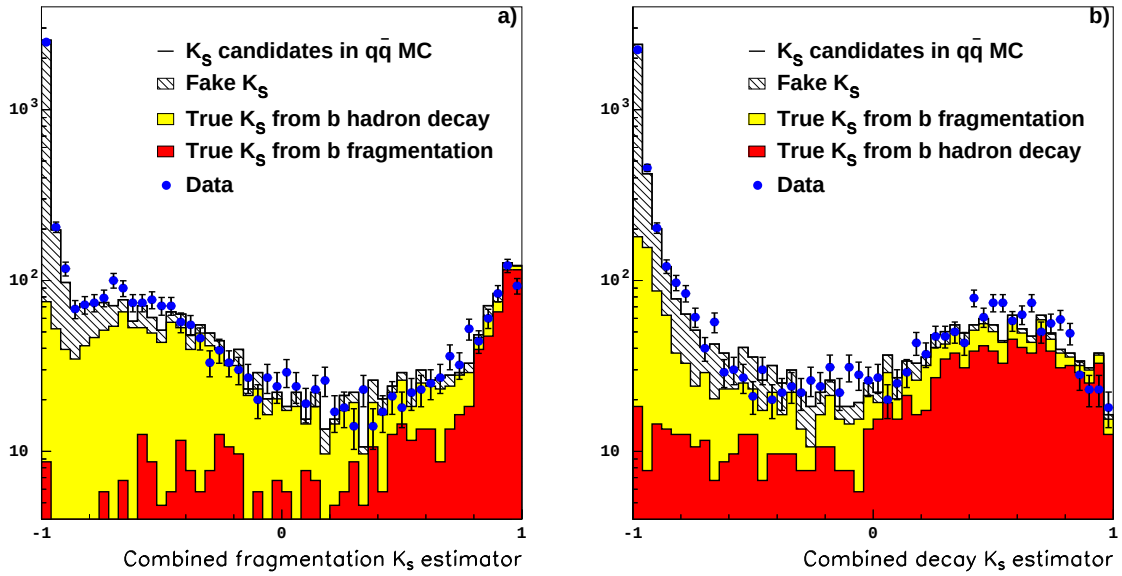

Figure 6.38: Combined neutral kaon estimators, a) for K_S from fragmentation, and b) for K_S from the b-hadron decay.

6.9 B_s enrichment

In an unbiased b-hadron sample only about 10% of the hemispheres contain a B_s [13]. The sensitivity of the oscillation analysis can be improved through the identification of variables which discriminate B_s from other b-hadron decays, and the evaluation, event by event, of the probabilities that a given candidate belongs to each b-hadron species.

The B_s enrichment proceeds in two streams: *i*) variables related to the electric charge and the charge particle multiplicity of the secondary and tertiary vertices distinguish charged from neutral b-hadrons; *ii*) the presence of kaons among fragmentation and decay particles distinguish B_s decays from the other b-hadrons. For the charged decay kaons, the charge correlation with the lepton candidate is the best discriminant.

A set of discriminant variables is identified and combined with a neural network to obtain an optimal separation performance. The description of the variables in the two streams follows.

Separation between charged and neutral b-hadrons

The number of charged particles in the charm vertex distinguishes between neutral and charged b-hadrons. With the assumption of perfect charged particle separation between those originated from the b-hadron decay and the others, an odd number of charged particles is expected in the charm vertex for neutral b-hadrons, and an even number for charged b-hadrons.

A new charged particle separation variable is built, based on the same particle properties as that described in Section 6.3.2. Two modifications are introduced to improve its performance. On one hand, the neural network training is done exclusively on events which fulfil all the selection criteria described in the previous Sections, and on the other, the impact parameter significance with respect to the charm vertex (if any) is taken as a variable to measure the compatibility of a charged particle with originating from the b-hadron decay. If no charm vertex is reconstructed, the impact parameter significance with respect to the inclusive tertiary vertex is used, as in Section 6.3.2. This track separation variable is used to build a set of "weighted" charm vertex estimators: the secondary vertex charge, the weighted vertex multiplicity, and the charge weighted momentum. The definition of these three estimators follows.

$$\text{Vertex charge} = q_\ell \sum_i^{N_{\rm ch}} w_i^{\kappa_1}\, q_i \ , \tag{6.8}$$

$$\text{Vertex multiplicity} = \sum_i^{N_{\rm ch}} w_i^{\kappa_2} \ , \tag{6.9}$$

$$\text{Charge weighted momentum} = q_\ell \sum_i^{N_{\rm ch}} w_i\, p_i^{\kappa_3}\, q_i \ , \tag{6.10}$$

In the three expressions, $N_{\rm ch}$ is the number of charged particles in the hemisphere, w_i is the value of the track separation variable for the i-th particle (expected to be close to 1 for true decay particles and close to 0 otherwise), p_i is the momentum of the i-th particle, q_i is the

charge of the i-th particle, q_ℓ is the charge of the lepton, and the κ_j values are optimized to obtain the best separation performance, ($\kappa_1 = 1.6,\ \kappa_2 = 2.2,\ \kappa_3 = 1.1$). The vertex charge is expected to be peaked at -1 for neutral b-hadrons (the charmed particle and the lepton have opposite electric charge sign) and at zero otherwise. The vertex multiplicity is expected to be concentrated around integer odd values for neutral b-hadrons. The charge weighted momentum distribution is expected to have a negative mean for neutral b-hadrons, and to be centred at zero otherwise.

The distributions of these variables in data and simulation are shown in Figs. 6.39 and 6.40. The fraction of B_s events in the simulation as a function of each variable is also displayed.

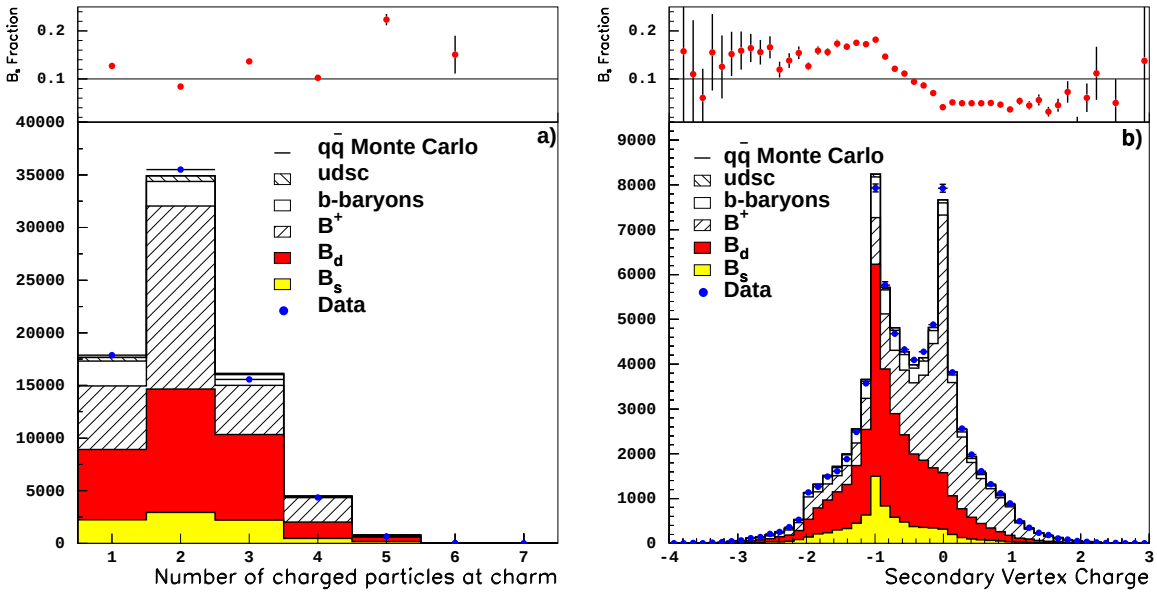

Figure 6.39: a) Number of charged particles at the charm vertex, and b) charm vertex charge estimator.

Separation between B_s and other neutral b-hadrons

The presence of a kaon among the fragmentation particles is an indication of a B_s decay. The fragmentation kaon can be charged or neutral. A variable is defined to identify charged kaons and another for K_S issued from the fragmentation process. For the charged fragmentation kaon, all tracks inside the "big" b-jet (Section 6.8) and not included in the charmed track selection are considered. The track with the highest value of the fragmentation kaon estimator (Section 6.8.1) is chosen and its estimator value is taken as a discriminant variable for the B_s enrichment. A similar treatment is applied to identify neutral kaons. The distributions of the fragmentation kaon estimators in data and simulation, and the fraction of B_s events as a function of the these estimators, are shown in Fig. 6.41.

The decay products of a semileptonic B_s decay include, in most cases, two kaons from the D_s^- decay, while the B_d decay products include typically a single kaon at most (apart from

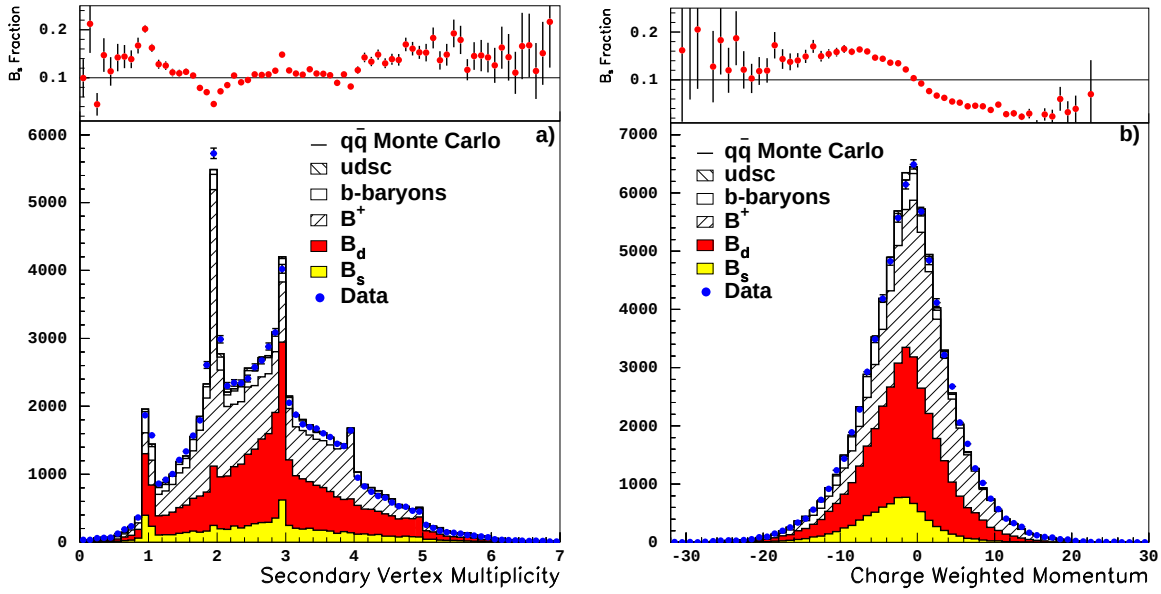

Figure 6.40: a) Charm vertex weighted multiplicity, and b) charm vertex charge weighted momentum estimators.

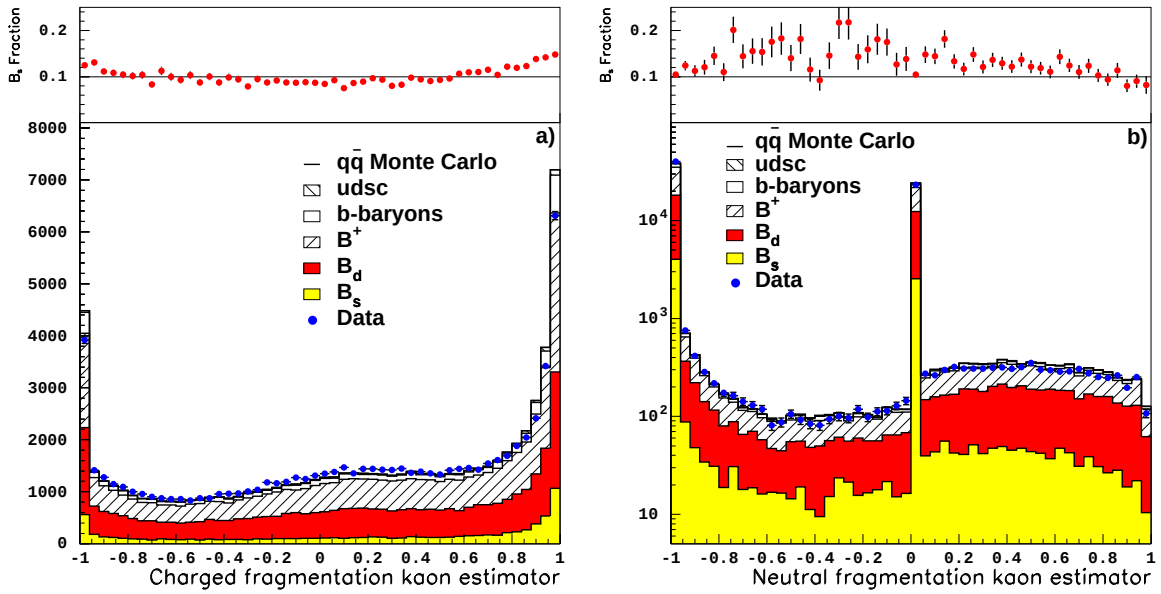

Figure 6.41: Fragmentation kaon estimators, a) charged kaon candidate, and b) neutral kaon candidate.

rare decays, thus not relevant here). The number of kaons found in the decay products of the b-hadron therefore distinguishes B_s decays from other b-hadron decays. Moreover, charged kaons from B_d decays have the same electric charge sign as the lepton, while in the case of B_s decays approximately the same amount of charged kaons of each sign are expected. The presence of a ϕ among the b-hadron decay products is a clear tag for a B_s decay. Finally, these properties are addressed by the variables defined below.

- Largest charged kaon candidate estimator for kaons with electric charge sign opposite to that of the lepton.
- Largest charged kaon candidate estimator for kaons with the same electric charge as the lepton.
- Largest two neutral kaon estimators.
- Invariant mass of the best (according to the kaon estimator) oppositely charged kaon pair formed to look for a ϕ candidate.
- $\sigma = \sum_i w_i\,(1 - w_i)$ the track separation estimator for the hemisphere, used as control variable.

The distributions of all these variables in data and simulation, and the fraction of B_s events as a function of the discriminant variables are shown in Figs. 6.42 to 6.44.

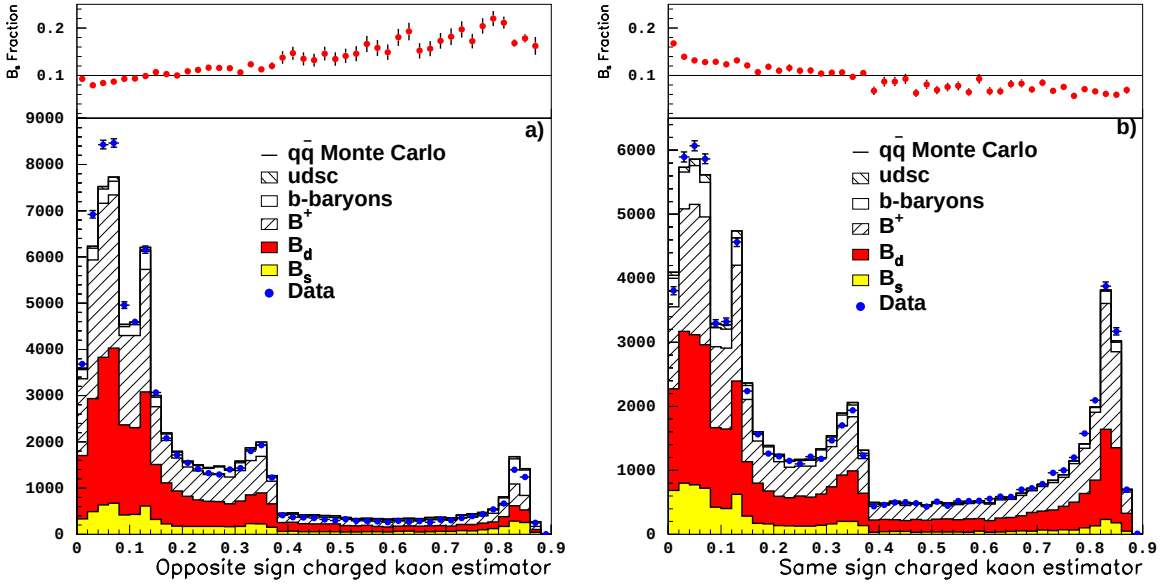

Figure 6.42: Charged decay kaon estimators, a) kaon with charge opposite to that of the lepton, and b) kaon with the same charge as the lepton.

The above twelve variables are combined with a neural network. The training is performed with $B_s \to \ell$ events as signal, and $b \to \ell$ decays, where the b quark forms any other b-hadron, as background. The distribution of the combined variable, N_{enrich} in data and simulation is shown in Fig. 6.45.

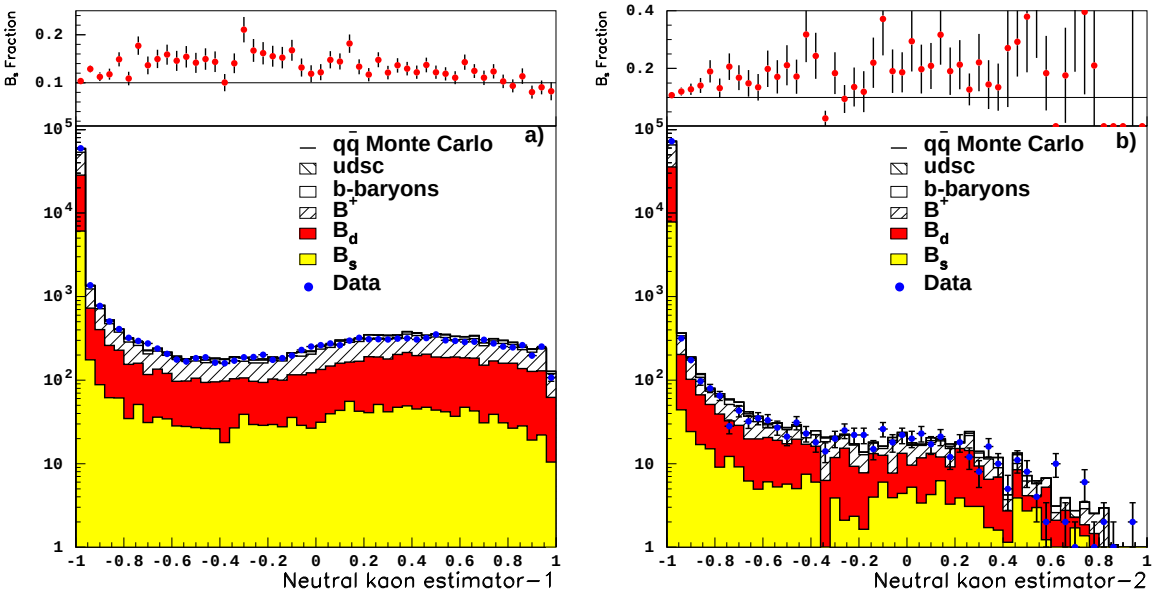

Figure 6.43: Neutral decay kaon estimators for, a) the best candidate, and b) the second best candidate (if any). The variable is set to zero when no neutral candidate is found.

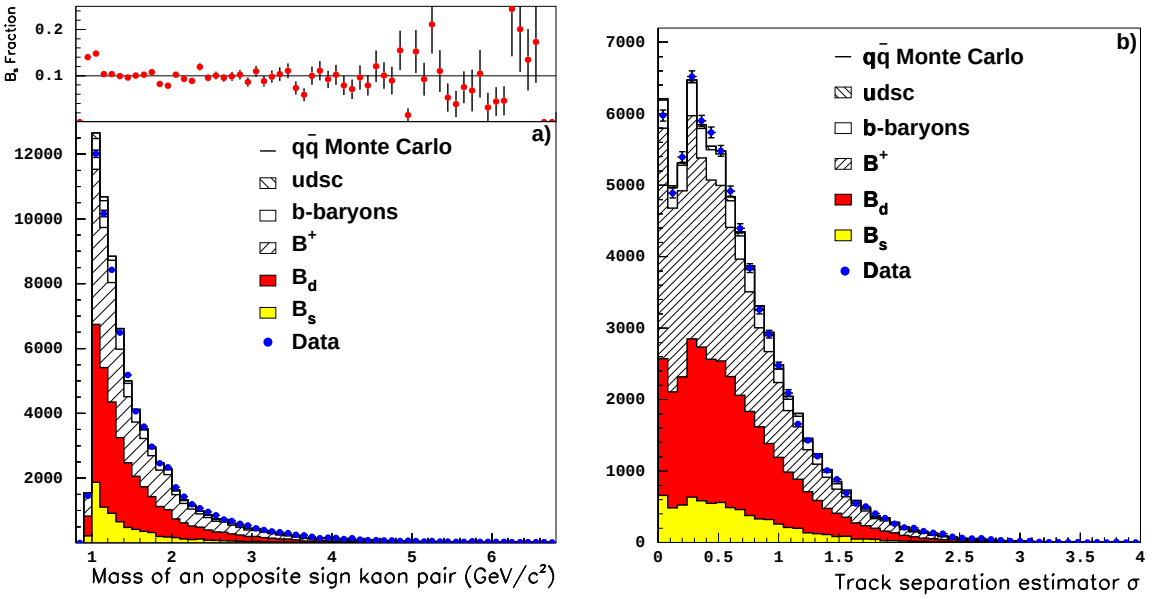

Figure 6.44: a) Invariant mass of the best ϕ candidate, and b) charged particle separation estimator in the B_s hemisphere.

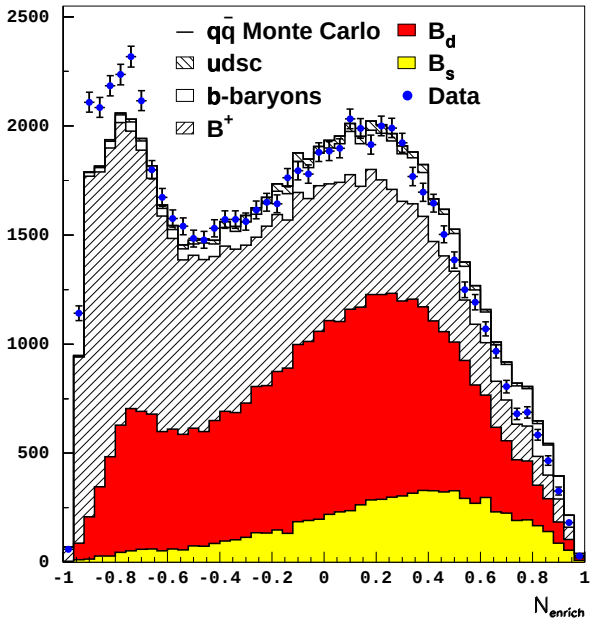

Figure 6.45: Combined enrichment variable.

A disagreement between data and simulation is seen the N_{enrich} distribution. A deficit of simulated events in the region mostly populated by B^+ mesons is observed. Although the biggest discrepancy between data and simulation in the input variables is seen in the charged decay kaon estimators (Fig. 6.42), a problem with the kaon identification is not expected to differentiate between charged and neutral b-hadrons. The explanation must therefore involve the charge estimators. All input charge estimators (Figs. 6.39 and 6.40) used to separate neutral from charged b-hadrons show some disagreement between data and simulation. In all cases, this disagreement favours the hypothesis of having less charged b-hadrons selected in the simulation than in the data.

The secondary vertex charge variable (Fig. 6.39b) was used to fit the amount of charged and neutral b-hadrons seen in the data sample selected. This distribution obtained from the simulation was found to fit nicely the data if the fraction of charged b-hadron was increased by 8%. All the variables which distinguish charged from neutral b-hadrons show a better agreement data/simulation with this adjustment. The distributions of the charge estimators with the correction applied are shown in Figs. 6.46 and 6.47.

The distribution of N_{enrich} after this correction (Fig. 6.48) still shows some deviation, which is however reduced by more than a factor of two with respect to the original one. This residual discrepancy is probably due to differences in shape for the variables that distinguish B_s decays from the other neutral b-hadrons (in particular for the charged kaons estimators). The effect of a difference in shape on the event-by-event estimated B_s purity is smaller than the effect of a bias in the selection efficiency. A possible systematic effect due to the residual discrepancy is estimated by removing the weights which re-adjust the charged-to-neutral ratio (Section 8.2).

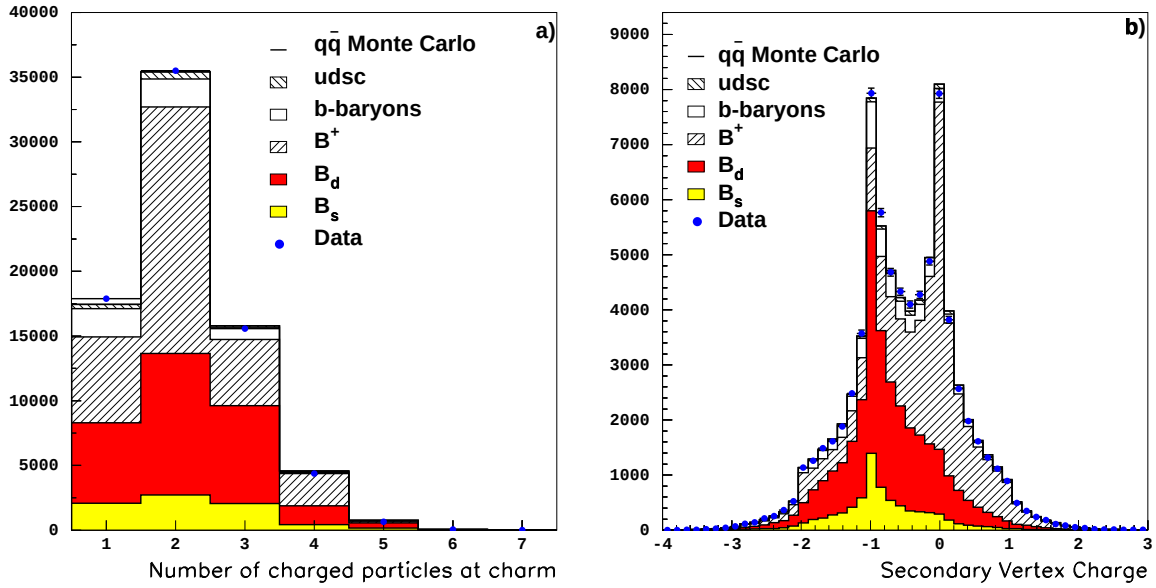

Figure 6.46: a) Number of charged particles at the charm vertex and b) charm vertex charge. Simulated events are reweighted to increase by 8% the B^+ population, as favoured by the data.

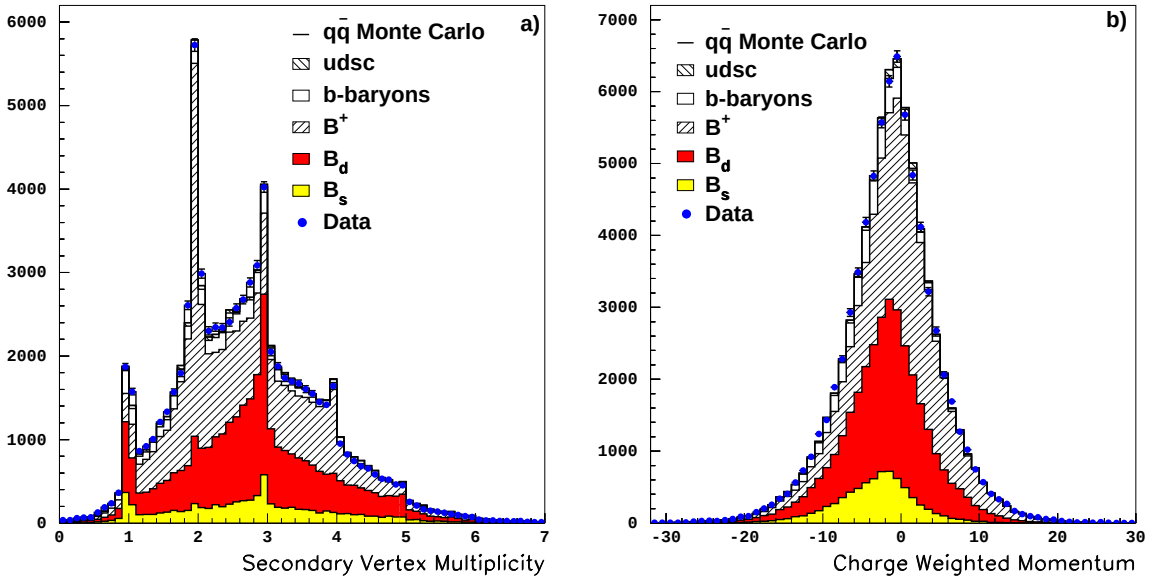

Figure 6.47: a) Charm vertex weighted multiplicity, and b) charm vertex charge weighted momentum. Simulated events are reweighted to increase by 8% the B^+ population, as favoured by the data.

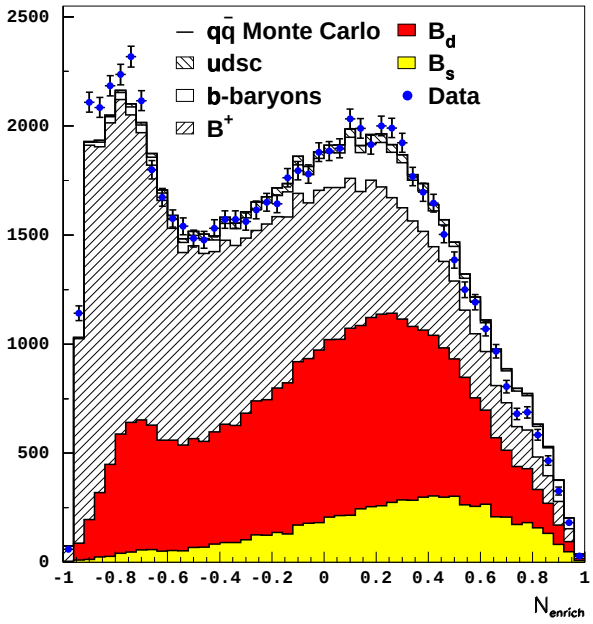

Figure 6.48: Combined enrichment variable with simulated events reweighted to increase by 8% the B^+ population, as favoured by the data.

The B_s enrichment estimator is used event by event to get the probability for the candidate to be a B_s, B_d, B^+, or b-baryon decay. In Fig. 6.49 the probability of each b-hadron species is shown as a function of the enrichment estimator. For approximately half of the event sample ($N_{\text{enrich}} > 0$) the B_s purity is estimated to be higher than the average 10%. For 17% of the events ($N_{\text{enrich}} > 0.4$), it is 20% or higher. The effective increase of the B_s purity on the whole sample is evaluated to be $\sim 25\%$.

6.10 Initial state tag

The B_s oscillation analysis requires the determination of whether the selected b-hadron contains a b or a $\bar{b}$ quark at production time. This determination is achieved with a study of the properties of the whole event, with the exclusion of the lepton candidate and the other b-hadron decay particles.

Properties of the hemisphere opposite to the semileptonic decay are used to determine the $b/\bar{b}$ content of the b-hadron produced together with the B_s candidate, and thus to indirectly determine the production state of the B_s candidate. The same side hemisphere carries information on the fragmentation process which produces the B_s candidate, and can also be used in the initial state flavour determination.

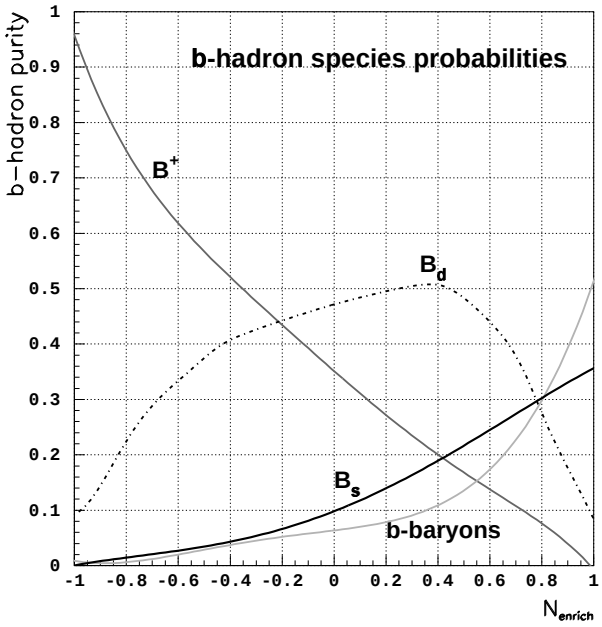

Figure 6.49: Probability of each b-hadron species is shown as a function of N_{enrich}.

6.10.1 Opposite hemisphere

The hemispheres opposite to the $B_s \to \ell$ candidate are unbiased b-hadron decays. Therefore a "general purpose" method for flavour tagging has to be used. The method described here was developed to study B_s oscillations and CP violation in the B_d system [85].

The calculation of the flavour estimator starts with the jet reconstruction. The JADE algorithm is used with a jet-resolution parameter $y_{\text{cut}} = 0.02$. The jet which includes the selected $B_s \to \ell$ candidate is identified and the jet which forms the highest invariant mass with this B_s jet is selected in the opposite hemisphere. The b-hadron decay vertex is searched for with a topological vertexing algorithm (either VNFIT or QVSRCH). This vertex is constrained to be on the axis of the selected jet in the hemisphere. Each track is then assigned a relative probability $\mathcal{P}_v$ that it originates from the secondary vertex.

The JADE jet definition with $y_{\text{cut}} = 0.0044$ is used to estimate the b-hadron flight direction. If more than one jet is found in the hemisphere considered (opposite to the B_s candidate), the b-jet candidate is chosen on the basis of the kinematic properties of its tracks and/or the presence of lepton candidates. The leading track of the b-jet candidate is used as a seed for BTCONE. The resulting jet axis is taken as an estimate of the b-hadron flight direction. In two-jet events (with Section 6.2 definition), the thrust axis is chosen instead. The uncertainties on the reconstructed angles are parametrized as a function of the jet momentum. A second estimate of the direction is taken as the vector joining the primary and secondary vertices, and its uncertainty is parametrized as a function of the measured decay length. The two estimates are averaged using their parametrized uncertainties, and the result is taken as the final b-hadron flight direction.

The b-hadron flight direction described above is used to build a track separation estimator

similar to that described in Section 6.3.2, but optimized on unbiased b-hadron decays. The estimator is converted to a probability $\mathcal{P}_{\rm sec}$ that a track originates from the secondary vertex.

The opposite side tag is built to have similar performance for all b-hadron decays. However, the charged particle distribution in an b-hadron decay hemisphere depends on the b-hadron type (B^+, B^0, or b-baryon). A variety of charge estimators tailored to the different b-hadron decays present in unbiased hemispheres are combined with information intended to differentiate between charged and neutral b-hadron decays.

A set of nine variables sensitive to the b-hadron charge are used for the opposite side tag. Two jet charge estimators are defined using all tracks in the hemisphere.

$$Q_{\rm J} = \frac{\sum_i q_i (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa}{\sum_i (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa} , \qquad \kappa = 0.5,\, 1. , \tag{6.11}$$

where q_i and $\vec{p}_i$ are the charge and momentum of the i-th track, and $\vec{d}_{\rm J}$ is the b-hadron flight direction. The jet charge is expected to have a negative (positive) mean for generic b ($\bar{\rm b}$) -hadrons. Two secondary vertex charges, one weighted with each track momentum along the b-hadron flight direction, are defined on the whole hemisphere

$$Q_{\rm vtx} = \sum_i \mathcal{P}_{\rm sec} q_i , \tag{6.12}$$

$$Q_{\rm vtx'} = \frac{\sum_i \mathcal{P}^i_{\rm sec} q_i (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa}{\sum_i \mathcal{P}^i_{\rm sec} (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa} , \qquad \kappa = 0.3 . \tag{6.13}$$

The weighted secondary vertex charge is expected to have the same correlation with the b-hadron flavour as the jet charge. The pure secondary vertex charge distinguishes between neutral and charged b-hadrons. For charged b-hadrons, it is also sensitive to the b ($\bar{\rm b}$) content. The weighted primary vertex charge, defined as

$$Q_{\rm pri} = \frac{\sum_i (1 - \mathcal{P}^i_{\rm sec}) q_i (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa}{\sum_i (1 - \mathcal{P}^i_{\rm sec}) (\vec{p}_i \cdot \vec{d}_{\rm J})^\kappa} , \qquad \kappa = 1.0 , \tag{6.14}$$

which also includes all tracks in the hemisphere. For charged b-hadrons, the primary vertex charge has the opposite charge correlation with the b quark, than that of the secondary vertex charge. For neutral b-hadrons, the primary vertex charge is especially important, as the secondary vertex charge provides almost no information.

Apart from the hemisphere-based charge estimators defined above, the total charge of the jet is also used. It is defined as $Q_{\rm tot} = \sum_i q_i$, where i runs over the charged particles in the b-jet. It is expected to have the same charge correlation with the b-hadron charge as the jet charge estimators.

The b ($\bar{\rm b}$) content of the b-hadron can also be estimated with $p_\parallel \cdot q$, where $p_\parallel$ is the longitudinal momentum with respect to the b-hadron flight direction of the leading particle in the b-hadron hemisphere, and q the sign of its electric charge. Identified leptons (Section 5.3) and kaons (Section 6.8.1) are also used, if present.

The variables described above are combined, together with a set of control variables ($\cos\theta_{\rm thurst}$, the charged particle multiplicity, the combined track separator estimator for the hemisphere, the reconstructed b-hadron momentum, the reconstructed decay length, and the

lepton momentum, in the case a lepton was found) with a neural network to obtain the combined opposite side tag variable. The distribution of this variable for the selected data and simulation samples are shown in Fig. 6.50.

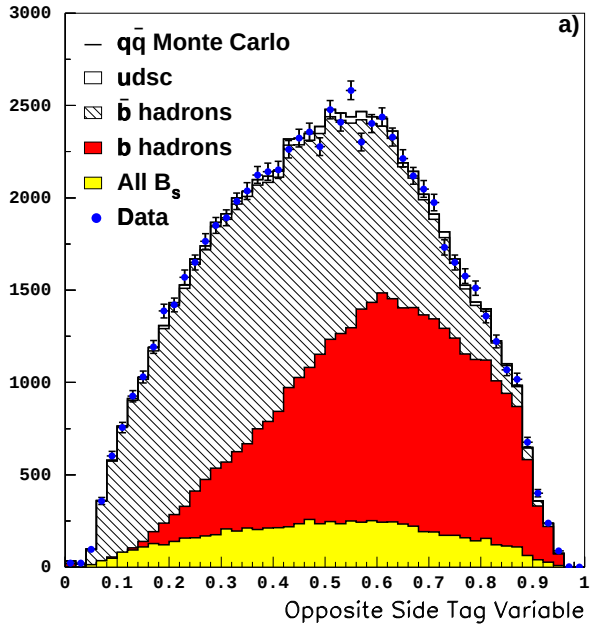

Figure 6.50: Opposite side initial state tag variable.

The effective mistag probability (defined in the next Section) for the opposite tag estimator in generic hemispheres is $\eta_{\text{eff}} = 29\%$.

6.10.2 Same hemisphere

On the B_s candidate hemisphere, flavour estimators are built as insensitive as possible to the decay products. The estimators can be divided in two types, *i*) primary vertex charge estimators, and *ii*) a charged fragmentation kaon estimator. In the first case, the estimators are built to be sensitive to the overall charge of fragmentation particles in the hemisphere of the B_s candidate. If the fragmentation kaon produced together with the B_s candidate is electrically charged, the charge sign is correlated with the b flavour of the B_s at production.

Five different estimators of the primary vertex charge are defined. Three of them use all tracks in the "big" b-jet weighted with their probability to be b-hadron decay products (defined as explained in Sections 6.3.2 and 6.9). The other two use tracks inside the same b-jet after the exclusion of the selected c tracks. These estimators are defined as

$$\begin{aligned} \text{PV Charge}_{\text{all}} &= \textstyle\sum_i^{N_{\text{jet}}} (1-w_i)\, q_i\, p_{i_\parallel}^{\kappa} \qquad \kappa = 0,\ 0.6,\ 1\ , \\ \text{PV Charge}_{\text{c excl}} &= \textstyle\sum_i^{N_{\text{jet}}} q_i\, p_{i_\parallel}^{\kappa} \qquad \kappa = 0,\ 0.3\ , \end{aligned} \tag{6.15}$$

where N_{jet} is the number of charged particles in the "big" b-jet, q_i is the charge of the i-th particle, $p_{i_\parallel}$ is the i-th particle momentum projected on the reconstructed b-hadron

flight direction. The κ values have been optimized to obtain the best overall separation performance.

The distributions of these five estimators are shown in Figs. 6.51 to 6.55. On the left (a) plots, the variable distribution is shown for all selected events, the data distribution is compared to that of simulated events; on the right (b), the variable distribution for $B_s \to \ell$ and $\overline{B}_s \to \ell$ simulated events are shown.

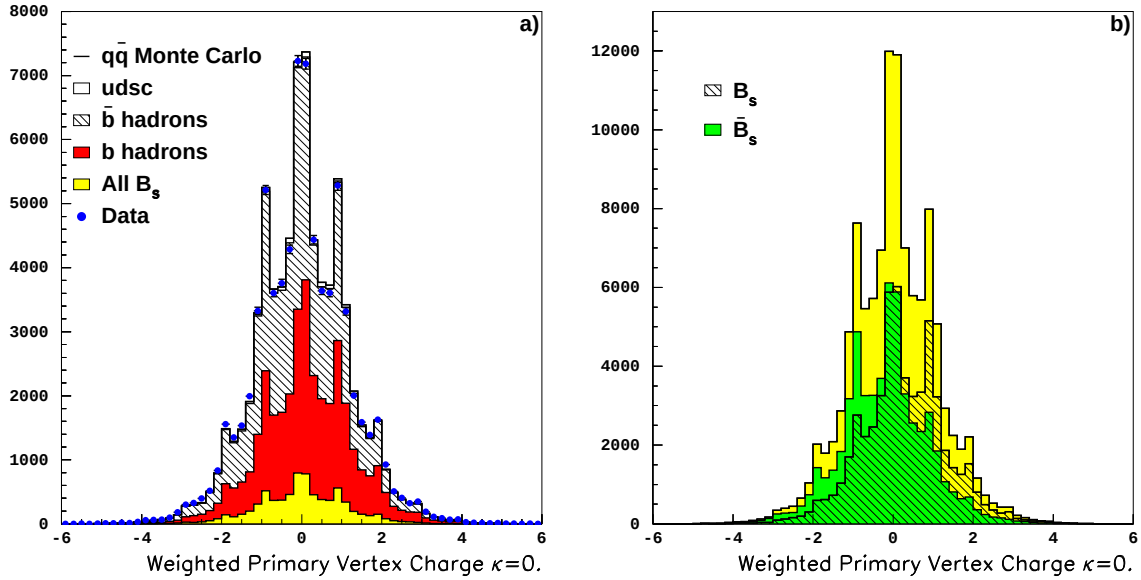

Figure 6.51: Weighted primary vertex charge with $\kappa = 0$. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

Charged fragmentation kaons are looked for in the selected events within particles in the "big" jet but not included in the c track selection. If their fragmentation kaon estimator (Section 6.8.1) value is positive, this value multiplied by the electric charge of the kaon candidate track is taken as fragmentation kaon variable for the flavour tagging. The variable is set to zero otherwise. The distribution of the charged fragmentation estimator used here is shown in Fig. 6.56a for the selected data sample, and in Fig. 6.56b for $B_s \to \ell$ and $\overline{B}_s \to \ell$ simulated events.

In addition, some control variables are used, *i*) the measured decay length l; *ii*) the measured momentum p_B; and *iii*) $\sigma = \sum_i w_i(1-w_i)$ which estimates the quality of the track separation.

An additional variable, not specific to any of the two hemispheres, is considered. The polar angle of the event thrust axis is used as a global flavour estimator profiting from the b forward-backward asymmetry. The distribution of the signed thrust polar angle is shown in Fig. 6.57.

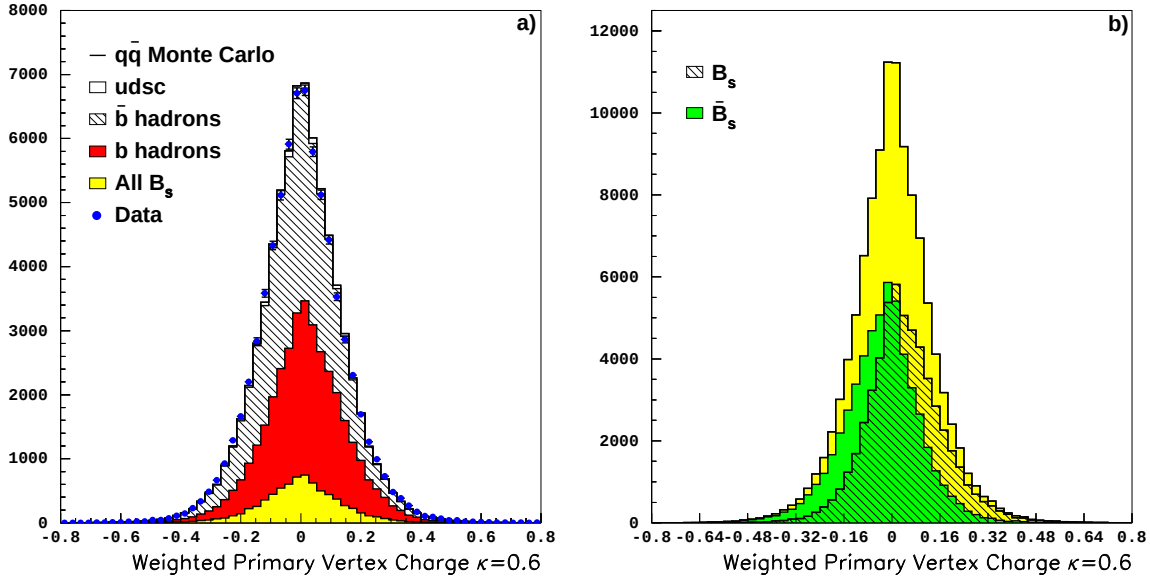

Figure 6.52: Weighted primary vertex charge with $\kappa = 0.6$. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

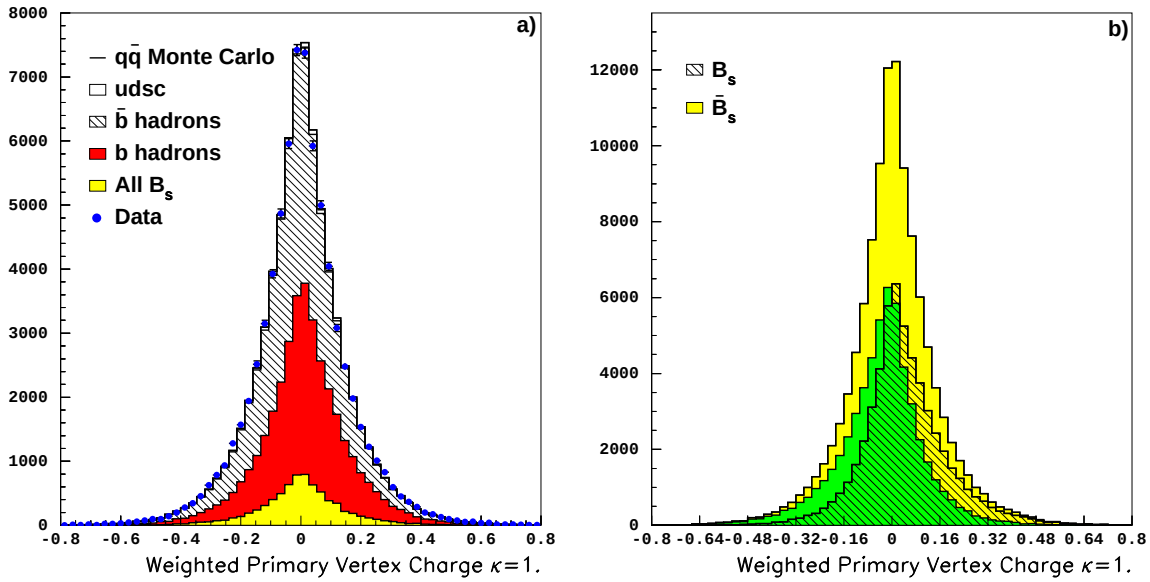

Figure 6.53: Weighted primary vertex charge with $\kappa = 1$. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

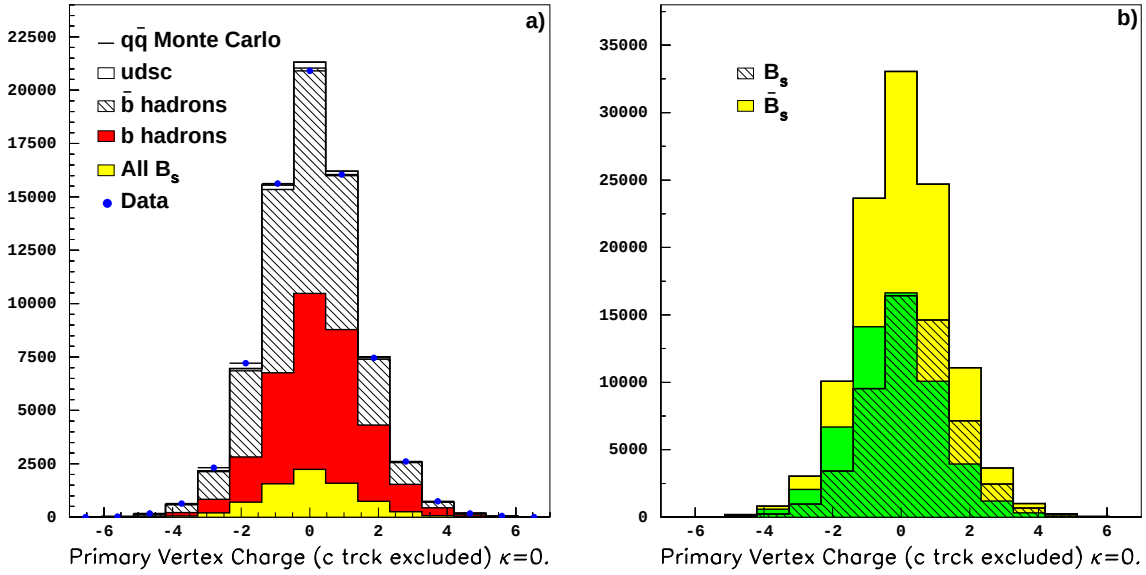

Figure 6.54: Primary vertex charge, c tracks excluded, with $\kappa = 0$. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

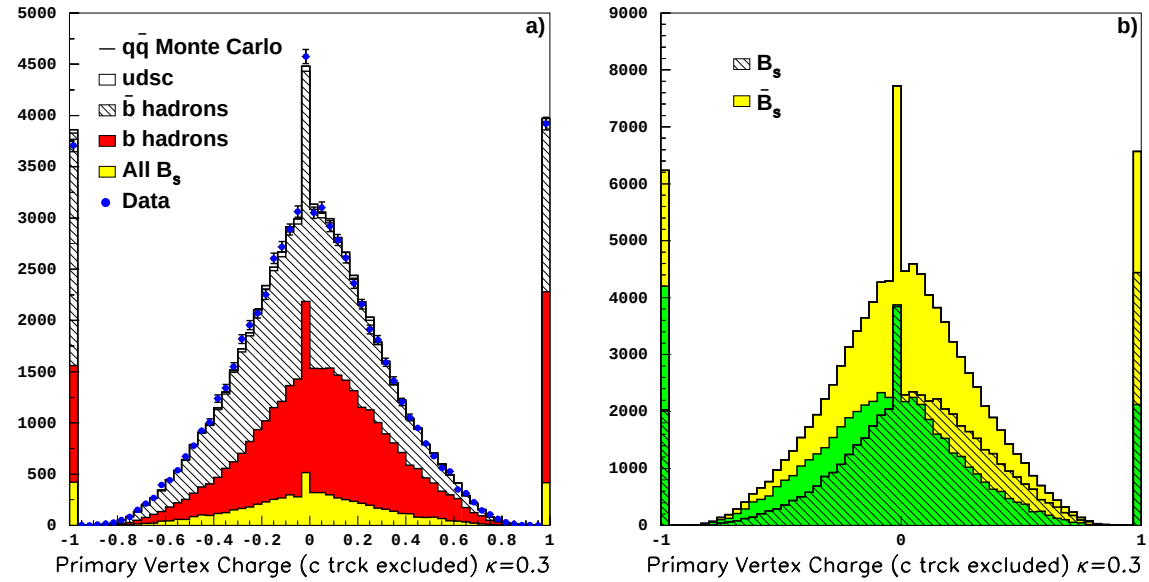

Figure 6.55: Primary vertex charge, c tracks excluded, with $\kappa = 0.3$. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

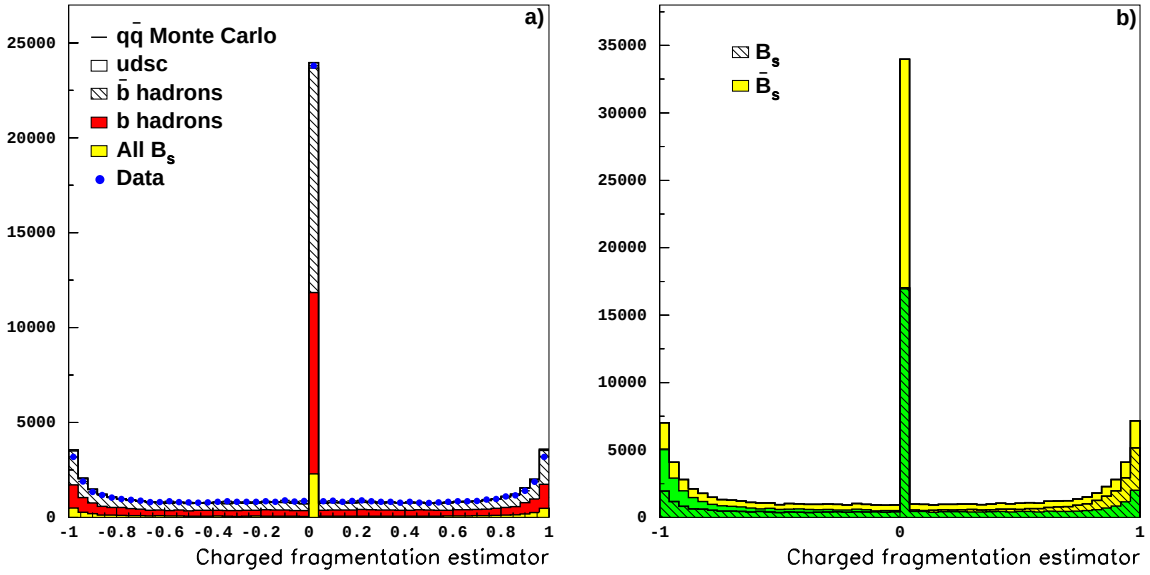

Figure 6.56: Charged fragmentation kaon estimator. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

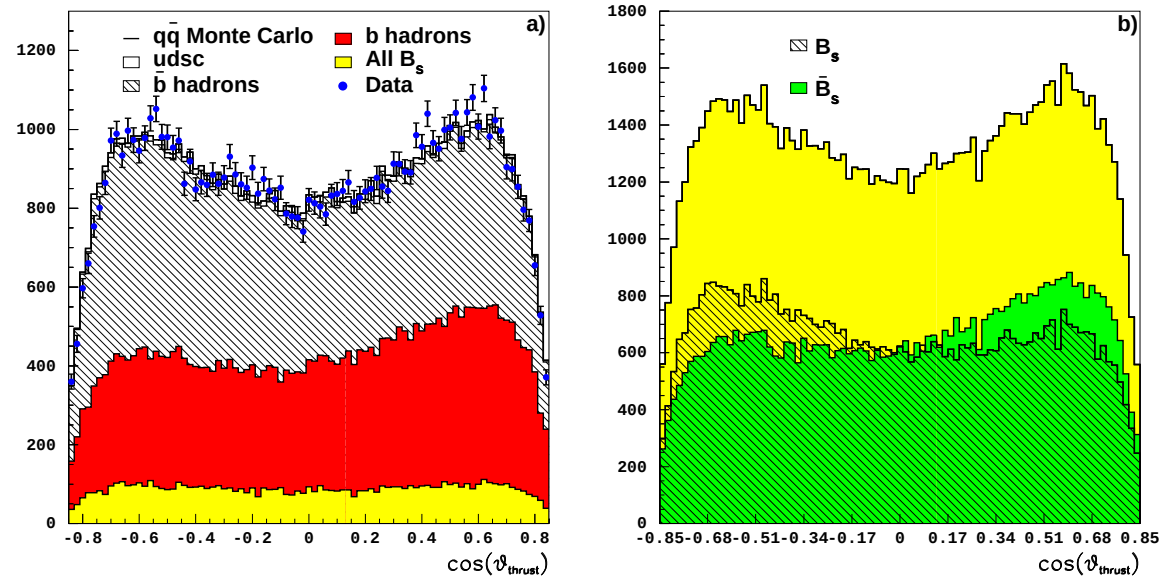

Figure 6.57: Signed thrust polar angle. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

6.10.3 Flavour tag performance

The variables described in the previous Section are combined together with the opposite side flavour tag variable (Section 6.10.1) to obtain the combined initial state tag estimator. The combination is done with a neural network trained to distinguish $B_s \to \ell$ from $\overline{B}_s \to \ell$. The same side discriminant variables are sensitive to the true b-hadron species of the B_s decay candidate. The charged fragmentation kaon is only expected to be found in the case of a $B_s/\overline{B}_s$ decay. In the case of $B^{+/-}$ decays, the primary vertex charge is expected to have a correlation with the production flavour opposite to that found for $B_d/\overline{B}_d$ or $B_s/\overline{B}_s$.

The distribution of the combined initial state tag variable, N_{is}, is shown in Fig. 6.58a for the selected sample, and in Fig. 6.58b for $B_s \to \ell$ and $\overline{B}_s \to \ell$ simulated events.

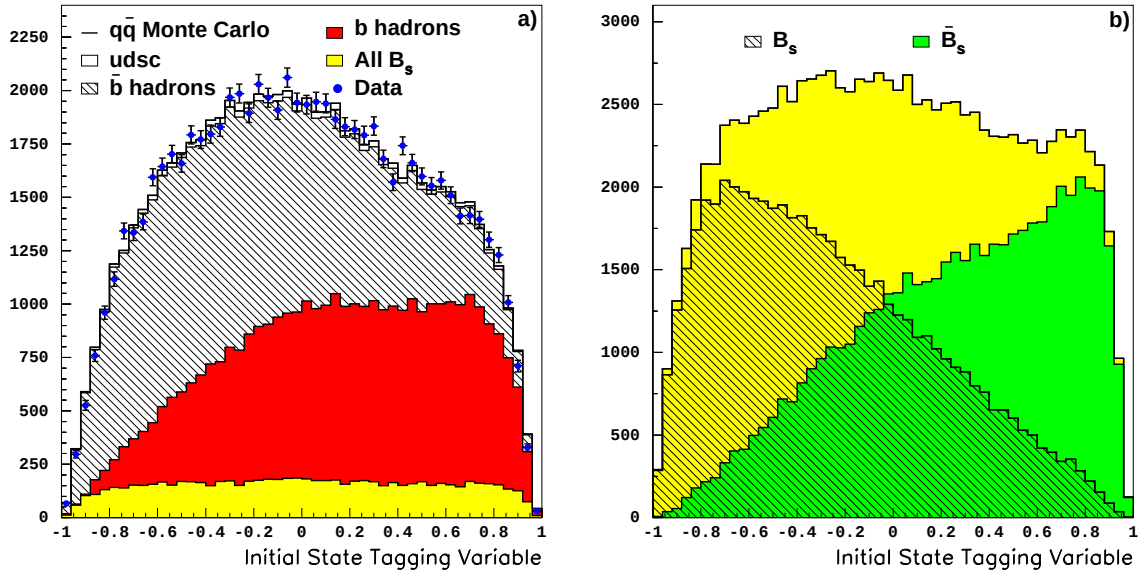

Figure 6.58: Initial state tag combined variable. a) For all selected events in data and simulation, and b) for $B_s \to \ell$ simulated events.

Events with $N_{is} < 0$ are considered as originating from a $\bar{b}$-quark. This initial state flavour is compared to the electric charge of the lepton, if $N_{is} < 0$ and the lepton charge is positive (negative), or if $N_{is} > 0$ and the lepton charge is negative (positive), the event is labeled as "unmixed" ("mixed"). This tag is used in Chapter 8 to define the probability density function to be used for each event.

The initial state tag variable, N_{is}, is translated to an event-by-event mistag probability to be used in the oscillation fit (Section 8.1.1). For $B_s \to \ell$ signal events, N_{is} is calibrated as shown in Fig. 6.59. Therefore, no particular parametrization is needed.

The effective mistag probability is defined as

$$\eta_{\text{eff}} = \frac{1}{2}\left[1 - \sqrt{\frac{\sum_i (1-\eta_i)^2 N_i^T}{\sum_i N_i^T}}\right] . \tag{6.16}$$

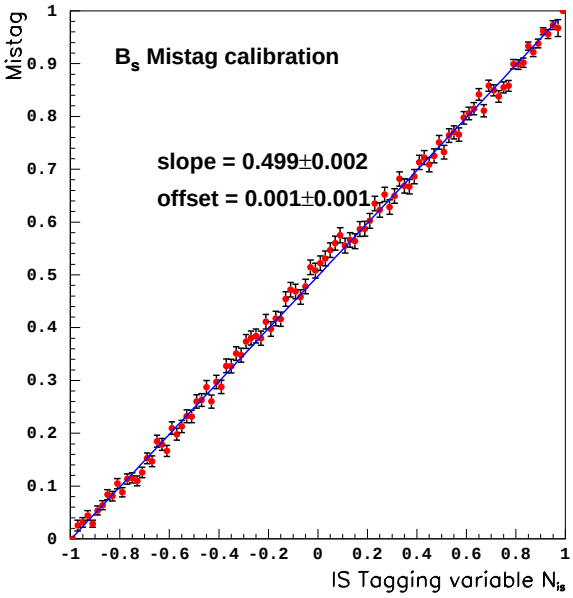

Figure 6.59: Calibration of the Initial State tag variable for $B_s \to \ell$ simulated events. The slope of the straight line fit and the offset at $N_{is} = -1$ are given.

The distribution of the discriminant variable N_{is} is divided in bins, η_i is the fraction of wrongly tagged events in each bin, and N_i^T is the total number of events in each bin. The effective mistag is used to measure the effective statistical significance which results from the tagging variable. On $B_s \to \ell$ events it is evaluated to be $\eta_{eff} = 24\%$ (while the average mistag probability is $\eta_{avg} = 28\%$).

The mistag probability needs to be parametrized as a function of the tagging variable for the other b-hadron species. The global (all vertex classes together) calibration curves for B_d, B^+, and b-baryons are shown in Fig. 6.60.

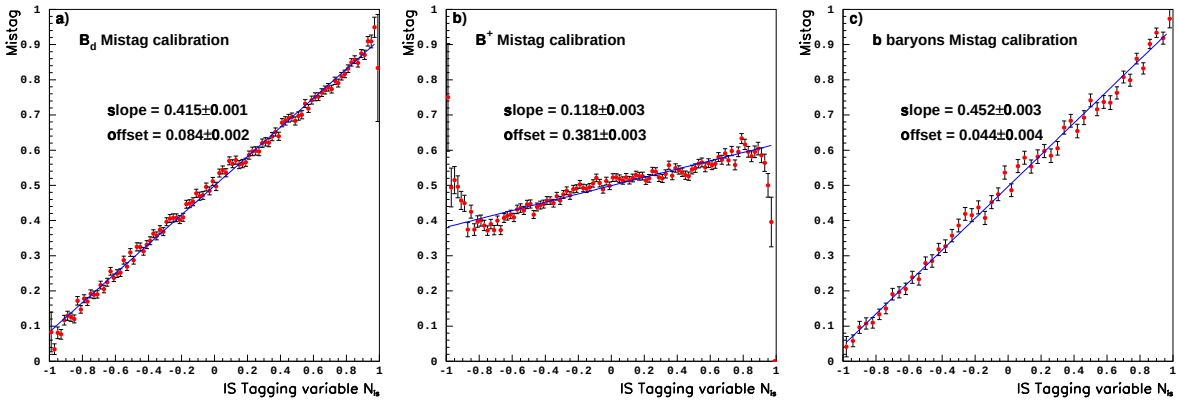

Figure 6.60: Calibration of the N_{is} variable for: a) B_d, b) B^+, and c) b-baryons. The slope of the straight line fit and the offset at $N_{is} = -1$ are given.

None of the curves in Fig. 6.60 is a perfect straight line with slope 0.5 and zero offset at $N_{\rm is} = -1$, as in the case of $B_s \to \ell$. The worst performance is observed for B^+. It is due to the opposite correlation between the primary vertex charge and the b-hadron flavour of B^+ with respect to B_s. The fact that the initial state tagging performance for the other b-hadrons is not optimal is not a problem for the B_s oscillation analysis. However the parametrization of the performance has to be done carefully not to affect the analysis result, a detailed parametrization is performed for each b-hadron species. In particular, it has been seen that, in the case of B_d mesons, the performance of the tagging depends on the measured proper time of the meson. The slope of the calibration curve is below 0.4 at low proper time ($t < 0.6\,\text{ps}$) and above 0.45 for high proper time ($t > 3\,\text{ps}$). This effect is taken into account in the event characterization used for the oscillation fit.

Chapter 7

Supplementary B_s selections

Data samples enriched in B_s meson decays can be used, in supplement to the inclusive semileptonic selection described in Chapter 6, for a B_s oscillation study. Two selections of this type, available in ALEPH, are described in this Chapter. In Section 7.1 a selection of $B_s \to D_s^- \ell^+ \nu$, where the D_s is completely reconstructed, is described. Some specific hadronic B_s decays to flavour eigenstates such as $B_s \to D_s^- \pi^+$ can be fully reconstructed. A selection of fully reconstructed B_s candidates is presented in Section 7.2. The method used to estimate the b-flavour at production for both selections is described in Section 7.3.2.

7.1 Selection of $D_s \ell$ pairs

The selection of $D_s \ell$ pairs aims at the reconstruction of $B_s \to D_s^- \ell^+ \nu(X)$ decays (charge conjugate states are implied along this Chapter, unless otherwise stated).

The D_s^- is reconstructed in six hadronic decay modes, $\phi\pi^-$, $K^{*0}K^-$, $K_S K^-$, $\phi\rho^-$, $K^{*0}K^{*-}$, and $\phi\pi^+\pi^-\pi^-$, and two semileptonic modes, $\phi e^- \nu_e$ and $\phi \mu^- \nu_\mu$. The sources of background in the selected sample are *i*) $b \to D_s^- D_{(s)} X$ events, where $D_{(s)} \to \ell^+ \nu(X)$; *ii*) kinematic reflections in the $K^{*0}K^{*-}$ ($K_S K^-$) channel, from $D^- \to K^{*0}\pi^-$ ($K_S \pi^-$) decays; and *iii*) combinatorial background.

A high purity charged kaon and pion identification is needed for this selection. In both cases a cut in the particle momentum and in the expected dE/dx deposition in the TPC are placed. For kaons these cuts are $p > 1.5\,\mathrm{GeV}/c$ (to ensure a good kaon/pion dE/dx separation), and $\chi_\pi + \chi_K < 1$, where χ_h is the difference between the measured energy loss and the expectation for the hadron h, divided by the expected uncertainty. For pions candidates these cuts are $p > 0.5\,\mathrm{GeV}/c$ for the particle momentum, and $|\chi_\pi| < 3$ for the measured ionization. The lepton candidate is identified as described in Section 5.3, both electrons and muons are required to have a momentum larger than $2.5\,\mathrm{GeV}/c$.

The unstable decay products of the D_s^- are reconstructed in the following channels:

$$\phi \to K^+K^- \qquad K^{*0} \to K^+\pi^- \qquad K_S \to \pi^+\pi^- \qquad K^{*-} \to K_S\pi^- \qquad \text{and} \qquad \rho^- \to \pi^0\pi^- \ .$$

Their reconstructed mass is required to lie within a specific range around the measured value [13] (between two and three standard deviations). The decay of the pseudoscalar

meson D_s^- into a vector meson (ϕ or K^{*0}) and a pseudoscalar meson (π^- or K^-) follows a distribution proportional to $\cos^2(\lambda)$, where λ is the helicity angle defined as the angle between the π^- (K^-) from the D_s^- and one of the ϕ (K^{*0}) daughters in the ϕ (K^{*0}) rest frame. To reduce the combinatorial background, $|\cos(\lambda)| > 0.4$ is required.

The reconstructed D_s^- candidate is extrapolated back and vertexed together with the lepton to obtain the B_s decay vertex. The vertex probability associated to the D_s^- and to the B_s candidates must be at least 1%.

7.1.1 Sample composition

The reconstructed D_s^- mass peak is fit with two Gaussians for the signal and a second order polynomial for the background. The background polynomial is fit on the wrong charge $D_s\,\ell$ combinations, for those channels in which statistics are sufficient, and on simulated events otherwise. For the leptonic modes a similar fit is performed to the ϕ mass peak. For the hadronic modes $K^{*0}K^-$, K^-K_S, and $K^{*0}K^{*-}$ a third Gaussian is included for the D^- peak. In the channel $\phi\,\pi^+\pi^-\pi^-$ an additional Gaussian is included to take into account reflections of D^{*-}, in which a pion in the decay $D^0 \to K^-\pi^+\pi^+\pi^-$ is misidentified as a kaon. The mass spectra on the ALEPH data is shown in Fig 7.1, together with the fit result.

Only those events for which the reconstructed D_s (or ϕ) mass is found in a window of two (or three, depending on the channel) standard deviations around the measured D_s mass [13], are selected. The sample composition achieved with the selection described above is summarized in Table 7.1.

Channel	Efficiency (%)	Signal (%)	$b \to D_s D X$ (%)	D^- (%)	Combinatorial (%)	Events observed
$\phi\,\pi^-$	12.1	52.0	21.8	0	26.2	68
$K^{*0}\,K^-$	8.2	34.0	12.8	13.3	39.9	113
$K_S\,K^-$	2.5	40.8	15.5	2.9	40.8	24
$\phi\,\rho^-$	1.1	28.2	6.4	2.7	62.7	22
$K^{*0}\,K^{*-}$	3.6	38.9	15.8	2.1	43.2	19
$\phi\,\pi^+\pi^-\pi^-$	6.3	27.7	10.0	0	62.3	13
$\phi\,e^-\bar{\nu}_e$	6.0	40.0	12.1	0	47.9	19
$\phi\,\mu^-\bar{\nu}_\mu$	4.0	51.1	20.0	0	28.9	19
Total		39.8	15.1	5.6	39.5	297

Table 7.1: Signal efficiency and estimated composition of the $D_s\,\ell$ data sample, evaluated from the simulation. In the last column, the number of candidates found in the data for each channel is shown.

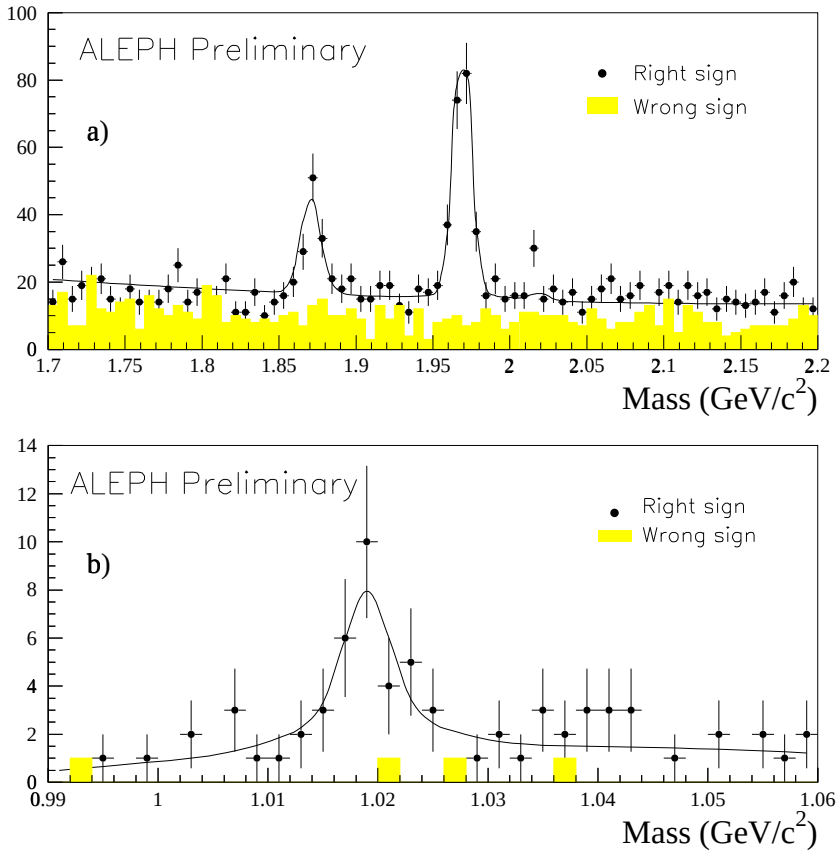

Figure 7.1: Reconstructed mass spectra for the D_s selected sample. a) D_s reconstructed mass for the hadronic D_s decays, and b) ϕ candidate mass for the semileptonic D_s decays. In both cases, a fit to the signal and background distributions is performed.

7.1.2 Event characterization

Signal purity

The average purity of the sample is evaluated from the simulation to be $\sim 40\%$. The statistical power of the analysis is enhanced via an event-by-event evaluation of the signal probability for each channel. The relevant variables used for this purpose are the lepton momentum and transverse momentum, and the number of charged particles (excluding those from the D_s candidate) which form a good vertex with the lepton.

Proper time measurement

The B_s decay length is measured by projecting the distance between the primary vertex and the B_s vertex along the direction of the total momentum of the $D_s \ell$ pair. The decay length uncertainty is obtained for each event from the vertex fit. The average resolution is $250\,\mu$m. The decay length uncertainty is found to be underestimated by 10% in the simulation. A

degradation of this amount is applied to the event-by-event measured uncertainty to account for it. The decay length resolution in the data is found to be slightly worse than that suggested by the simulation [61]. Therefore, an additional $5 \pm 3\%$ degradation is applied to correct for this disagreement.

The B_s momentum is computed from the reconstructed D_s momentum, the lepton momentum, and an estimate of the neutrino energy obtained as explained in Section 6.7.1. A bias on the reconstructed momentum is observed in the simulation, it accounts, on average, for $1\,\mathrm{GeV}/c$. A correction as a function of the reconstructed momentum is applied to all events. The relative momentum uncertainty is parametrized as a function of the estimated momentum. The average relative momentum resolution obtained is 10%.

About half of the data sample selected here is also selected by the analysis presented in Chapter 6. For the presentation of the averaged ALEPH oscillation results, the events in common between the two semileptonic selections are excluded from the inclusive semileptonic selection to avoid any statistical correlation between the two data samples.

7.2 Selection of exclusive hadronic flavour final states

The hadronic B_s selection aims at the reconstruction of the following decay channels:

$$\mathrm{B_s} \to \mathrm{D_s^{-(*)}}\pi^+ \qquad \mathrm{B_s} \to \mathrm{D_s^{-(*)}}a_1^+ \qquad \mathrm{B_s} \to \mathrm{D_s^{-(*)}}\rho^+,$$

with the D_s candidates reconstructed as

$$\mathrm{D_s^-} \to \phi\,\pi^- \qquad \mathrm{D_s^-} \to \mathrm{K^{*0}K^-} \qquad \mathrm{D_s^-} \to \mathrm{K_S K^-}\ .$$

The ϕ, $\mathrm{K^{*0}}$, a_1^+ and $\mathrm{K_S}$ are reconstructed in their charged decay modes: $\phi \to \mathrm{K^+K^-}$, $\mathrm{K^{*0}} \to \mathrm{K^+}\pi^-$, $a_1^+ \to \rho^0\pi^+$, and $\mathrm{K_S} \to \pi^+\pi^-$.

To reduce the combinatorial background, cuts are applied to the $\mathrm{d}E/\mathrm{d}x$ estimator for kaons and pions, $\chi_\pi + \chi_\mathrm{K} < 1.6$ for kaons, and $|\chi_\pi| < 3$ for pions.

All track candidates for the $\mathrm{D_s^{-(*)}}\pi^+$ channel reconstruction are required to have a momentum in excess of $1\,\mathrm{GeV}/c$. Then the reconstruction of the D_s candidates is performed in a similar manner as that in Section 7.1. The reconstructed D_s candidate is required to have a momentum greater than $3\,\mathrm{GeV}/c$.

The momentum of the pion candidates used in the a_1^+ reconstruction is required to be greater than $0.5\,\mathrm{GeV}/c$, and the momentum of the reconstructed ρ^0 and a_1^+ in excess of $1\,\mathrm{GeV}/c$. The three pions, two of which have an invariant mass within $\pm 150\,\mathrm{GeV}/c^2$ of the nominal ρ^0 mass [13], are required to form a common vertex with a χ^2 probability of the fit greater than 1%. The difference between the invariant mass and the nominal mass of the a_1^+ [13] should be within the natural width of the a_1^+ ($\pm 300\,\mathrm{MeV}/c^2$).

The B_s candidates are reconstructed with the combination of $\mathrm{D_s^-}$ and a π^+ or an a_1^+ candidates. The B_s vertex fit is required to have a probability in excess of 1%. The D_s vertex is imposed to be in front of the B_s vertex. To ensure that the B_s flight path is along the B_s momentum direction, $\cos(\alpha) \geq 0.95$ is required, where α is the angle between the B_s decay length direction and the B_s momentum; in the case of the $\mathrm{D_s^-}a_1^+$ channel, the angular cut is

tighter: $\cos(\alpha) \geq 0.97$. The B_s momentum is required to be larger than $20\,\mathrm{GeV}/c$ for the $D_s^-\pi^+$ channel; the cut is increased to $25\,\mathrm{GeV}/c$ for the $D_s^- a_1^+$ ($D_s^- \to \phi\pi$) channel, and to $30\,\mathrm{GeV}/c$ for the other $D_s^- a_1^+$ channels. If more than one combination is present in an event, that with the highest momentum is taken.

7.2.1 Sample composition

Events in which the B_s decays to a D_s^{*-} ($\to D_s^-\gamma$), or to a $D_s^-\rho^+$ pair have a low reconstructed mass because of the loss due to the non-reconstructed photon or neutral pion from the decay of the D_s^{*-} or the ρ^+. The simulation predicts these events to populate the mass region between $5\,\mathrm{GeV}/c^2$ and $5.3\,\mathrm{GeV}/c^2$. This region is referred in the following as the satellite peak. About 80% of the $D_s^{*-} \to D_s^-\gamma$ events are recovered through the identification of a photon close in phase space from the D_s^-, and with a $D_s^- + \gamma$ mass close to the nominal D_s^{*-} mass [13]. To improve the resolution on the reconstructed B_s mass, the photon energy is recomputed imposing the D_s^{*-} mass constraint. Only 40% of the $B_s \to D_s^-\rho^+$, where a π^0 is missing in the final state, are recovered with a similar technique.

Only those events for which the reconstructed B_s mass is in a range between $5.0\,\mathrm{GeV}/c^2$ and $5.44\,\mathrm{GeV}/c^2$ are taken. If the reconstructed mass falls within $\pm 70\,\mathrm{MeV}/c^2$ of the measured B_s mass ($5369.6 \pm 2.4\,\mathrm{MeV}/c^2$ [13]), the event is considered to be in the "main peak", and in the satellite peak otherwise. The selection efficiencies in the different channels are presented in Table 7.2.

Channel	Efficiency (%)
$B_s \to D_s^-\pi^+(\phi\,\pi^-)$	21 ± 1
$B_s \to D_s^-\pi^+(K^{*0}K^-)$	15 ± 1
$B_s \to D_s^-\pi^+(K_S K^-)$	12 ± 1
$B_s \to D_s^- a_1^+(\phi\,\pi^-)$	11 ± 1
$B_s \to D_s^- a_1^+(K^{*0}K^-)$	6 ± 1
$B_s \to D_s^- a_1^+(K_S K^-)$	8.0 ± 0.4

Table 7.2: Event selection efficiencies for each channel.

The reconstructed B_s mass spectra for $D_s^-\pi^+$, $D_s^- a_1^+$, and the sum of them are shown in Fig. 7.2. The first plot includes $D_s^{-(*)}\pi^+$ and $D_s^{-(*)}\rho^+$ decay channels, and the second includes $D_s^{-(*)} a_1^+$.

The purity of the signal and the composition of the background are evaluated from the simulated samples and are given in Table 7.3. The background contribution which is not explicited in the Table is due to combinatorics. The data sample in the $D_s^- \to \phi\,\pi^-$ channel is divided into two classes depending on the decay length significance of the reconstructed B_s candidate. In Table 7.3, these classes are indicated as (1) for $l_{B_s}/\sigma_l > 3$ and (2) for $0 < l_{B_s}/\sigma_l < 3$. For the other channels, only events events with high decay length significance ($l_{B_s}/\sigma_l > 3$) are selected.

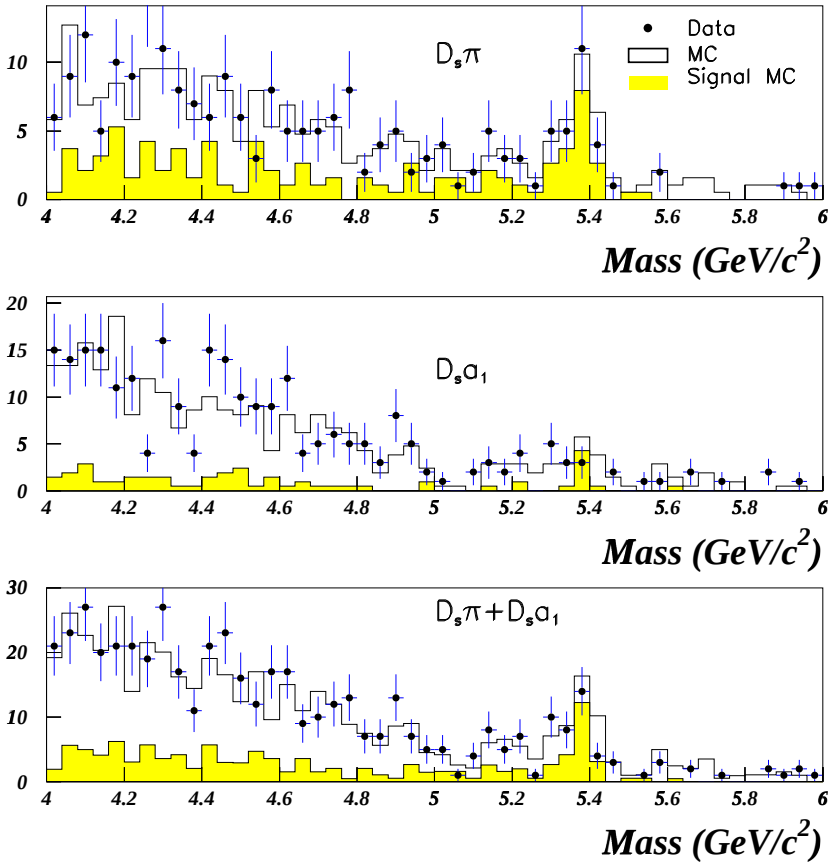

Figure 7.2: Mass spectra of the $D_s^{-(*)}\pi^+ + D_s^{-(*)}\rho^+$, and $D_s^{-(*)}a_1^+$ candidates. The points and error bars are the data, the histograms correspond to the simulation. The shaded histograms show the contribution of the signal events. The first plot refers to the $D_s^{-(*)}\pi^+ + D_s^{-(*)}\rho^+$ channels, the second to the $D_s^{-(*)}a_1^+$ channel, and the last one is the sum of the other two.

7.2.2 Event characterization

Signal purity

The average purity for each channel in the main and satellite peaks is shown in Table 7.3. An effective purity is evaluated for each event in every signal region according to some discriminant variables like: the helicity angle (for D_s decays to a pseudoscalar and a vector) and the reconstructed D_s mass. The statistical power of the sample for the B_s oscillation analysis is increased by the use of this effective purity.

Proper time measurement

The decay length of the B_s candidates is estimated as the distance between the B_s vertex and the primary vertex projected on the direction of the reconstructed B_s momentum. The decay length resolution is evaluated to be about 150 μm. As in the case of D_s ℓ pairs, a degradation

Region	B_s purity (%)	B_d fraction of background (%)	B^+ fraction of background (%)	Events Observed
Main peak $D_s^-\pi^+$				
$\phi\pi^+$ (1)	81	17	10	7
$K^{*0}K^-$	50	22	11	11
K_SK^-	28	7	0	6
$\phi\pi^+$ (2)	2	5	3	6
Satellite peak $D_s^-\pi^+$				
$\phi\pi^+$ (1)	82	33	67	4
$K^{*0}K^-$	32	41	24	7
K_SK^-	30	32	5	9
$\phi\pi^+$ (2)	15	12	0	8
Main peak $D_s^-a_1^+$				
$\phi\pi^+$ (1)	47	26	13	6
$K^{*0}K^-$	44	24	24	3
K_SK^-	37	53	18	2
$\phi\pi^+$ (2)	3	3	3	6
Satellite peak $D_s^-a_1^+$				
$\phi\pi^+$ (1)	24	32	28	6
$K^{*0}K^-$	15	52	17	9
K_SK^-	8	33	25	10

Table 7.3: Fraction of signal and background events in the two mass peak regions for the two decay channels as predicted from the simulation.

of $5\pm3\%$ is applied to correct for possible disagreements between the data and the simulation. The momentum reconstruction is obtained from the sum of the vector-momenta of all the B_s decay products, it leads to an excellent relative momentum resolution, evaluated to be of 0.5%.

7.3 Flavour tagging

7.3.1 Final state tag

No ambiguity exists on the final state flavour of the B_s signal events, the electric charge sign of the reconstructed D_s candidate is taken. Events where the D_s^- is produced via a W-boson have always the wrong tag. In the rest of the background, the tag depends the real nature of each event, and it is evaluated from the simulation.

7.3.2 Initial state tag

The flavour state at production time is estimated using information from the two event hemispheres, the hemisphere where the B_s candidate is reconstructed and the hemisphere

opposite to it. The relevant information is combined into a single tag variable.

For the opposite side hemisphere, the combined variable described in Section 6.10.1 is used.

For the hemisphere of the B_s candidate decay, discriminant variables are defined and combined with the opposite side flavour estimator with a neural network. All the charged particles from the B_s decay are identified, and therefore excluded from the charge estimators. An approach similar to that developed in Section 6.10.2 for the inclusive lepton selection is taken. A "big jet" built to contain all the particles from the fragmentation and hadronization of the b-quark is defined with the JADE algorithm, (Section 5.2.1) with a jet-resolution parameter $y_{\rm cut} = 0.02$. Charged particles in this jet, which are not included among the B_s decay products, are used to define four charge estimators: *i*) three jet charges with $\kappa =$ 0.0, 0.5, and 1.0; and *ii*) a fragmentation kaon estimator defined as in Section 6.8.1 (trained on hadronic B_s decays), the neural network estimator of the most likely fragmentation kaon is taken with the sign of the electric charge of the candidate.

These four variables are combined with other initial state charge estimators: the polar angle of the decay axis of the B_s candidate and the opposite side flavour estimator. In addition, three control variables are also used in the neural network training, *i*) the reconstructed momentum of the B_s candidate; *ii*) the number of charged particles in the "big jet"; and *iii*) the spread of weights of the charged particles in the opposite hemisphere (see Section 6.10.1) to obtain the initial state tag variable.

The value of the initial state tag variable for an event is used both to determine the initial b-quark flavour of the B_s candidate at production, and to estimate the probability of this flavour determination to be wrong. The mistag probability is parametrized as a function of this variable for all the sample components both for the $D_s\,\ell$ selection and fully reconstructed B_s candidates. The effective mistag probability evaluated on simulated events for the signal is in both cases $\eta_{\rm eff} \simeq 24\%$.

A B_s candidate is tagged as unmixed (mixed) when the reconstructed initial and final flavour states are the same (opposite).

Chapter 8

Results on B_s Oscillations

The selection and characterization of an inclusive semileptonic event sample is presented in Chapter 6. This data sample, together with the supplementary samples described in Chapter 7, are used to constrain the B_s oscillation frequency. In this Chapter, the results obtained with the three data samples separately are presented, followed by the ALEPH combination and the world combination.

In Section 8.1, the likelihood function used to fit the B_s oscillation frequency is described. Details on the input parameters in the case of the inclusive semileptonic event sample are given. The results obtained on this sample, with both the likelihood and the amplitude method, are presented in Section 8.1.2. The systematic uncertainties are discussed in Section 8.2, and the final result for the inclusive semileptonic event sample is described in Section 8.3. The fitting procedure was checked with the help of simulated events. The results of these checks are shown in Section 8.4.1. The impact of the event-by-event evaluation of the decay length uncertainty (Section 6.3) is explained in Section 8.4.2. The results obtained with the $D_s\,\ell$ events and the selection of fully reconstructed B_s mesons is presented in Section 8.5. The B_s oscillation results are combined first within the ALEPH experiment, and then with all results available to date in Section 8.6. The global combination presents a hint for a B_s oscillation signal in the frequency region around $17.5\,\mathrm{ps}^{-1}$. The interpretation of this hint in terms of a possible statistical fluctuation is given in Section 8.6.3.

8.1 The Δm_s fit

The likelihood function is built from the proper time probability density functions defined in Section 3.2 for every sample component. Each b-hadron species is described by its theoretically expected *p.d.f.* folded with the expected experimental effects. The non-b background proper time distribution is taken from the simulation. The likelihood function therefore reads

$$L \;=\; \prod_{i}^{N_{\mathrm{unmix}}} p.d.f.^{\mathrm{unmix}}(t) \;\times\; \prod_{i}^{N_{\mathrm{mix}}} p.d.f.^{\mathrm{mix}}(t)\;. \tag{8.1}$$

where $p.d.f.^{\mathrm{unmix}}(t)$ and $p.d.f.^{\mathrm{mix}}(t)$ are the sum of the contributions of the four b-hadron species and the non-b background. As explained in Section 6.10.3, each event is attributed a

"mixed" or "unmixed" label depending on the charge correlation between the initial and final state flavour estimators. This label is used to choose the *p.d.f.* to be used for each event. In the following, $N_{\text{mix (unmixed)}}$ represents the number of events labeled as mixed (unmixed) in the sample.

The *p.d.f.*'s which describe the proper time distribution of the four b-hadron species in Eq. 8.1 are folded with a resolution function and with the proper time reconstruction efficiency (Section 6.7.3). In the case of the inclusive semileptonic event sample, the resolution on the reconstructed b-hadron momentum is obtained from the simulation. The decay length uncertainty is that obtained from the event-by-event vertex fit with a pull correction (Section 6.3.8). As explained in Section 6.7.1, the relative momentum resolution is described by the sum of two Gaussians of widths σ_{p_1} and σ_{p_2}, and relative fractions f_{p_1} and f_{p_2} (with $f_{p_2} = 1 - f_{p_1}$). The proper time resolution function therefore reads

$$\mathcal{R}(t, t_0) = \sum_{i=1}^{2} f_{p_i} \frac{1}{\sqrt{2\,\pi}\,\sigma_{t_i}} \exp\left[-\frac{1}{2}\left(\frac{t - t_0}{\sigma_{t_i}}\right)^2\right] , \tag{8.2}$$

where σ_{t_i} is the proper time resolution obtained as the convolution of the decay length and momentum resolutions. It can be written as

$$\sigma_{t_i} = \sqrt{\left(\frac{m_{\mathrm{B}}}{p_{\mathrm{B}}\,c}\,\sigma_l\right)^2 + \left(t\,\frac{\sigma_{p_i}}{p_{\mathrm{B}}}\right)^2} . \tag{8.3}$$

The small effect of the correlation between the reconstructed momentum and decay length is neglected.

The proper time distribution expected for light-quark hadrons and for c-hadrons is *a priori* different because their characteristic lifetime are different. However, the selection bias introduced by the inclusive semileptonic event selection on the light-quarks and charm events is such that the two distributions are found to be very similar in the simulation, as shown in Fig. 8.1.

A single function is parametrized from the simulation for the non-b background component. This function includes the resolution and reconstruction inefficiency effects by construction, and it is used to build the complete likelihood function.

8.1.1 Input description

Each reconstructed b-hadron candidate is given a probability of originating from a B_s, a B_d, a B^+, a b-baryon or the non-b background. These probabilities are obtained from the b-tagging (Section 6.4) and the B_s enrichment (Section 6.9) neural network combined variables, as described in Section 6.4. The distributions of these probabilities for the selected data sample, compared to that of the simulation sample, are shown in Fig. 8.2. The disagreement observed between data and simulation for each sample component probability is due to the discrepancy in the B_s enrichment variable previously discussed. This discrepancy is taken into account in the evaluation of the systematic uncertainties (Section 8.2).

The mistag probability is estimated for each event and each sample component from the $b \to \ell$ enrichment (Section 6.5) and the initial state tag (Section 6.10) variables. The mistag probability for the four b-hadron species in data and simulation is shown in Fig. 8.3.

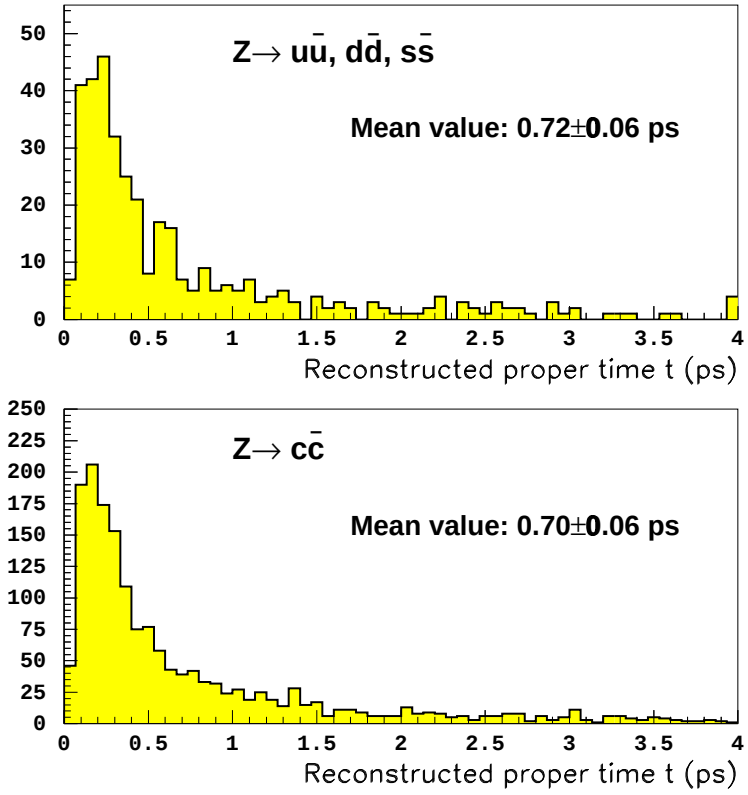

Figure 8.1: Proper time distribution for $u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and charm background from simulated events.

The mistag probability for the non-b background is found to be close to 50%, as intuitively expected.

The flavour tagging of the inclusive semileptonic sample analysis was optimized to obtain the best performance on the B_s signal. As a consequence, the smallest values of the mistag probability are indeed obtained for the B_s component of the sample. The combined effective mistag for the B_s mesons in the whole sample is $\eta_{\rm eff} \simeq 30\%$, while it is $\eta_{\rm eff} \simeq 33\%$ for B_d mesons, and $\eta_{\rm eff} \simeq 45\%$ for B^+ mesons. It was checked that the fraction of events tagged as mixed in the data sample agrees with the simulation prediction:

$$\left.\frac{N^{\rm mix}}{N^{\rm mix}+N^{\rm unmix}}\right|_{\rm Data} = 0.453 \pm 0.003 \quad \text{and} \quad \left.\frac{N^{\rm mix}}{N^{\rm mix}+N^{\rm unmix}}\right|_{\rm Sim} = 0.450 \pm 0.002 \ . \quad (8.4)$$

8.1.2 Results on the inclusive semileptonic event sample

The Likelihood

Once the likelihood function L for the data sample is built (Section 8.1), the most probable value for Δm_s is looked for as the value which maximizes the likelihood (or which minimizes the negative log likelihood $\mathcal{L} \equiv -\log L$ instead). The shape of the function $\mathcal{L}$ for the inclusive

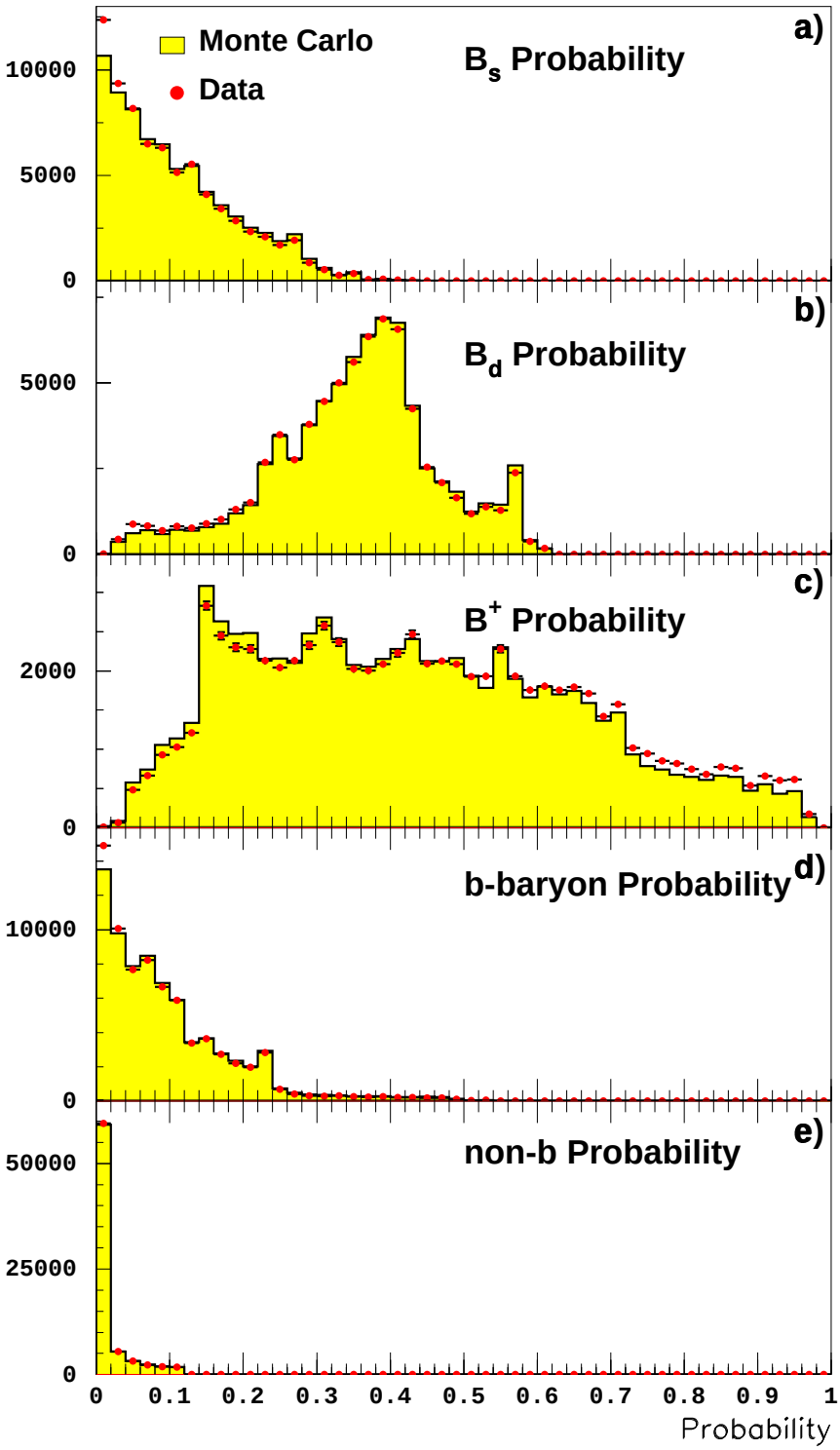

Figure 8.2: Probability distribution for all sample components in the inclusive semileptonic event sample; a) B_s, b) B_d, c) B^+, d) b-baryon, and e) non-b probability for all events.

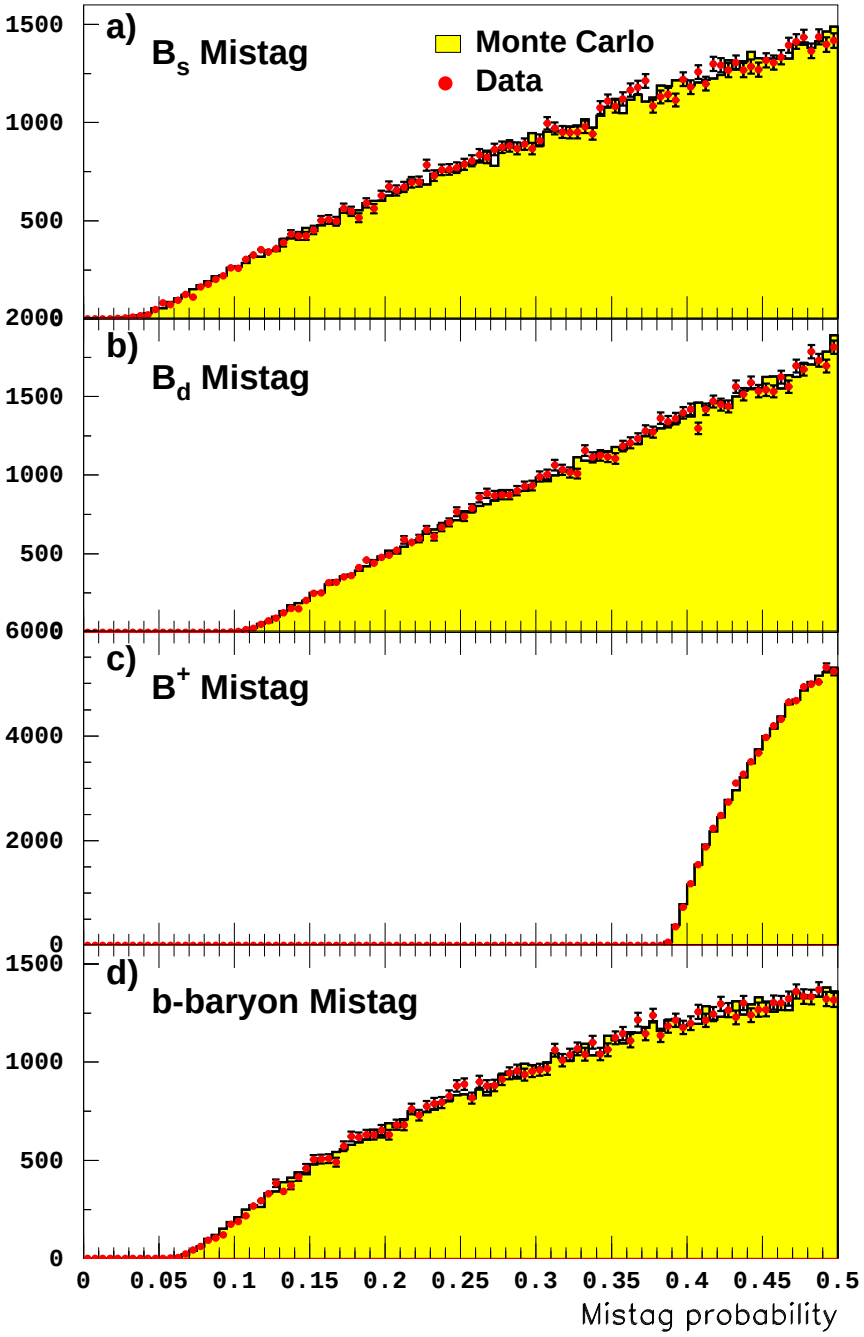

Figure 8.3: Mistag probability distribution for all b-hadron species in the inclusive semileptonic event sample; a) B_s, b) B_d, c) B^+, and d) b-baryon, mistag probability for all events.

semileptonic data sample as a function of the Δm_s free parameter is shown in Fig. 8.4.

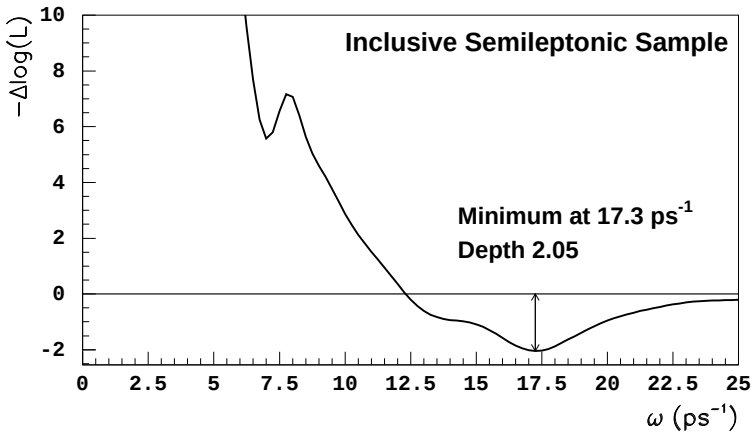

Figure 8.4: Function $\mathcal{L}(\Delta m_s)$ for the inclusive semileptonic data sample as a function of Δm_s.

A minimum which corresponds to slightly more than two standard deviations is found at $\Delta m_s = 17.3^{+1.3}_{-1.5}\,\mathrm{ps}^{-1}$. The uncertainties quoted here correspond to the values for which $\mathcal{L} = \mathcal{L}_{\min} + 1/2$. They would provide the one standard deviation statistical uncertainty if the likelihood function was parabolic in a wide range around the minimum, which is not the case with the sensitivity obtained today (the same procedure is followed for the other likelihood minima discussed in this Chapter). Two sigmas is not considered to be sufficient to claim the observation of an oscillation signal, and more data would be needed to confirm this possible measurement.

The Amplitude Method

In the amplitude method (Section 3.3), the likelihood function described above is modified. An amplitude $\mathcal{A}$ is introduced in front of the oscillating term of the probability density function of the B_s signal (as shown in Eq. 3.10). The amplitude is fit by minimizing $\mathcal{L}(\Delta m_s)$ with respect to $\mathcal{A}$ for each value of the oscillation frequency in the range of interest.

This method is used in this thesis to set limits from the individual event samples and to perform the combination of several results. The amplitude spectrum obtained for the inclusive semileptonic event sample is shown in Fig. 8.5. Only statistical uncertainties are considered at this stage. The dashed line in Fig. 8.5 presents the total uncertainty multiplied by 1.645 as a function of the frequency. The crossing point of this line with 1 defines the sensitivity (or expected lower limit at 95% C.L.) of the analysis. As expected, a significant deviation from $\mathcal{A} = 0$ is observed in the frequency range in which the minimum of the likelihood was found.

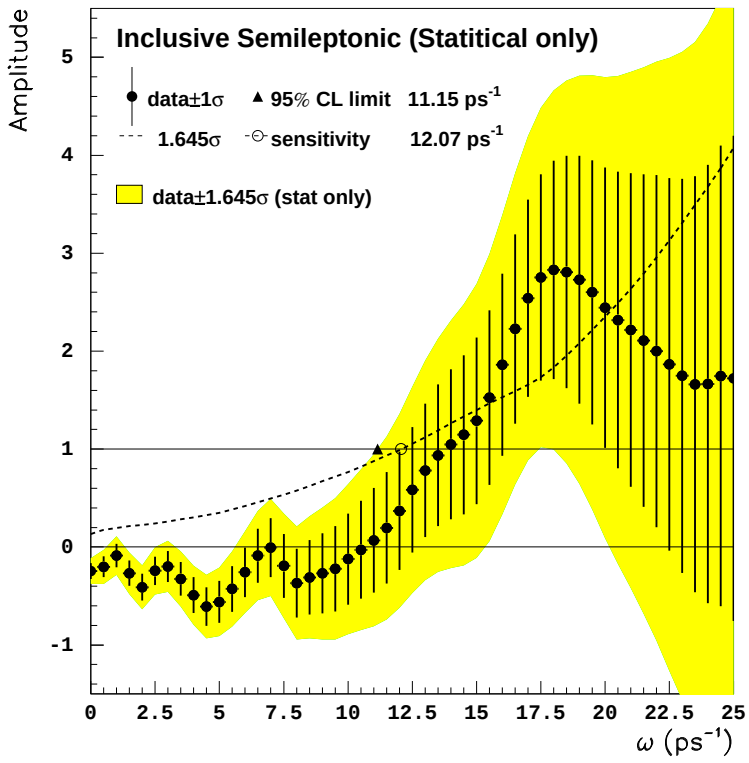

Figure 8.5: Amplitude spectrum for the inclusive semileptonic sample. Only statistical uncertainties are considered.

8.2 Systematic uncertainties

The systematic uncertainties are evaluated on the measured amplitudes rather than on the position of the likelihood minimum, because of the low significance of the minimum observed. The input parameters in the fit are varied within their uncertainties and the amplitude spectrum is rebuilt in each case. The variations in the fit amplitude with respect to Fig. 8.5 are taken as the systematic uncertainty. All systematic uncertainties studied for the inclusive semileptonic event sample are summarized in Table 8.1. The statistical uncertainty on the amplitude at $\Delta m_s = 0, 10, 15, 20$ and 25 ps^{-1} is compared with each of the systematics at the same point.

With the inclusive semileptonic sample, the most important systematic uncertainty on Δm_s, in particular at high values of the oscillation frequency, comes from f_s, the fraction of B_s mesons in an unbiased b-hadron sample. As of today, it has still a relatively large uncertainty, $f_s = (10.7 \pm 1.4)\%$ [13].

For each event, the decay length is obtained from the b vertex fit and from the bias correction, and the decay length uncertainty from the b vertex fit and from the pull correction (Sections 6.3.6 and 6.3.8). Both corrections are obtained from simulated events. They are therefore reliable only to the extent that the vertexing algorithm has the same performance on data and simulation. To check the latter point, a specific analysis was performed. Events

Δm_s	0 ps^{-1}	10 ps^{-1}	15 ps^{-1}	20 ps^{-1}	25 ps^{-1}
$\sigma_{\mathcal{A}}^{stat}$	$\pm$0.081	$\pm$0.466	$\pm$0.851	$\pm$1.430	$\pm$2.477
f_s	$\pm$0.087	$\pm$0.017	$\pm$0.206	$\pm$0.334	$\pm$0.348
Decay length Resolution	$\pm$0.001	$\pm$0.015	$\pm$0.125	$\pm$0.202	$\pm$0.111
Momentum Resolution	$\pm$0.008	$\pm$0.008	$\pm$0.117	$\pm$0.214	$\pm$0.209
τ_{B_s}	$\pm$0.006	$\pm$0.016	$\pm$0.069	$\pm$0.122	$\pm$0.145
Mistag	$\pm$0.070	$\pm$0.047	$\pm$0.010	$\pm$0.020	$\pm$0.101
$b \to \overline{c} \to \ell$	$\pm$0.065	$\pm$0.046	$\pm$0.069	$\pm$0.023	$\pm$0.103
$b \to \ell / b \to c \to \ell$	$\pm$0.014	$\pm$0.007	$\pm$0.009	$\pm$0.018	$\pm$0.026
Enrichment	$\pm$0.051	$\pm$0.051	$\pm$0.010	$\pm$0.023	$\pm$0.063
Δm_d	$\pm$0.036	$\pm$0.001	$\pm$0.001	$\pm$0.001	$\pm$0.003
Total Systematic Uncertainty	$\pm$0.145	$\pm$0.088	$\pm$0.286	$\pm$0.463	$\pm$0.473

Table 8.1: Systematic uncertainties on the amplitude at different Δm_s values compared to the statistical uncertainty at the same point in the case of the inclusive semileptonic event sample.

arising from a Z decay into light quarks were selected and the vertexing algorithm was used on these events to reconstruct the primary vertex. The "decay length" was defined as the distance between the primary vertex reconstructed with the vertexing algorithm and that with the standard method (Section 5.4). The "decay length" distribution was used to compare the performance of the vertexing algorithm in data and simulation. The event selection to perform this test was almost the same as that described in Sections 6.2 and 6.3. The only difference concerned the lepton identification. In light-quark decays of the Z, the leading track is not expected to be a lepton, therefore the electron and muon particle identification criteria were not applied here. Then, a cut on the b-tagging variable $N_{\rm btag} < -0.75$ (Section 6.4) was used to reduce the amount of $Z \to b\overline{b}$ events. The resulting sample composition is 86.2% $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, 12.4% $c\overline{c}$ and 1.4% $b\overline{b}$. The distribution of the distance between the primary vertex reconstructed with the vertexing algorithm developed for this thesis and the standard one is shown in Fig. 8.6 for data and simulation.

The two distributions in Fig. 8.6 are similar. It indicates that the vertexing algorithm has the same performance on data and simulation, at least for the events selected to produce Fig. 8.6. This exercise is, however, not sensitive to the effect of misidentified particles in building the "charmed" particle.

Finally, a variation of the decay length resolution by $\pm$3% was considered to take into account possible remaining discrepancies on the vertex algorithm performance between data and simulation. A variation of $\pm$10% was taken for the relative momentum resolution.

The initial state tagging variable was, for the $B_s \to \ell$ signal events, perfectly calibrated with slope 1. For systematic studies this slope is varied by 5%, following the prescription in Ref. [85].

As explained in Section 6.9, a discrepancy in the enrichment variable distribution between data and simulation is seen. Weights are applied to simulation events to increase by 8% the amount of charged b-hadrons and improve the agreement with the data. These weights are

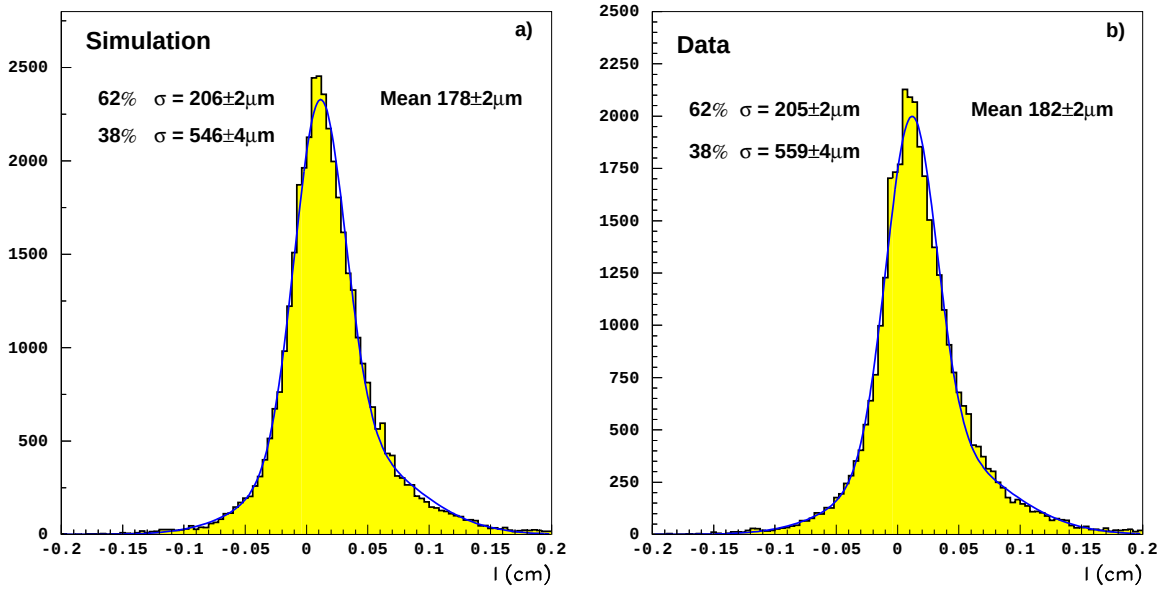

Figure 8.6: Distribution of the distance between the primary vertex reconstructed with the vertexing algorithm and the standard one, a) for the simulation and b) for the data.

removed for the systematic uncertainty evaluation.

The relevant physical quantities for the B_s oscillation fit considered as possible sources of systematic uncertainties are shown, with their uncertainties, in Table 8.2.

f_s	(10.7 ± 1.4) % [13]
τ_{B_s}	(1.54 ± 0.07) ps [13]
$\mathrm{Br}(b \to \overline{c} \to \ell)$	(1.62 ± 0.44) % [86]
$\mathrm{Br}(b \to \ell)$	(10.62 ± 0.17) % [86]
$\mathrm{Br}(b \to c \to \ell)$	(8.07 ± 0.25) % [86]
Δm_d	(0.464 ± 0.018) ps^{-1} [13]

Table 8.2: Physical inputs to the B_s oscillation fit.

The values of the $b \to \ell$ and $b \to c \to \ell$ branching ratios in Table 8.2 are correlated; the correlation coefficient is -0.37. The systematic uncertainty quoted in Table 8.1 is the combination of the two effects taking into account their correlation.

A variation of Δm_d only affects the amplitude spectrum in the frequency region $\omega \sim \Delta m_d$. It is one of the dominant systematic effects at low frequency, but it is completely negligible at the frequency range of interest ($\omega \gtrsim 15\,\mathrm{ps}^{-1}$).

8.3 Final result for the inclusive semileptonic sample

The systematic uncertainties evaluated in the previous Section are included to produce the complete amplitude spectrum for the inclusive semileptonic event sample. This spectrum is shown in Fig. 8.7

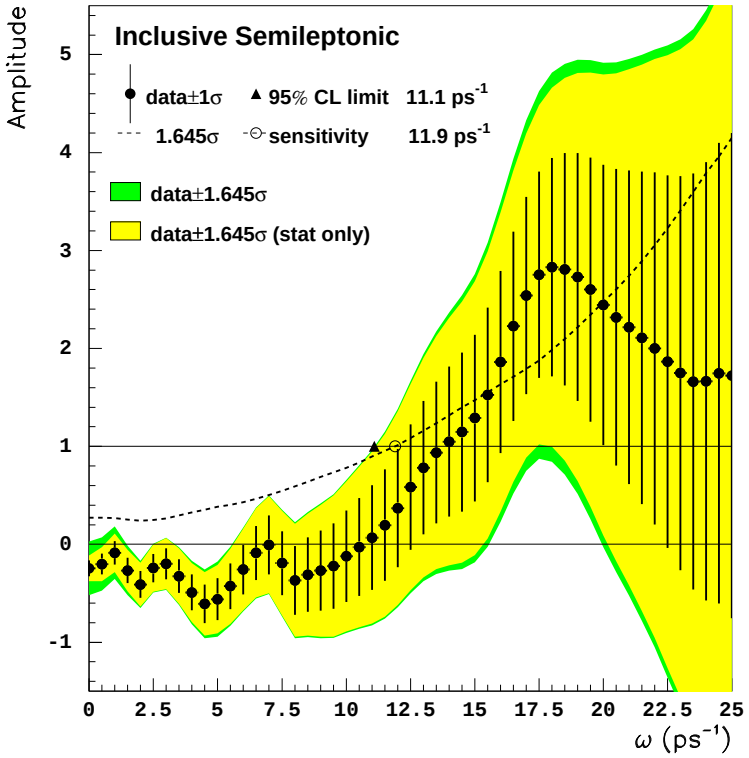

Figure 8.7: Amplitude spectrum for the inclusive semileptonic sample, statistical and systematic uncertainties are shown.

The expected lower limit at 95% C.L. (or sensitivity) of the analysis is $11.9\,\mathrm{ps}^{-1}$. The observed lower limit is smaller than that expected because of the amplitude excursion around $\Delta m_s \sim 17.5\,\mathrm{ps}^{-1}$. A lower limit of $11.1\,\mathrm{ps}^{-1}$ is set at 95% C.L. on Δm_s with the inclusive semileptonic event sample.

As already seen in Table 8.1, systematic uncertainties are only dominant at small values of the oscillation frequency. The high frequency range is largely dominated by the statistical uncertainties. Therefore, a significant gain can be expected with a combination of independent analyses.

The curve of in Fig. 8.7 can be compared to that obtained by the previous analysis [10], reproduced for completeness in Fig. 8.8. In Fig. 8.7 and Fig. 8.8, a deviation from $\mathcal{A} = 0$ is observed in the same frequency range. All the improvements in the selection and characterization of the events (Chapter 6) are translated into a substantially higher analysis sensitivity. A sensitivity gain of more than $2\,\mathrm{ps}^{-1}$ is achieved. The performance of the two analyses can also be compared by the uncertainty on the measured amplitude at a given value of the test

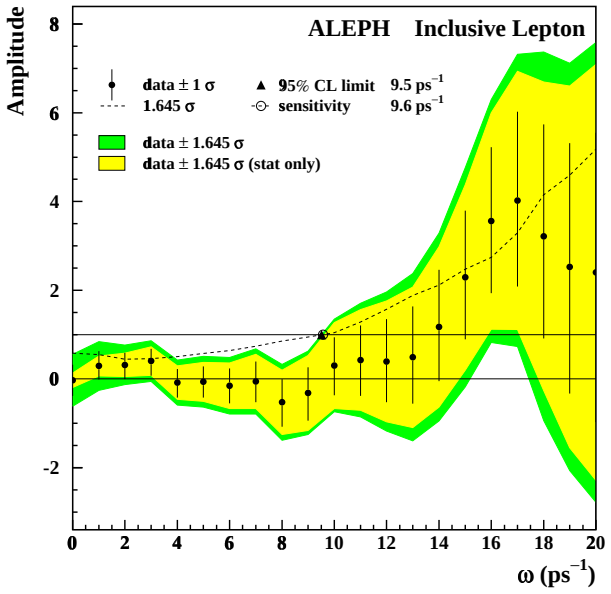

Figure 8.8: Amplitude spectrum for the *previous* inclusive semileptonic sample.

frequency. For the previous analysis this uncertainty amounts to ± 3.28 at $20\,\mathrm{ps}^{-1}$, and it is reduced to ± 1.50 here (the same comparison can be done at $17\,\mathrm{ps}^{-1}$, the amplitude uncertainty was ± 1.99 and it is now ± 1.11). The uncertainty on the amplitude at high frequency has been reduced by more than a factor of two.

The amplitude spectrum (including the systematic uncertainties) can be converted back into a likelihood function using Eq. 3.17 (Section 3.3). The corresponding curve is very similar to that in Fig. 8.4, with a minimum at the same position. The significance of the minimum is reduced due to the inclusion of systematic uncertainties: the depth of the likelihood is now -1.8 .

8.4 Checks

8.4.1 Checks with simulated events

Checks were performed on the simulated samples to ensure that the results obtained above are robust. The fitting procedure was tested with events in which the oscillation frequency is known.

A pure sample of $\mathrm{B_s}$ events was simulated, with $\Delta m_\mathrm{s} = 14\,\mathrm{ps}^{-1}$. The likelihood fit and the amplitude spectrum analysis were performed on that sample. The results are shown in Figs. 8.9 and 8.10.

The minimum of the likelihood is found close to $\Delta m_\mathrm{s} = 14\,\mathrm{ps}^{-1}$, as expected. The measured value of the oscillation frequency in that sample is $\Delta m_\mathrm{s} = 14.04^{+0.21}_{-0.17}\,\mathrm{ps}^{-1}$. The amplitude spectrum follows also the expectation: the measured amplitude is compatible with

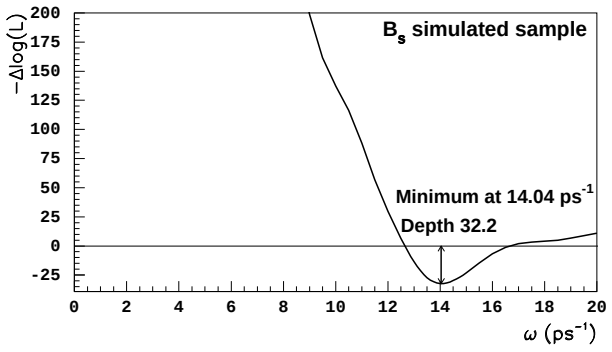

Figure 8.9: Function $\mathcal{L}(\Delta m_s)$ for a simulated sample of B_s events as a function of Δm_s.

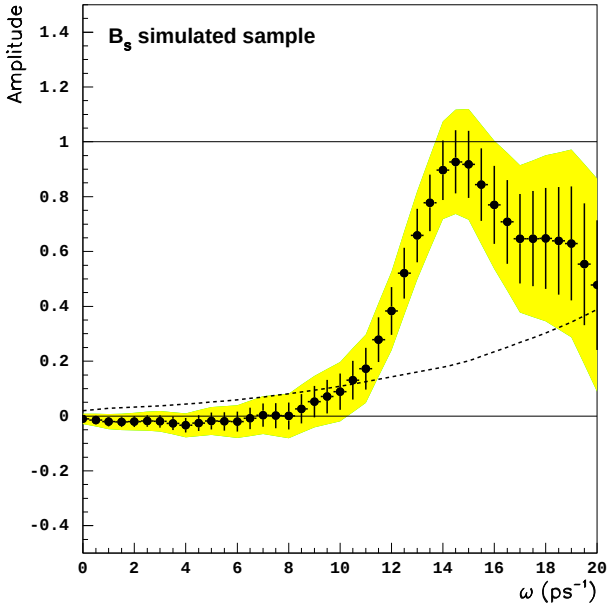

Figure 8.10: Amplitude spectrum for a sample of B_s simulated events.

zero for values below the true oscillation frequency and it is compatible with one at the true frequency. The fact that both, the likelihood function and the amplitude spectrum, have the expected shape for $\Delta m_s = 14\,\mathrm{ps}^{-1}$ indicates that the B_s signal events are correctly treated in building the likelihood.

The same checks were performed with the $Z \to q\overline{q}$ simulated sample, in which the oscillation frequency for the B_s events was also fixed at $\Delta m_s = 14\,\mathrm{ps}^{-1}$. The results both for the likelihood function and for the amplitude spectrum are shown in Figs. 8.11 and 8.12.

A minimum is "observed" in the likelihood function at $\Delta m_s \sim 14\,\mathrm{ps}^{-1}$. Its depth (4.5) corresponds to three standard deviations. The "measured" value of the oscillation frequency in that sample is $\Delta m_s = 14.0 \pm 0.8\,\mathrm{ps}^{-1}$. The measured value in the $q\overline{q}$ simulation indicates

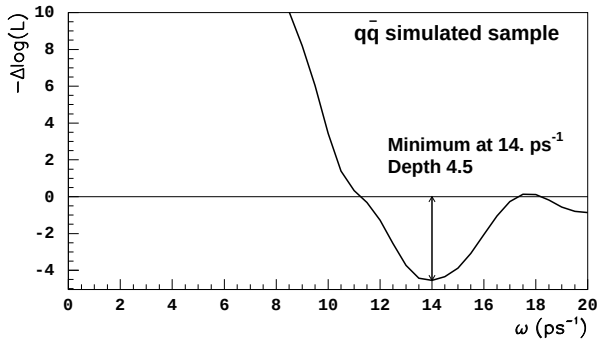

Figure 8.11: Function $\mathcal{L}(\Delta m_s)$ for a simulated sample of $q\overline{q}$ events as a function of Δm_s.

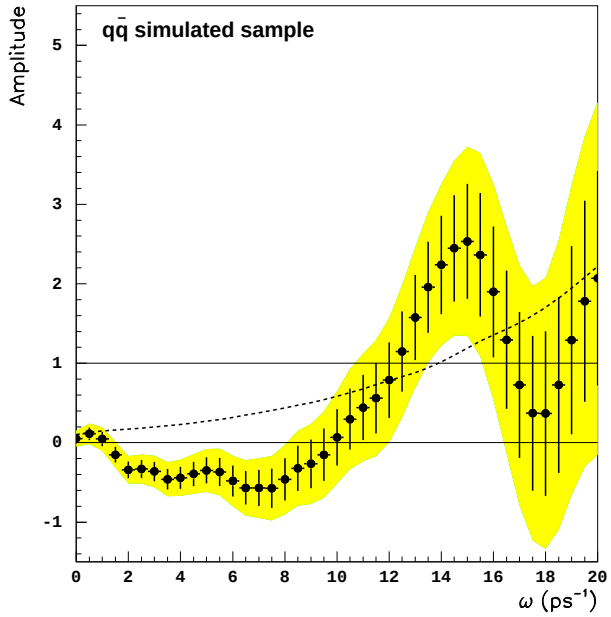

Figure 8.12: Amplitude spectrum for a sample of $q\overline{q}$ simulated events with $\Delta m_s^{\rm true} = 14\,{\rm ps}^{-1}$. The statistics is this sample are a factor 1.6 larger than the data sample.

that the complete description of the simulated sample is correctly implemented in the fitting procedure, at least for what concerns the B_s oscillation measurement.

The amplitude spectrum in Fig. 8.12 presents a deviation from $\mathcal{A} = 0$ around the true oscillation frequency, as expected. However, in a frequency range between $2\,{\rm ps}^{-1} - 8\,{\rm ps}^{-1}$, the measured amplitude is significantly below zero. A compatible feature is found in the data (Fig. 8.7). As this effect is not observed in the pure B_s sample, it is understood to be due to a misparametrization of some properties of the background components. This misparametrization does not affect the B_s component and does not bias the measured Δm_s, as seen from the likelihood analysis. It is, however, being studied in more detail for the final publication of ALEPH results on B_s oscillations.

8.4.2 Impact of the uncertainty treatment

A rough estimation of the improvement achieved here with respect to the previous analysis can be obtained with the comparison of the statistical power of the two analyses. If resolution effects are ignored, three parameters determine the statistical power of a B_s oscillation analysis (Section 3.3), *i*) the number of selected events; *ii*) the effective signal purity; and *iii*) the effective mistag probability. The number of events selected has doubled. The effect of the B_s enrichment was evaluated in the previous analysis to bring an effective increase of the sample B_s purity of 13%. This increase is $\sim 25\%$ for the present analysis (Section 6.9). Improvements on the flavour tagging were also performed (Sections 6.5 and 6.10). However, the average fraction of $b \to \ell$ decays in the data sample selected here is smaller than that of the previous analysis. All in all, the overall effective mistag probability is similar in the two analyses.

With all these numbers, a reduction factor of ~ 1.5 in the statistical uncertainty is expected from Eq. 3.27. As discussed in Section 8.3, the actual improvement observed is larger than this expectation. The improvement factor at $\omega = 20\,\mathrm{ps}^{-1}$ is ~ 2.2. The new vertexing algorithm and the parametrization of the pull corrections account for most of the additional uncertainty reduction.

In Section 3.3.5, an exercise with toy experiments was performed to illustrate the effect of the event-by-event estimate of the decay length uncertainty, with some choice of the resolution parameters. In the case of this analysis, the effect is even more sizeable. For illustration, the amplitude analysis is performed on the data sample with the average uncertainty used for all events (73% of the events in a core with $267\,\mu$m resolution and a tail with $806\,\mu$m). The decay length bias is also corrected globally. The uncertainty on the measured amplitude as

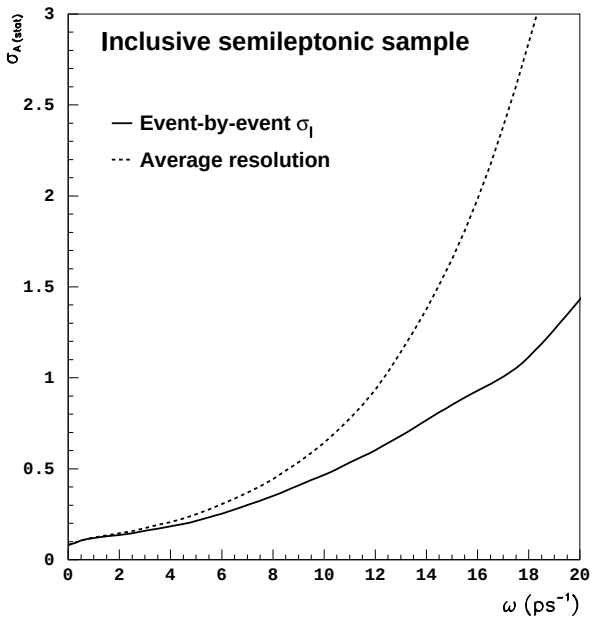

Figure 8.13: Amplitude uncertainty as a function of the test frequency. The results obtained using an event-by-event estimate of the decay length uncertainty are compared with those obtained with the average decay length uncertainty.

a function of the test frequency is shown in Fig. 8.13 compared to the uncertainty obtained when the event-by-event decay length uncertainties are used.

As the frequency increases, the difference between the amplitude uncertainties found in the two cases becomes larger. At $\omega = 20\,\mathrm{ps}^{-1}$, the uncertainty obtained with the average decay length uncertainty is almost a factor three larger than that obtained with the event-by-event treatment.

In the previous analysis, an intermediate treatment between the use of the average uncertainty with a global pull correction and the event-by-event treatment described in Section 6.3 was performed. The decay length uncertainty was evaluated event by event but a global pull correction was applied to the whole event sample. As a consequence, a dilution was introduced on the event-by-event treatment and the sensitivity of the analysis was reduced.

8.5 Results with supplementary event selections

The same likelihood function described in Section 8.1 is used here to study the B_s oscillation frequency in the event samples described in Chapter 7. The evaluation of systematic uncertainties is in both cases similar to that in Section 8.2.

The likelihood function and the amplitude spectrum obtained with the $D_s\,\ell$ data sample are shown in Figs. 8.14 and 8.15.

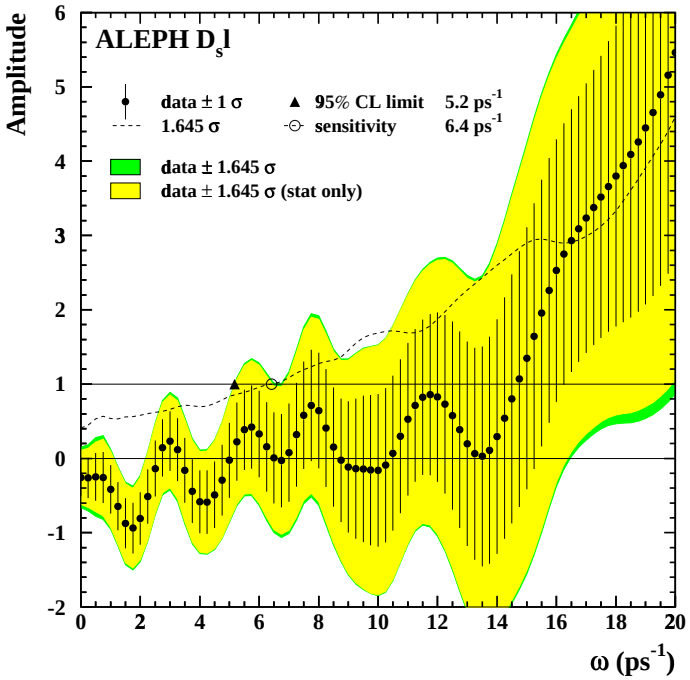

Figure 8.14: Amplitude spectrum for the $D_s\,\ell$ event selection.

A minimum in the likelihood function (and a deviation from zero in the amplitude spectrum) is observed at $\Delta m_s = 17.5\,\mathrm{ps}^{-1}$. This minimum is less deep than that found in the

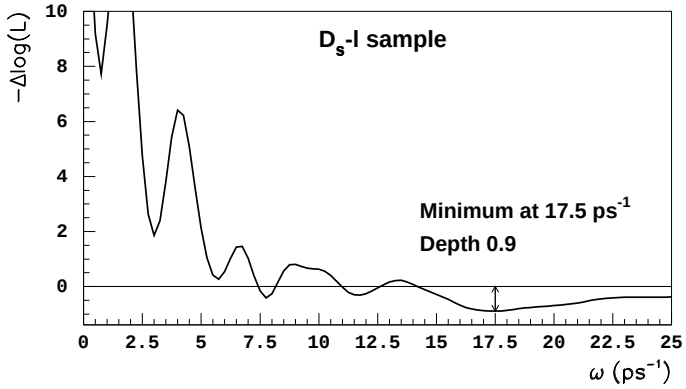

Figure 8.15: Function $\mathcal{L}(\Delta m_s)$ for the $D_s\,\ell$ data sample as a function of Δm_s.

inclusive semileptonic event sample, but it is located in the same frequency range. The sensitivity ($6.4\,\mathrm{ps}^{-1}$) is not as large as that of the inclusive semileptonic sample, but the performance of the analysis is good enough to make it relevant for the complete ALEPH results on B_s oscillations.

The amplitude spectrum obtained with the exclusive B_s selection is shown in Fig. 8.16. The corresponding likelihood function does not present a significant minimum in the frequency range of interest.

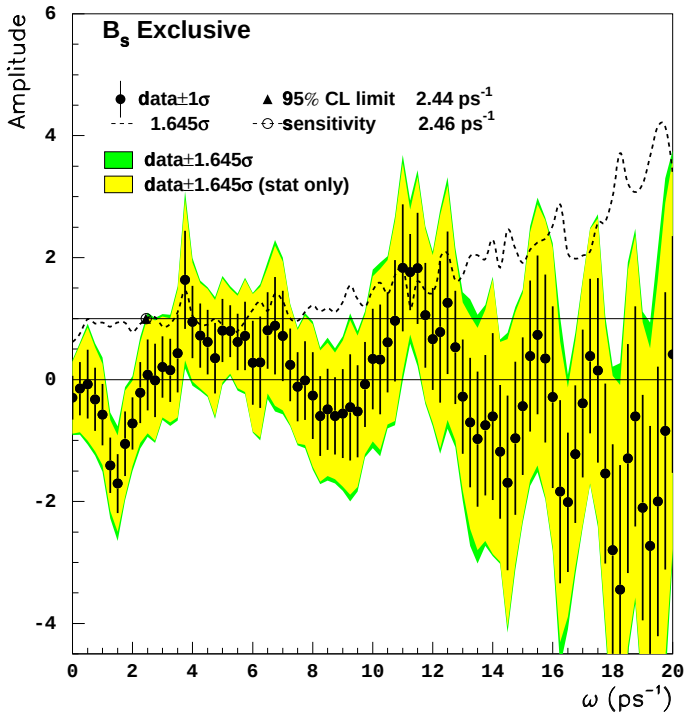

Figure 8.16: Amplitude spectrum for the selection of exclusive hadronic B_s decays.

Very few events are selected in this sample, but their excellent resolution, both for the decay length and the momentum, provides a rather flat shape of the amplitude uncertainty as a function of the test frequency. The analysis is therefore relevant in the ALEPH combination, in particular at large frequency values, even though it does not constrain significantly the value of Δm_s by itself. Because of the accurate momentum measurement, the correlation between two consecutive measured amplitude values is much smaller than in the other two data samples (Section 3.3.4).

A comparison of the performance of the three analyses presented in this thesis is shown in Fig. 8.17. In this figure, the statistical uncertainty on the measured amplitude is displayed as a function of the test frequency for the three analyses.

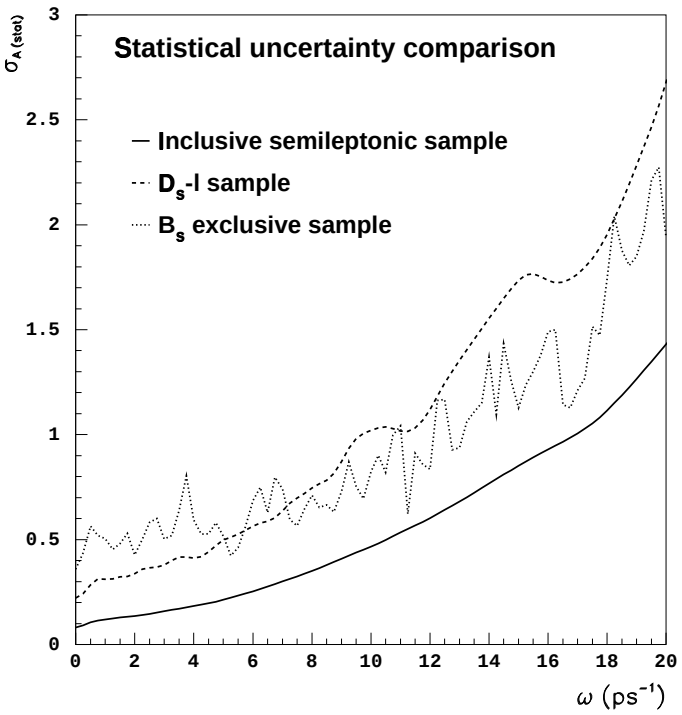

Figure 8.17: Amplitude uncertainty for the three event selections.

The inclusive semileptonic event selection is the most sensitive in the whole frequency range explored. At small values of the frequency, the $D_s\,\ell$ sample is the second most sensitive selection. However, at large frequency values, the exclusive selection becomes more sensitive than the $D_s\,\ell$ as a consequence of its excellent proper time resolution. The fast variation of the amplitude uncertainty as a function of the test frequency in the case of the fully reconstructed B_s mesons sample is due to the reduced statistics of that sample. Different subsets of events drive the precision on the measured amplitude at different values of the test frequency. The number of events in each subset, their resolution and their mistag probability determine the uncertainty on the measured amplitude in each case.

8.6 Combination with other results

The results of the three B_s oscillation analyses presented here are combined first with the other ALEPH analysis on the subject, and then with all available results in the world to obtain the most accurate result possible today.

8.6.1 Combination with other ALEPH results

In addition to the three analyses described in this thesis, another analysis on B_s oscillations was performed with the ALEPH data [61]. Fully reconstructed D_s candidates were combined with an oppositely charged hadron to form B_s decay candidates. The performance of this analysis is substantially worse than that of the analyses presented here, especially at high frequency (*e.g.*, the statistical uncertainty on the amplitude at $\omega = 20\,\mathrm{ps}^{-1}$ is more than seven times larger than in the inclusive semileptonic event sample). Nonetheless, the analysis is combined, for completeness, with the three discussed in this thesis.

Some of the events selected in the $D_s\,\ell$ sample are also selected in the inclusive semileptonic sample. To avoid any statistical correlation between the two samples, the events in common (~ 150) were removed from the sample in which their treatment is less accurate: the inclusive semileptonic event sample. The amplitude spectrum for this latter sample was redone before the combination was performed. A display of such an event is presented in Fig. 8.23, at the end of this Chapter.

The amplitude spectra of the four ALEPH analyses are combined. The result of this combination and the corresponding likelihood function are displayed in Figs. 8.18 and 8.19.

The combined ALEPH sensitivity for B_s oscillations is $13.8\,\mathrm{ps}^{-1}$ and a lower limit for the oscillation frequency is set at $\Delta m_s > 10.7\,\mathrm{ps}^{-1}$ at 95% C.L. (this limit is lower than that set with the inclusive semileptonic event sample alone because of the positive statistical fluctuation around $\omega \sim 11\,\mathrm{ps}^{-1}$ observed in the B_s exclusive sample). A minimum which corresponds to slightly more than 2.5 standard deviations is found at $\Delta m_s = 17.3\,^{+0.4}_{-0.3}\,\mathrm{ps}^{-1}$.

8.6.2 World combination

The ALEPH results on B_s oscillations are the most sensitive (or second most) in the world in the oscillation frequency range explored. However, a significant gain on the overall sensitivity is obtained by combining these results with those from other experiments. In particular the SLD experiment [60, 87], and to a lesser extent the DELPHI experiment [57, 59, 88], have B_s results competitive in the high frequency range.

A comparison of the statistical uncertainty on the measured amplitude as a function of the test frequency for all experiments which have results on the subject is shown in Fig. 8.20. The world combination is also shown. It can be noted that the SLD uncertainty curve is not a steep function of the frequency as those from LEP (in particular from DELPHI and OPAL). This difference is due to the decay length resolution: it is on average much better at SLD than at LEP ($\sim 100\,\mu$m compared to $\sim 250\,\mu$m for an inclusive event selection).

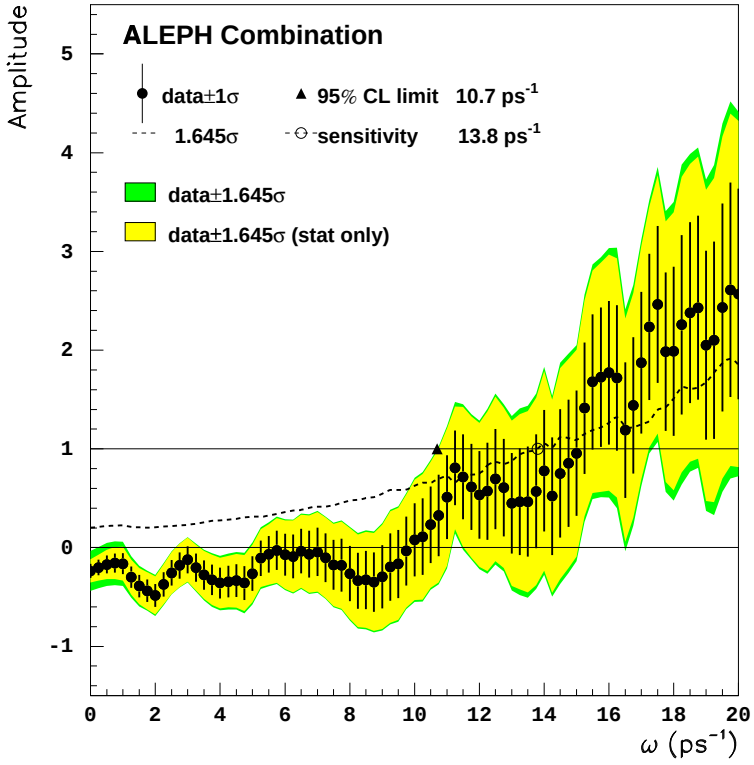

Figure 8.18: Combination of ALEPH results on B_s oscillations, amplitude spectrum.

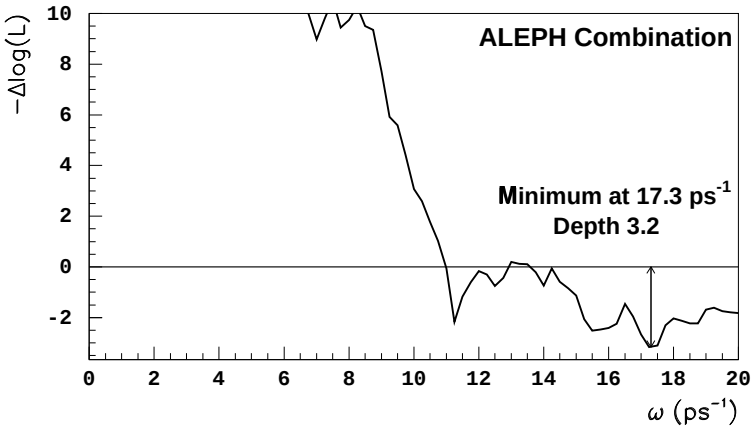

Figure 8.19: Likelihood function for the combination of ALEPH results on B_s oscillations.

The ALEPH results have the smallest statistical uncertainty on the measured amplitude up to a frequency of $\sim 17\,\mathrm{ps}^{-1}$ where SLD results become more competitive. Therefore, the ALEPH results drive most of the the world combination on B_s oscillations.

The amplitude spectrum for the combination of all B_s oscillations results up to date is

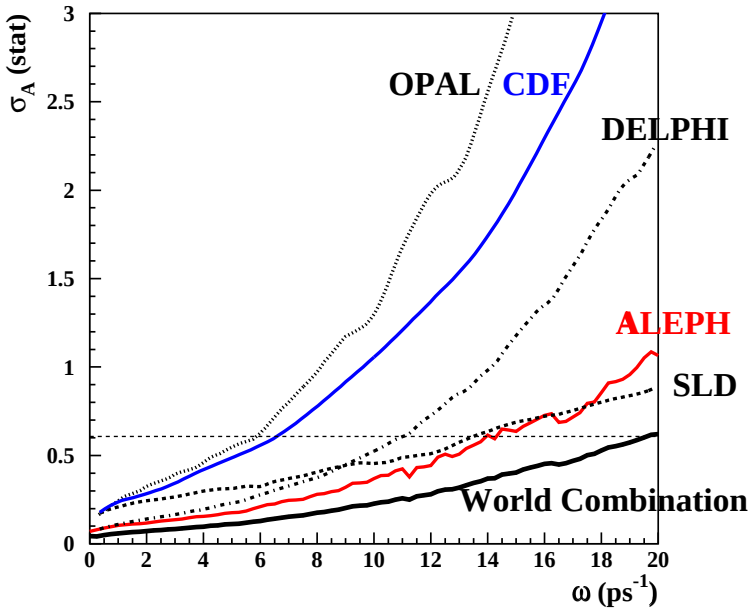

Figure 8.20: Comparison of the statistical uncertainty on the measured amplitude of all B_s results available today. The analyses are combined to provide a single figure per experiment.

shown in Fig. 8.21. The corresponding likelihood function is displayed in Fig. 8.22.

The world combined sensitivity is $19.3\,\mathrm{ps}^{-1}$. A lower limit on the B_s oscillation frequency is set at $\Delta m_s > 15.1\,\mathrm{ps}^{-1}$ at 95% C.L. This limit is significantly smaller than the expected limit because of the deviation from $\mathcal{A} = 0$ observed around $\omega \sim 17.5\,\mathrm{ps}^{-1}$. The combined likelihood function presents a minimum with a significance of ~ 2.3 standard deviations. The measured value of the B_s oscillation frequency in the world combined sample is $\Delta m_s = 17.3 \pm 0.3\,\mathrm{ps}^{-1}$. As in the case of the inclusive semileptonic event sample, the significance of the likelihood minimum is not considered sufficient to claim a measurement of Δm_s, more data or more refined analyses would be needed to confirm the present observed hint.

8.6.3 Interpretation

At the time of writing, as shown in the previous Section, the minimum in the world combined likelihood function is not deep enough to claim a measurement of the B_s oscillation frequency. However, the minimum observed, together with the shape of the amplitude spectrum in Fig. 8.21, may indicate that the world combination is close to be sensitive to the actual Δm_s value. The probability that the effect observed is due to a statistical fluctuation is studied with the method described in Section 3.3.4.

Although the world combined amplitude spectrum evolved substantially since winter 1999 (Fig. 3.7 compared to Fig. 8.21), the probability of observing a minimum in the likelihood more unlikely than that actually observed is, as it was at that time, around 3%. The depth of the minimum has decreased, it was 2.9 and it is now 2.7, and the frequency value in which the minimum is found is now larger than it was. The two changes have effects on the probability

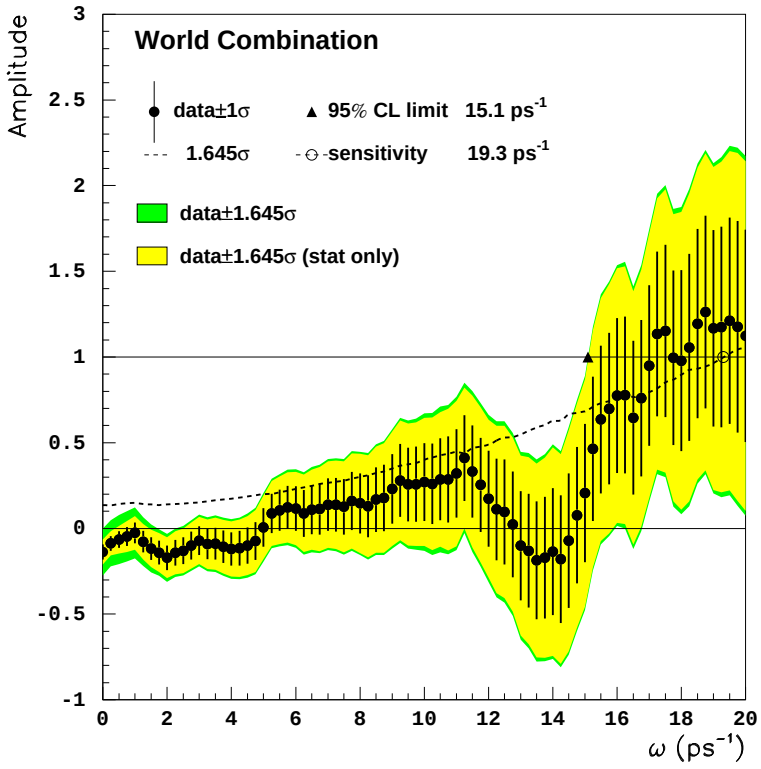

Figure 8.21: Amplitude spectrum for the world combination of B_s oscillation results.

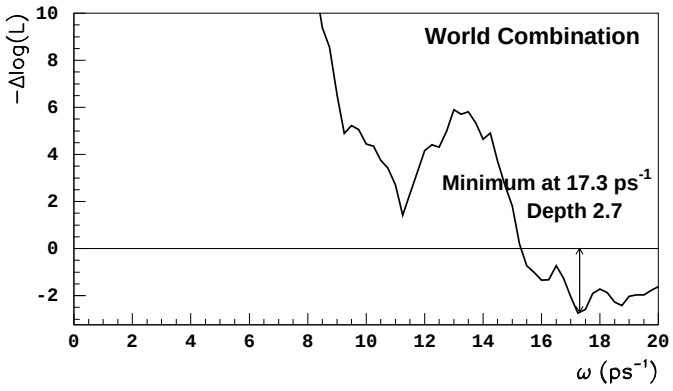

Figure 8.22: Likelihood function for the world combination of B_s oscillations results.

of a fluctuation which compensate, and result in the same estimated probability that what is seen is due to a statistical fluctuation.

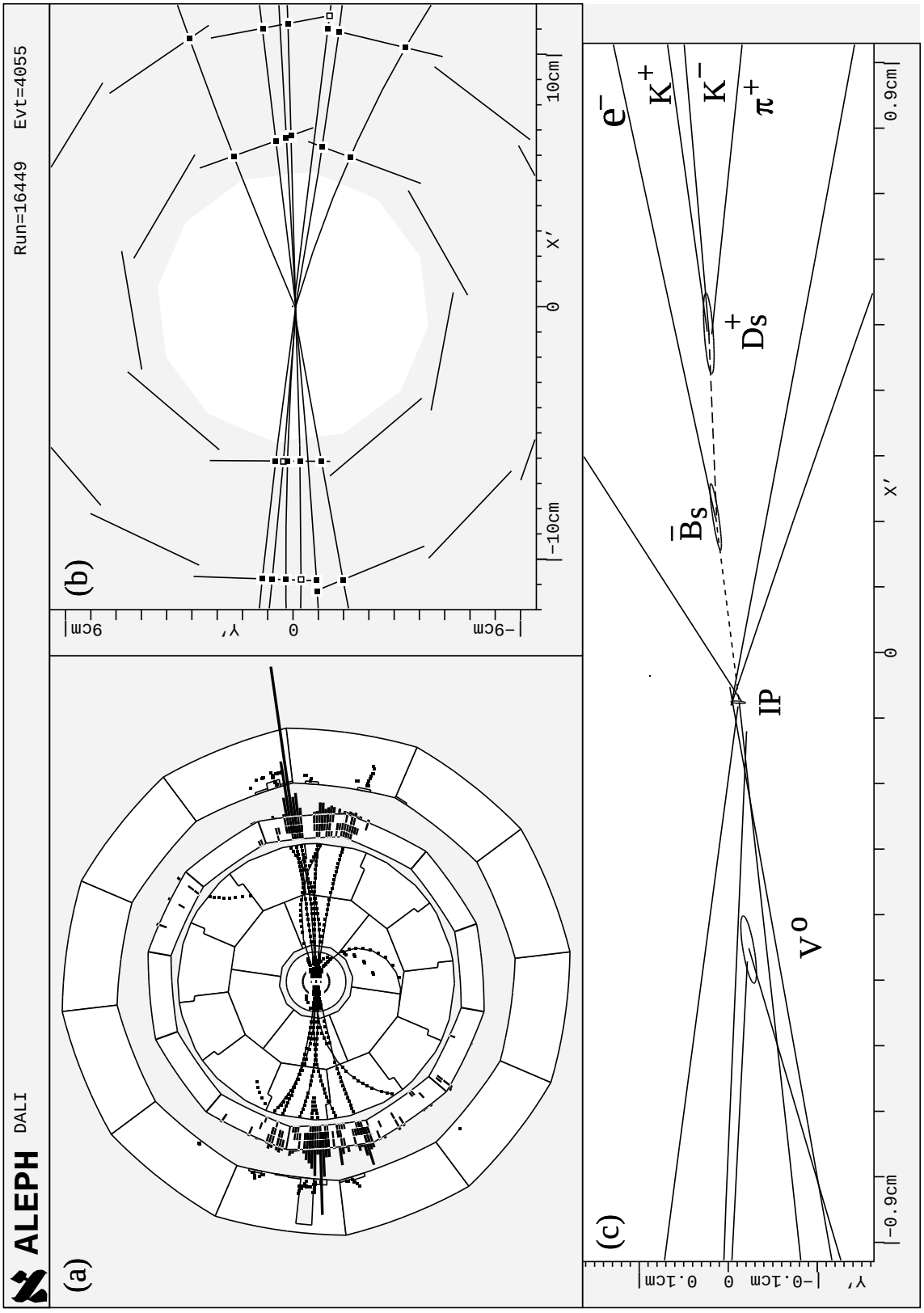

Figure 8.23: Display of a $D_s\,\ell$ candidate.

Chapter 9

Conclusions

A complete picture of the study of B_s oscillations, performed with the ALEPH data taken at energies close to the Z mass between 1991 and 1995, was presented. The emphasis of this thesis was both on the method used for establishing and combining results (the amplitude method) and on a new analysis based on an inclusive semileptonic event sample. The latter provides the single most sensitive B_s oscillation result at LEP. Two supplementary event samples enriched on B_s mesons were described. Finally, the ALEPH- and world-combined results on B_s oscillations were presented.

Some mis-concepts present in the previous literature on the amplitude method were clarified. The expected shape of the amplitude as a function of the test frequency was derived for the first time, also for the case when the oscillation frequency is within the frequency range explored.

The previous ALEPH analysis based on an inclusive semileptonic event sample, which was published in 1998, provided the world's single most sensitive result on B_s oscillations at that time. The analysis presented in this thesis is significantly more sensitive than its predecessor. Most of the increase in sensitivity is due to a new vertexing algorithm especially developed for this analysis and a careful event-by-event treatment of the decay length uncertainties. The use of neural networks for the event selection, the initial and final state flavour determination, and the B_s enrichment also contribute to the increased performance of the analysis. A quantitative estimate of the improvement is provided either by the 95% C.L. sensitivity of both analyses (from $\omega^{95} = 9.6\,\mathrm{ps}^{-1}$ to $\omega^{95} = 11.9\,\mathrm{ps}^{-1}$) or by the amplitude uncertainty for a specific value of the test frequency, $\omega = 20\,\mathrm{ps}^{-1}$ for instance (from ± 3.28 to ± 1.50).

This thesis is probably the last detailed report on B_s oscillations studies with the ALEPH detector at LEP. It is therefore a good opportunity to summarize the evolution of the results on the subject since the first analysis was performed with the ALEPH data back in 1994. Before the LEP experiments started taking data, a complete report on the prospective results was produced [90]. At that time, B mixing was already established. The B_d oscillation frequency measurement was still rather poor, with a 50% relative uncertainty [91]. The B_s oscillation frequency was known to be much larger than that of B_d mesons, but no specific results had ever been produced on the subject. With the assumption of 10 million hadronic Z decays collected at LEP, and the status of the general knowledge on b-hadron physics at that time, it was foreseen that only exclusive reconstruction of hadronic B_s decays into

$D_s^- \pi^+$ would be feasible. No clear statement of a prediction for the final LEP results was attempted, due to the lack of knowledge of many of the crucial input parameters. The experience has shown that the exclusive reconstruction of $B_s \to D_s^- \pi^+$ decays is indeed useful for B_s oscillation studies at LEP, but a much more sensitive, as well as more complicated, analysis can be constructed starting from inclusive semileptonic b-hadron decays.

The time evolution of the lower limit and the sensitivity on Δm_s from the ALEPH experiment since 1994 is presented in Fig. 9.1. The figures shown correspond to the combination of

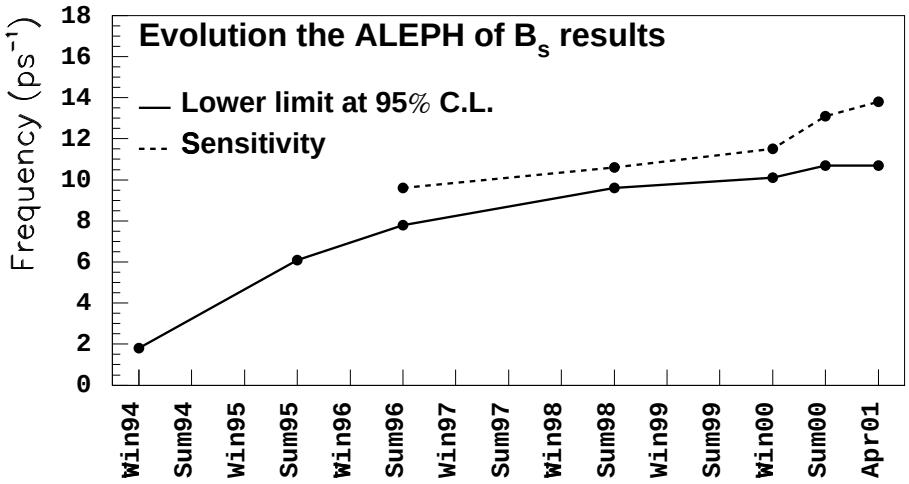

Figure 9.1: Evolution of the ALEPH B_s oscillation results.

the published ALEPH analyses and of those preliminarily released. The step in summer 2000 corresponds to the release of the analysis based on an inclusive semileptonic event selection discussed in this thesis as preliminary ALEPH result for summer 2000 conferences. The last step comes from a recent improvement in the selection of fully reconstructed B_s candidates.

The combined ALEPH sensitivity for B_s oscillations is now $\omega^{95} = 13.8\,\mathrm{ps}^{-1}$, with an uncertainty on the amplitude at $\omega = 20\,\mathrm{ps}^{-1}$ of ± 1.12. A hint of a B_s oscillation signal is observed both in the results from the analysis of the inclusive semileptonic event sample and in the combined ALEPH results (Figs. 8.4, 8.7, 8.18 and 8.19). With the ALEPH combination, a likelihood minimum which corresponds to slightly more than 2.5 standard deviations is found at $\Delta m_s = 17.3\,^{+0.4}_{-0.3}\,\mathrm{ps}^{-1}$ (it was $\Delta m_s = 17.3\,^{+1.3}_{-1.5}\,\mathrm{ps}^{-1}$ for the inclusive semileptonic event selection alone). The uncertainties quoted here, as in Chapter 8, correspond to the values for which $\mathcal{L} = \mathcal{L}_{\min} + 1/2$. They do not provide the one standard deviation statistical uncertainty because the likelihood function is not parabolic in a wide enough range.

The figures obtained with the ALEPH analyses can be compared with those of the other LEP experiments which also performed B_s oscillations analyses: $\omega^{95} = 11.0\,\mathrm{ps}^{-1}$, and $\sigma_{\mathcal{A}}(20\,\mathrm{ps}^{-1}) = \pm 2.3$ for DELPHI, and $\omega^{95} = 8.0\,\mathrm{ps}^{-1}$, and $\sigma_{\mathcal{A}}(20\,\mathrm{ps}^{-1}) = \pm 5.4$ for OPAL (L3 never released any result on B_s oscillations). From these numbers, it is clear that the ALEPH results, and in particular the results from the analysis based on an inclusive semileptonic sample presented in this thesis, dominate the LEP combined B_s oscillations results.

In addition to the LEP experiments, only the SLD Collaboration produced competitive

B_s oscillations results up to now. The SLD combined sensitivity is $\omega^{95} = 13.4\,\mathrm{ps}^{-1}$ and the amplitude uncertainty at high frequency is $\sigma_{\mathcal{A}}(20\,\mathrm{ps}^{-1}) = \pm 0.9$. These results are comparable with those from ALEPH at high frequency.

The amplitude spectrum corresponding to the combination of all results on B_s oscillations available at the moment was presented in Fig. 8.21. The world combined sensitivity is $19.3\,\mathrm{ps}^{-1}$. A lower limit on the B_s oscillation frequency is set at $\Delta m_s > 15.1\,\mathrm{ps}^{-1}$ at 95% C.L. The likelihood function for the world combination has a minimum with a significance of ~ 2.3 standard deviations at $\Delta m_s = 17.3 \pm 0.3\,\mathrm{ps}^{-1}$. However, the significance of the minimum is not quite large enough to claim the observation of an oscillation signal (even though it is larger than expected with the overall sensitivity: $2.1\,\sigma$ significance). New data or new analyses of the existing data would be needed to confirm this result.

In the near future, final results from the inclusive semileptonic event selection and from the $D_s\,\ell$ selection from ALEPH will be available. The DELPHI Collaboration is also expected to update and finalize some of their analyses, and therefore improve their combined sensitivity. The SLD Collaboration will most probably finalize their analyses, although not much improvement is expected. The last word on B_s oscillations from the Z peak data is therefore to be given soon. The present situation may evolve into a deeper minimum of the likelihood, and finally reveal the value of Δm_s.

In one or two years time, the first results from CDF and D0 at Tevatron RunII will most probably be ready. If Δm_s is in the frequency range predicted by the Standard Model (as hinted by the Z peak data), and it has not been measured before, the Tevatron experiments will not miss it.

Bibliography

[1] S.L.Glashow, *Partial-Symmetries of weak interactions*; *Nucl. Phys.* **A22** (1961) 579.
S.Weinberg, *Model of leptons*; *Phys. Rev. Lett.* **19** (1967) 1264.
A.Salam, *Elementary Particle Theory*; *Proc. Nob. Symp.* (1968) 367.

[2] J.Iliopoulos *Progress in Gauge Theories*; Proc. 17th Internat. Conf. on High-Energy Physics, London, 1974 (ed. J.R. Smith).

[3] N.Cabbibo, *Unitarity symmetry and leptonic decays*; *Phys. Rev. Lett.* **10** (1963) 531.
M.Kobayashi and T.Maskawa, *CP-violation in the renormalizable theory of weak interaction*; *Progr. Theor. Phys.* **49** (1973) 652.

[4] I.I.Bigi and A.I.Sanda, *CP Violation*; Cambridge monographs on particle physics, nuclear physics and cosmology. Cambridge University Press 2000.

[5] A.D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the Universe*; *JETP Lett.* **5** (1967) 24.

[6] M.B.Gavela, et al., *Standard Model CP-violation and baryon asymmetry*; *Mod. Phys. Lett.* **A9** (1994) 795.

[7] A.Ali, *Flavour changing neutral current processes in B decays*; *Nucl. Phys.* **B** (Proc. Suppl.) **59** (1997) 86.
J.L. Rosner, *CKM matrix and Standard Model CP violation*; *Nucl. Phys.* **B** (Proc. Suppl.) **59** (1997) 1.

[8] The B Oscillation Working Group. `http://lepbosc.web.cern.ch/LEPBOSC/` .

[9] D.Abbaneo and G.Boix, *The B_s oscillation amplitude analysis*; *Journal of High Energy Physics* **JHEP08** (1999) 004.

[10] The ALEPH Coll., *Search for B_s^0 oscillations using inclusive lepton events*; *Euro. Phys. J.* **C7** (1999) 553.

[11] I.S.Towner and J.C.Hardy, *The current status of V_{ud}*; talk at WEIN98 and `nucl-th/9809087`.

[12] A.García, et al., *Neutron beta decay and the current determination of $|V_{ud}|$*; *Phys. Lett.* **B500** (2001) 66.

[13] The Particle Data Group, *Review of Particle Physics*; *Euro. Phys. J.* **C15** (2000) 1-878.

[14] A.J.Buras and R.Fleischer, *Quark mixing, CP violation and rare decays after the top quark discovery*; `hep-ph/9704376`. Published in Heavy Flavours II, World Scientific (1997).

[15] The ALEPH Coll., *A direct measurement of $|V_{cs}|$ in hadronic W decays using a charm tag*; *Phys. Lett.* **B465** (1999) 349.
The DELPHI Coll., *Measurement of $|V_{cs}|$ using W decays at LEP2*; *Phys. Lett.* **B439** (1998) 209.
The OPAL Coll., *A Measurement of the Rate of Charm Production in W Decays*; *Phys. Lett.* **B490** (2000) 71.

[16] A.Ealet, *WW cross sections and W branching ratios*; talk at ICHEP2000, Osaka, July 2000.

[17] N.Isgur and M.B.Wise, *Weak decays of heavy mesons in the static quark approximation*; *Phys. Lett.* **B232** (1989) 113.
N.Isgur and M.B.Wise, *Weak transition form factors between heavy mesons*; *Phys. Lett.* **B237** (1990) 527.

[18] I.I.Bigi, *Memo on extracting $|V_{cb}|$ and $|V_{ub}/V_{cb}|$ from semileptonic B decays*; `hep-ph/9907270`.

[19] E.Barberio, *$|V_{cb}|$ from $B \to D^* \ell \bar{\nu}_\ell$ and new $B \to D^{**} \ell \bar{\nu}_\ell$ results at LEP*; talk at ICHEP2000, Osaka, Japan.

[20] M.Battaglia, *Determinations of $|V_{ub}|$ with inclusive techniques at LEP*; *Nucl. Phys.* **B** (Proc. Suppl.) **93** (2001) 291.

[21] The CLEO Coll., *Measurement of $B \to \rho \ell \nu$ Decay and $|V_{ub}|$*; *Phys. Rev.* **D61** (2000) 052001.

[22] A.P.Heinson, *Measuring the CKM Matrix Element $|V_{tb}|$ at D0 and CDF*; talk at 2nd Int. Conf. on B Physics and CP Violation, Honolulu, Hawaii, March 1997 `hep-ex/9707026`.
The CDF Coll., *First Measurement of the Ratio $BR(t \to Wb)/BR(t \to Wq)$ and associated limit on the CKM element $|V_{tb}|$*; Submitted to *Phys. Rev. Lett.*, `hep-ex/0012029`.

[23] J.Swain and L.Taylor, *First determination of the quark mixing matrix element $|V_{td}|$ independent of assumptions of unitarity*; *Phys. Rev.* **D58** (1998) 093006.

[24] D.Wyler, *Goals and methods of flavour physics*; 1999 European School of High-Energy Physics.

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

[26] A.J.Buras, M.E.Lautenbacher, and G.Ostermaier, *Waiting for the Top Quark Mass, $K^+ \to \pi^+ \nu \bar{\nu}$, $B_s^0 - \overline{B_s^0}$ Mixing, and CP Asymmetries in B-Decays*; *Phys. Rev.* **D50** (1994) 3433.

[27] M.Gell-Mann and A.Pais, *Behaviour of neutral particles under charge conjugation*; *Phys. Rev.* **97** (1955) 1387.
K.Landé, et al., *Observation of long-lived neutral V particles*; *Phys. Rev.* **103** (1956) 1901.

[28] M.K.Gaillard and B.W.Lee, *Rare decay modes of the K mesons in gauge theories*; *Phys. Rev.* **D10** (1974) 897.

[29] The $\Delta\Gamma_s$ Working Group. `http://lepbosc.web.cern.ch/LEPBOSC/deltagamma_s/` .

[30] A.J.Buras, W.Słominski, and H.Steger, *B-meson decay, CP violation, mixing angles and the top quark mass*; *Nucl. Phys.* **B238** (1984) 529.
A.J.Buras, W.Słominski, and H.Steger, *$B^0 - \overline{B^0}$ mixing, CP violation and the B-meson decay*; *Nucl. Phys.* **B245** (1984) 369.

[31] T.Inami and C.S.Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \overline{K^0}$*; *Prog. Theo. Phys.* **65** (1981) 297.

[32] A.J.Buras, M.Jamin, and P.H.Weisz, *Leading and next-to-leading QCD corrections to ε-parameter and $B^0 - \overline{B^0}$ mixing in the presence of a heavy top quark*; *Nucl. Phys.* **B374** (1990) 491.

[33] E.A. Paschos, B.Stech, and U.Türke, *The charged current couplings and CP-violation in the B-meson system*; *Phys. Lett.* **B128** (1983) 240.

[34] L.Lellouch and C.-J.D.Lin (UKQCD Coll.), *$B^0 - \overline{B^0}$ mixing and decay constants from Lattice QCD*; talk at Heavy Flavours 8, Southampton, UK 1999. `hep-ph/9912322`.

[35] P.Ball, et al., *B decays at the LHC*, in *Standard model physics (and more) at the LHC*; CERN-TH-2000-101, Mar 2000; `hep-ph/0003238`.

[36] An extensive review of the literature about CKM fits is given in:
M.Ciuchini, G.D'Agostini, E.Franco, V.Lubicz, G.Martinelli, F.Parodi, P.Roudeau and A.Stocchi; *2000 CKM Triangle analysis: a critical review with updated experimental inputs and theoretical parameters*; submitted to JHEP; `hep-ph/0012308`.

[37] A.I. Sanda, *CP Violation in B decays*; talk presented in Moriond 2001, Les Arcs, France.

[38] F.Caravaglios, F.Parodi, P.Roudeau, and A.Stocchi, *Determination of the CKM unitarity triangle parameters by end 1999*; `hep-ph/0002171`.

[39] The UA1 Coll., *Search for $B^0 - \overline{B^0}$ oscillations at the CERN proton-antiproton collider*; *Phys. Lett.* **B186** (1987) 247.

[40] The ARGUS Coll., *ARGUS results on $B^0 - \overline{B^0}$ mixing*; talk to the Intern. Symp. on Production and decay of heavy hadrons (Heidelberg, May 1986).

[41] The CLEO Coll., *Limits on $B^0 - \overline{B^0}$ mixing and τ_{B^0}/τ_{B^+}*; *Phys. Rev. Lett.* **58** (1987) 183.

[42] The MARKII Coll., *Upperlimit on $B^0 - \overline{B^0}$ mixing in e^+e^- annihilation at 29 GeV*; *Phys. Lett.* **B160** (1985) 188.

[43] The ARGUS Coll., *Observation of $B^0 - \overline{B^0}$ mixing*; *Phys. Lett.* **B192** (1987) 245.

[44] The CLEO Coll., *$B^0 - \overline{B^0}$ mixing at the $\Upsilon(4S)$*; *Phys. Rev. Lett.* **62** (1989) 2233.

[45] The ALEPH Coll., *An investigation for B_d and B_s oscillation*; *Phys. Lett.* **B322** (1994) 441.

[46] The UA1 Coll., *Measurement of $B^0 - \overline{B^0}$ mixing at the CERN $SP\overline{P}S$ Collider*; *Phys. Lett.* **B262** (1991) 171.
The CDF Coll., *Measurement of $B - \overline{B}$ production correlations, $B^0 - \overline{B^0}$ mixing, and a limit on ϵ_B in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV*; *Phys. Rev.* **D55** (1997) 2546.

[47] The ALEPH Coll., *Measurement of the b forward-backward asymmetry and mixing using high p_T leptons*; *Phys. Lett.* **B384** (1996) 414.
The DELPHI Coll., *Measurement of the semileptonic b branching fractions and average b mixing parameter in Z decays*; Submitted to *Euro. Phys. J.* **C**.
The L3 Coll., *Measurement of the $e^+e^- \to Z \to b\bar{b}$ Forward-Backward Asymmetry and the $B^0 - \overline{B^0}$ Mixing Parameter Using Prompt Leptons*; *Phys. Lett.* **B448** (1999) 152.
The OPAL Coll., *Measurement of Heavy Quark Forward-Backward Asymmetries and Average B Mixing Using Leptons in Multihadronic Events*; *Z. Phys.* **C70** (1996) 357.

[48] The ALEPH Coll., *Precision measurement of the inclusive b lifetime and time-dependent mixing*; contributed paper to EPS-HEP Jerusalem, August 1997.
The ALEPH Coll., *Improved measurement of the $B_d - \overline{B_d}$ oscillation frequency*; *Z. Phys.* **C75** (1997) 397.
The DELPHI Coll., *Measurement of $B_d - \overline{B_d}$ oscillations*; *Z. Phys.* **C76** (1997) 579.
The L3 Coll., *Measurement of the $B_d - \overline{B_d}$ oscillation frequency*; *Euro. Phys. J.* **C5** (1998) 195.
The OPAL Coll., *Measurement of the B^0 Lifetime and Oscillation Frequency using $B^0 \to D^{*+}\ell^-\nu$ decays*; *Phys. Lett.* **B493** (2000) 266.
The SLD Coll., *Measurements of the Time Dependence of $B_d - \overline{B_d}$ Mixing with Kaon and Charge Dipole Tags* ; contributed paper to ICHEP96, Warsaw, July 1996.
The SLD Coll., *Measurement of Time Dependent $B_d - \overline{B_d}$ Mixing using Topology and Charge Selected Semileptonic B Decays*; contributed paper to ICHEP96, Warsaw, July 1996.
The SLD Coll., *Measurement of Time Dependent $B_d - \overline{B_d}$ Mixing using Inclusive Semileptonic Decays*; contributed paper to ICHEP96, Warsaw, July 1996.
The CDF Coll., *Time-dependent B^0 mixing in the $\bar{b} \to \ell^+ X, b \to \overline{B^0} \to D^{+(*)}X$ channel*; Preliminary result, CDF4526.
The CDF Coll., *Measurement of $B_d - \overline{B_d}$ oscillations frequency using dimuon data in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV*; *Phys. Rev.* **D60** (1999) 051101.

[49] The CLEO Coll., *Precise Measurement of $B^0 - \overline{B^0}$ Mixing Parameters at the $\Upsilon(4S)$*; *Phys. Lett.* **B490** (2000) 36.

[50] G.Boix, *B_s Physics at LEP, SLD, and CDF*; talk presented in BCP4, February 2001 Ise-Shima, Japan. To be published by World Scientific.

[51] W. Taylor, *Run II Beauty Physics at CDF and D0*; talk presented in BCP4, February 2001 Ise-Shima, Japan. To be published by World Scientific.

[52] I.Dunietz, *$B_s - \overline{B_s}$ mixing, CP violation, and extraction of CKM phases from untagged B_s data samples*; *Phys. Rev.* **D52** (1995) 3048.

[53] M.Gronau and D.London, *New physics in CP asymmetries and rare B decays*; *Phys. Rev.* **D55** (1997) 2845. And references there in.

[54] J.Papavassiliou and A.Santamaria, *Extra dimensions at the one loop level: $Z \to b\bar{b}$ and $B - \overline{B}$ mixing*; *Phys. Rev.* **D63** (2001) 016002.

[55] R.Barbieri, et al., *Alternative theories of CP violation*; *Phys. Lett.* **B425** (1998) 119.

[56] H.-G. Moser and A. Roussarie, *Mathematical Methods for $B^0\overline{B^0}$ Oscillation Analyses*; *Nucl. Instrum. and Methods* **A384** (1997) 491.

[57] The DELPHI Coll., *Study of $B_s^0 - \overline{B_s^0}$ oscillations and B_s^0 lifetimes using hadronic decays of B_s^0 mesons*; *Euro. Phys. J.* **C18** (2000) 229.

[58] The ALEPH Coll., *Study of $B_s^0 - \overline{B_s^0}$ oscillations using fully reconstructed B_s^0 and $D_s - \ell$ events*; CONF-2000-024. *Euro. Phys. J.* **C16** (2000) 555.

[59] The DELPHI Coll., *Measurement of the B_s^0 Lifetime and Study of $B_s^0 - \overline{B_s^0}$ Oscillations using $D_s - \ell$ Events*; *Euro. Phys. J.* **C16** (2000) 555.

[60] The SLD Coll., *Time dependent $B_s^0 - \overline{B_s^0}$ mixing using inclusive and semileptonic B decays at SLD*; contributed paper to ICHEP2000, Osaka, Japan.

[61] The ALEPH Coll., *Study of B_s^0 oscillations and lifetime using fully reconstructed D_s^- decays*; *Euro. Phys. J.* **C4** (1998) 367.

[62] *LEP Design Report*; CERN-LEP/84-01, Vol. I and II, 1984.

[63] T.Suzuki, *General Formulas of Luminosity for Various types of Colliding Beam Machines*; KEK-76-3, Jul. 1976.

[64] The ALEPH Coll., ALEPH HANDBOOK 1995; Vol. 1 Edited by Chris Bowdery, CERN 1995.

[65] The ALEPH Coll., ALEPH HANDBOOK 1995; Vol. 2, Edited by Chris Bowdery, CERN 1997.

[66] The ALEPH Coll., *Performance of the ALEPH detector at LEP*; *Nucl. Instrum. and Methods* **A360** (1995) 481.

[67] The ALEPH Coll., *A study of the decay width difference in the $B_s^0 - \overline{B_s^0}$ system using $\phi\phi$ correlations*; *Phys. Lett.* **B486** (2000) 286.

[68] T. Boccali, *Combination of the $\mathrm{d}E/\mathrm{d}x$ information from TPC pads and wires*; Internal note ALEPH 99-033.

[69] The ALEPH Coll., *Measurement of the Z Resonance Parameters at LEP*; *Euro. Phys. J.* **C14** (2000) 1.

[70] D.Bédérère, et al., *SiCAL, a high precision silicon-tungsten luminosity calorimeter for ALEPH*; *Nucl. Instrum. and Methods* **A365** (1995) 117.

[71] JADE Coll., W.Bartel, et al., *Experimental studies on multijet production in e^+e^- anihilations at PETRA energies*; *Z. Phys.* **C33** (1986) 23.

[72] The ALEPH Coll., *Measurement of the mean lifetime and branching fractions of b-hadrons decaying to J/ψ*; *Phys. Lett.* **B295** (1992) 396.

[73] The OPAL Coll., *QCD Studies Using a Cone-based Jet Finding Algorithm for e^+e^- Collisions at LEP*; *Z. Phys.* **C63** (1994) 197.

[74] The ALEPH Coll., *Measurement of the Forward-Backward Asymmetries in $Z \to b\bar{b}$ and $Z \to c\bar{c}$ Decays with Leptons*; contributed paper to HEP99, Tampere, Finland.

[75] E. Longo and I. Sestili, *Monte Carlo calculation of photon initiated electromagnetic showers in lead glass*; *Nucl. Instrum. and Methods* **128** (1987) 283.

[76] D.Brown, *QFNDIP, a primary vertex finder*; Internal note ALEPH 92-47.

[77] T.S.Mattison, *QVSRCH, a tool for inclusive secondary vertex finding*; Internal note ALEPH 92-173.

[78] T.Oest, *Inclusive Vertexing*; Internal note ALEPH 95-125.

[79] D.Brown and M.Frank, *Tagging b-hadrons using track impact parameters*; Internal note ALEPH 92-135.

[80] P. Gay, et al., *Tagging Higgs Boson in hadronic LEP2 events with neural networks*; talk presented in Engineering and artificial Intelligence for High Energy and Nuclear Physics Pisa (Tuscany), Italy, 1995. More information can be found on the web page `http://alephwww.cern.ch/~gaypas/NN/Welcome.html` .

[81] CN/ASD, *GEANT detector description and simulation tool*; CERN, 1994.

[82] T. Sjöstrand and M. Bengtsson, *The Lund Monte Carlo for jet fragmentation and e^+e^- physics -JETSET version 6.3- and update*; *Comput. Phys. Commun.* **43** (1987) 367.

[83] The ALEPH Coll., *Heavy quark tagging with leptons in the ALEPH detector*; *Nucl. Instrum. and Methods* **A346** (1994) 461.

[84] The ALEPH Coll., *Measurement of the B_s^0 lifetime*; *Phys. Lett.* **B322** (1994) 275.

[85] The ALEPH Coll., *Study of the CP asymmetry of $B^0 \to \psi K_S^0$ decays in ALEPH*; *Phys. Lett.* **B492** (2000) 259.

[86] Averages used for the Winter 2000 conference, see `http://lepewwg.web.cern.ch/LEPEWWG/heavy/` .

[87] The SLD Coll., *Time Dependent $B_s^0 - \overline{B_s^0}$ Mixing Using Exclusively Reconstructed D_s^+ Decays at SLD*; contributed paper to ICHEP2000, Osaka, Japan.

[88] The DELPHI Coll., *Search for $B_s^0 - \overline{B_s^0}$ oscillations in DELPHI*; DELPHI-CONF 403, contributed paper to ICHEP2000, Osaka, Japan.

[89] The OPAL Coll., *A Study of B_s^0 Meson Oscillation Using Hadronic Z^0 Decays Containing Leptons*; *Euro. Phys. J.* **C1** (1999) 587.
The OPAL Coll., *A Study of $B_s - \overline{B_s}$ Meson Oscillation using D_s Lepton Correlations*; CERN-EP-2000-136, accepted by *Euro. Phys. J* **C**.

[90] G.Altarelli, et al., *Z physics at LEP 1, Vol.1, Standard physics*; CERN 89-08.

[91] The Particle Data Group, *Review of particle properties*; *Phys. Lett.* **B204** (1988) 218.