# Decay Parameters of the $\mathrm{B}_s^0$ Meson in the Inclusive $\phi\ell^+$ Channel

**Ole Myren Røhne**


Thesis submitted in partial fulfillment
of the requirements for the degree Doctor Scientiarum
at the Department of Physics, University of Oslo

August 1998

SDG

# Abstract

In a sample of 2470K hadronic $Z^0$ decays registered by DELPHI during the years 1994 and 1995, an estimated signal of $127.9 \pm 9.7$ inclusive $B_s^0 \to \phi \ell^+$ decays are reconstructed. The $B_s^0$ lifetime is measured, $\tau_{B_s^0} = \left(1.40^{+0.31}_{-0.27}\ (stat) \pm 0.08\ (syst)\right)$ ps, and exclusion regions for the time dependent $B_s^0 \overline{B}_s^0$ mixing frequency is found:

$$\Delta m_s \notin \langle 0.9\ \text{ps}^{-1} \ldots 1.8\ \text{ps}^{-1} \rangle \cup \langle 2.2\ \text{ps}^{-1} \ldots 3.3\ \text{ps}^{-1} \rangle \cup \langle 8.0\ \text{ps}^{-1} \ldots 9.0\ \text{ps}^{-1} \rangle \ (95\%\ \text{C. L.}).$$

The sensitivity of the analysis is $\Delta m_s < 1.5\ \text{ps}^{-1}$, defined as the point where the exclusion probability drop below 50%.

# Acknowledgements

I am greatly indebted to all my collaborators in the DELPHI B experiment, particularly from the $B$ lifetimes/oscillation team. A few people deserve special mentioning: Achille Stocchi and Fabrizio Parodi for showing a personal interest in my work, providing countless suggestions for improvements, and Ph.D-student Anders Borgland for feeding me information from team meetings that I missed, in the spirit of friendly competition.

My Ph.D studies have been supported by the Department of Physics at the University of Oslo, through a combined research/teaching assistantship. A special grant from the Norwegian Research Council allowed me to spend the year 1996 at CERN, Geneva, working closer to the DELPHI experimental facilities.

My thesis supervisor at the University of Oslo has been professor Torleiv Buran. He introduced me to the physics of neutral $b$ mesons and suggested the study of $\mathrm{B}_s^0\overline{\mathrm{B}}_s^0$ mixing as the topic for my thesis work. His passion for $b$ physics and the discrete symmetries of Nature has been a source of inspiration. Many thanks also for reading and commenting the early thesis manuscript, giving a lot of constructive criticism.

Three more members of the University of Oslo Elementary Particle Physics group have been particularly important to my work: Ph.D-student Trond Myklebust has suffered from sharing his office with me for three years, still he has always been willing to help solving numerous computing problems. Professors Lars Bugge and Alex L. Read have created an inspiring atmosphere of open discussions and creative problem-solving. Special thanks to Alex for reading and commenting this manuscript.

Finally, I would like to thank my wife Anne Grethe. Without her continued support and encouragement, my thesis work would have remained uncompleted.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction and thesis outline

Elementary Particle Physics up to an energy scale of about 100 GeV is well understood as a quantum field theory known as the Standard Model of Elementary Particles in which matter, represented as fermion fields, interact through gauge bosons. Fermions are spin-$\frac{1}{2}$ particles obeying Fermi-statistics, the quarks and leptons, see figure 1.1. The gauge bosons are spin-1 particles having Bose-Einstein-statistics, see figure 1.2, that are associated with the generators of internal gauge symmetries. The unification of the electromagnetic and weak nuclear interactions into a single electro-weak interaction described by the $SU(2)_{\mathrm{W}} \otimes U(1)_{\mathrm{Y}}$ symmetry is one of the major triumphs of the Standard Model of Elementary Particles ([1], [2], [3]).

| | | | |
|---|---|---|---|
| $\frac{1}{2}$ 0 <br> $\nu_e$ <br> $e$-neutrino <br> $< 15 \cdot 10^{-6}$ | $\frac{1}{2}$ $-1$ <br> $e$ <br> electron <br> 0.000511 | $\frac{1}{2}$ $-\frac{1}{3}$ <br> $d$ <br> down <br> 0.005–0.015 | $\frac{1}{2}$ $+\frac{2}{3}$ <br> $u$ <br> up <br> 0.002–0.008 |
| $\frac{1}{2}$ 0 <br> $\nu_\mu$ <br> $\mu$-neutrino <br> $< 0.00017$ | $\frac{1}{2}$ $-1$ <br> $\mu$ <br> muon <br> 0.10566 | $\frac{1}{2}$ $-\frac{1}{3}$ <br> $s$ <br> strange <br> 0.1–0.3 | $\frac{1}{2}$ $+\frac{2}{3}$ <br> $c$ <br> charm <br> 1.0–1.6 |
| $\frac{1}{2}$ 0 <br> $\nu_\tau$ <br> $\tau$-neutrino <br> $< 0.024$ | $\frac{1}{2}$ $-1$ <br> $\tau$ <br> tau <br> 1.777 | $\frac{1}{2}$ $-\frac{1}{3}$ <br> $b$ <br> bottom <br> 4.1–4.5 | $\frac{1}{2}$ $+\frac{2}{3}$ <br> $t$ <br> top <br> $180 \pm 12$ |

Legend: spin charge / $S_{\mathrm{ymbol}}$ / name / mass ($\frac{GeV}{c^2}$)

Figure 1.1: *Periodic Table of the Fermions. Any similarity to the Periodic Table of the Elements is purely intentional: Particles listed in one column share physical properties while particle masses within one row increase (with one notable exception) from left to right.*

The photon, $\gamma^0$ — mediator of the electromagnetic interactions, is massless as required by local gauge invariance. On the other hand, the $W\pm$ and $Z^0$ bosons that mediate the weak nuclear interactions are observed [4] to have masses of $(80.41 \pm 0.10)$ GeV and $(91.187 \pm 0.007)$ GeV,

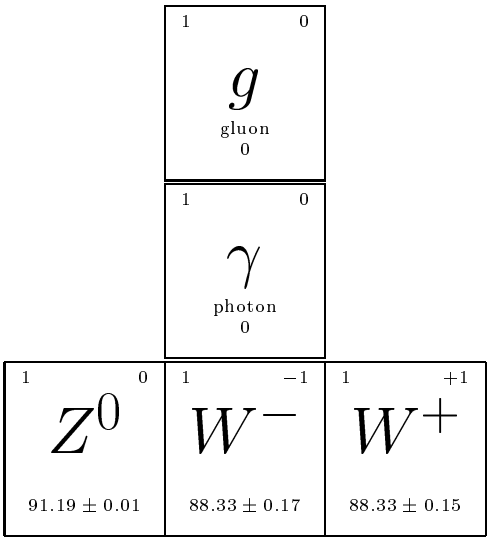


Figure 1.2: *Gauge bosons of elementary particle physics. The gluon mediate the strong nuclear interaction associated with the $SU(3)_{\mathrm{C}}$ gauge symmetry. The photon (mediating the electromagnetic interaction) and the $Z^0$ and $W^{\pm}$ (responsible for the weak nuclear interaction) are all remnants of the spontaneously broken $SU(2)_{\mathrm{W}} \otimes U(1)_{\mathrm{Y}}$ gauge symmetry*

respectively [5]. As massive gauge bosons are indeed incompatible with a naive local gauge symmetry, one concludes that the underlying $SU(2)_{\mathrm{W}} \otimes U(1)_{\mathrm{Y}}$ symmetry must be broken. The Higgs mechanism introduces *spontaneous* symmetry breaking by means of a complex scalar $SU(2)_{\mathrm{W}}$ doublet field, the Higgs field, which acquires a non-zero vacuum expectation value. In this process, $W^{\pm}$ and $Z^0$ become massive while $\gamma^0$ remains massless. This happens at the cost of only two arbitrary parameters: the vacuum expectation value of the Higgs field and the ratio of the $W^{\pm}$ and the $Z^0$ masses, and the theory retains the appealing properties of a local gauge theory. As an added bonus, it is possible to have the same spontaneous symmetry breaking generate mass terms for the leptons and quarks. The Yukawa couplings between the Higgs field and the fermions are however completely arbitrary. With 3 fermion generations, this arbitrariness gives 13 or 20 new parameters, depending on whether the neutrinos are massive or not.

Focusing on the quark sector, diagonalization of the up- and down-type mass matrices yields 6 arbitrary quark masses and 4 parameters (3 rotation angles and 1 phase) describing how up- and down-type quarks mix in the charged current coupling to the $W^{\pm}$. In general, with $n_g$ generations, the quark mixing is described by an unconstrained complex $n_g \times n_g$ matrix, commonly named after Cabibbo [6], Kobayashi and Maskawa [7]. In the Standard Model, the CKM-matrix should be unitary, and testing this unitarity is thus a test of the Standard Model itself. One part of the CKM-matrix that is particularly poorly known is the couplings of charged currents involving a 3rd generation quark with the $W^{\pm}$, as described by the 3rd row and column. These parameters are accessible through the study of $b$-hadron physics: The couplings of the $b$ quark itself to the $c$- and $u$ quark are directly measurable in $b$-decays; the couplings of the $t$ quark are accessible because virtual $t$ quarks turn up as quantum corrections in $\Delta B = 2$ ($b \leftrightarrow \overline{b}$) transitions.

The coupling $V_{tq}$ ($q = d, s$) between the $W^{\pm}$ and the $tq$ charged current determines the size of the $\mathrm{B}^0_q$ mass splitting $\Delta m_q$, which is experimentally manifest in the $\mathrm{B}^0_q\overline{\mathrm{B}}^0_q$ mixing. As $\Delta m_d$, the $\mathrm{B}^0_d$ mass splitting, is already well measured (see references in section 2.3), experimental interest turns to $\Delta m_s$. The motivation for measuring $\Delta m_s$ is partially to extract $V_{ts}$ directly, partially to extract $\left|\frac{V_{td}}{V_{ts}}\right|^2 \sim \frac{\Delta m_d}{\Delta m_s}$ with minimal hadronic uncertainties.

The main object of this thesis is the study of time dependent charge correlations in $\mathrm{B}^0_s$ decays, searching for time dependent $\mathrm{B}^0_s\overline{\mathrm{B}}^0_s$ mixing with oscillation frequency $\Delta m_s$. From measurements

of integrated mixing, the $B_s^0$ mass splitting is known to be sizeable, and several existing analyses report lower limits on $\Delta m_s$; see section section 4.1 for references. However, the region of the $\Delta m_s$ parameter open to exploration is currently limited by experimental statistics. Combining as many decay channels as possible from all available experiments is necessary to get the best obtainable limit, and possibly discover time dependent $B_s^0\overline{B}_s^0$ mixing. The ambition of the present analysis is to contribute to the available statistics of $B_s^0$ decays through the inclusion of another decay channel: $B_s^0$ mesons partially reconstructed in the inclusive $\phi\ell^+$ final state. The work is based on data collected by the DELPHI experiment at LEP, CERN during the 1994 and 1995 runs, representing the larger part of the hadronic $Z^0$ statistics, and registered with fully operational particle identification and three dimensional vertexing capabilities. As the study of time dependent mixing already involves decay time reconstruction, only a little additional work is required to measure the $B_s^0$ lifetime and to study the possible decay width splitting of the $B_s^0$ states.

The present thesis is outlined as follows: Chapter 2 tries to bridge the gap between fundamental parameters of the Standard Model and certain observables in the neutral B meson system. Chapter 3 gives a brief overview of the experimental setup: the DELPHI detector at LEP. The selection of $B_s^0$ events decaying into the $\phi\ell^+$ final state is treated in chapter 4 and the extraction of physical parameters from decay observations is described in chapter 5. Finally, chapter 6 concludes and compares the results to analyses of complementary decay channels.

# Chapter 2

# From Standard Model parameters to $b$-physics observables

The focus of this chapter is not the Standard Model as such, but the Cabibbo-Kobayashi-Maskawa matrix and how its parameters relate to experimental observables in neutral B mesons.

## 2.1 The origin of the Cabibbo-Kobayashi-Maskawa matrix in the Standard Model

In the Standard Model [1] [2] [3], the fermions acquire mass through the coupling to some Higgs field. The important constraint when specifying this interaction is that the Lagrangian is invariant under the gauge group. In the minimal model, the Higgs $SU(2)$ doublet $\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ couples the left-handed quark doublets to the up-type right-handed quarks while its charge conjugate field $\begin{pmatrix} \phi^{0*} \\ \phi^- \end{pmatrix} = i\sigma_2 \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}^*$ couples the left-handed quark doublets to the down-type quarks in a similar fashion:

$$\mathcal{L}(q,H) = Y_{ij}\left(\overline{u}_{iL}, \overline{d}_{iL}\right)\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} u_{jR} + Y'_{ij}\left(\overline{u}_{iL}, \overline{d}_{iL}\right)\begin{pmatrix} \phi^{0*} \\ \phi^- \end{pmatrix} d_{jR} + \text{h.c.} \tag{2.1}$$

Within the theory, the Yukawa-couplings $Y_{ij}$ and $Y_{ij}$ are completely arbitrary complex matrices of dimension $n_g$, the number of fermion generations.

In the process of spontaneous symmetry breaking process, the Higgs field acquires a finite vacuum expectation value. Three degrees of freedom are "eaten" by the gauge bosons as they become massive, while one real component remains as the Higgs particle. Defining $\langle\phi^0\rangle \equiv v$ and changing variables $\phi^0 \to v + \phi^0$, the Higgs doublet is then written as $\begin{pmatrix} 0 \\ v+\phi^0 \end{pmatrix}$. Inserting this in equation 2.1, quark mass terms are generated and quark couplings to the Higgs particle are seen to be proportional to the quark masses:

$$M_{ij}\overline{u}_{iL}u_{jR}(1 + \frac{1}{v}\phi^0) + M'_{ij}\overline{d}_{iL}d_{jR}(1 + \frac{1}{v}\phi^0) + \text{h.c.} \tag{2.2}$$

As $\mathbf{M} \equiv v\mathbf{Y}$ and $\mathbf{M'} \equiv v\mathbf{Y'}$ are general complex matrices, two sets of two unitary matrices are needed to diagonalize them:

$$\mathbf{M} \;=\; \mathbf{U_L}^\dagger \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \mathbf{U_R} \tag{2.3}$$

$$\mathbf{M}' = \mathbf{U}_{\mathbf{L}}'^{\dagger} \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \mathbf{U}_{\mathbf{R}}' \tag{2.4}$$

Hence, the original quark fields do not represent the massive or "physical" quarks. Changing basis to mass-eigenstates means that the hadronic charged current is no longer flavor-diagonal:

$$J_\mu^+ = \overline{u}_{iL}\gamma_\mu d_{iL} = \overline{u'}_{kL}\gamma_\mu d'_{jL}\mathbf{U}_{\mathbf{L}ki}\mathbf{U}_{\mathbf{L}ij}'^{\dagger}. \tag{2.5}$$

In the context of the Standard Model, this defines the Cabibbo-Kobayashi-Maskawa matrix [6] [7] $\mathbf{V}_{\mathrm{CKM}} \equiv \mathbf{U}_{\mathbf{L}}\mathbf{U}_{\mathbf{L}}'^{\dagger}$. On the other hand, it is possible to give a purely phenomenological definition of the CKM-matrix as the contributions of different up- and down-type quark-combinations to the charged hadronic current. Within the Standard Model $\mathbf{V}_{\mathrm{CKM}}$ is necessarily unitary because $\mathbf{U}_{\mathbf{L}}$ and $\mathbf{U}_{\mathbf{L}}'$ are unitary.

## 2.2 Parameterizations and unitarity conditions on $\mathbf{V}_{\mathrm{CKM}}$

In general, a $n_g \times n_g$ unitary matrix is described by $n_g^2$ parameters; $n_g \cdot (n_g-1)/2$ $O(3)$ rotation angles and $n_g \cdot (n_g+1)/2$ phases. However, $2n_g - 1$ phases are unobservable because they can be absorbed in a suitable re-phasing of the quark fields, and the number of observable phases is $(n_g-1)(n_g-2)/2$. Having observable phases is a requirement to have CP-violation, consequently there must exist at least $n_g = 3$ generations in order to explain CP-violation in the Standard Model.

For some unknown reason, the CKM-matrix has an interesting hierarchical structure: The elements that describe couplings within one generation are of order unity, the 1st to 2nd generation couplings are almost an order of magnitude smaller, the 2nd to 3rd generation elements are yet another order of magnitude smaller, and finally the 1st to 3rd generation couplings are once again an order of magnitude smaller:

$$\mathbf{V}_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

Various [8] [9] numerological schemes for the Yukawa-couplings have been proposed to explain this pattern, giving one motivation among others to measure each component as accurately as possible.

Defining $\theta_{ij}$ as the angle of rotation in the $i-j$-plane and writing $s_{ij} \equiv \sin\theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$ the particular parameterization advocated by [4] is recognized as a product of three familiar rotation matrices and a phase matrix together with its conjugate:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} c_{12}c_{13} & c_{13}s_{12} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{13}c_{23} \end{pmatrix}. \tag{2.6}$$

Numerous other parameterizations are of course also possible. For original work, see [10], [11] and further references in [4]. Series-expanding equation 2.6 using the rotation angles as parameters and keeping the dominant terms, a popular approximation of the CKM-matrix due to Wolfenstein [12] is obtained:

$$\mathbf{V}_{\mathrm{CKM}} \simeq \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{2.7}$$

The coupling hierarchy mentioned above can be expressed as the degree of $\lambda \simeq \sin\theta_c$, the parameters $A$ and $|\rho + i\eta|$ are of order unity and the argument of $\rho + i\eta$ is related to the CP-violating phase.

The unitarity-conditions on the CKM-matrix can be written down taking the scalar product of one row or column with the complex conjugate of another:

$$V_{ij}^* V_{jk} = \delta_{ik}. \tag{2.8}$$

Of particular interest is the combination of the 1st ($d$ quark) with the 3rd ($b$ quark) column. This condition involves the smallest and most poorly known matrix elements:

$$V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0. \tag{2.9}$$

Normalizing the second term to one and invoking the Wolfenstein-parameterization, this is written:

$$(\rho + i\eta) + (-1) + (1 - \rho - i\eta) = 0.$$

This 3-term complex equation can be visualized as a triangle in the $\rho$–$i\eta$ plane, see figure 2.1.

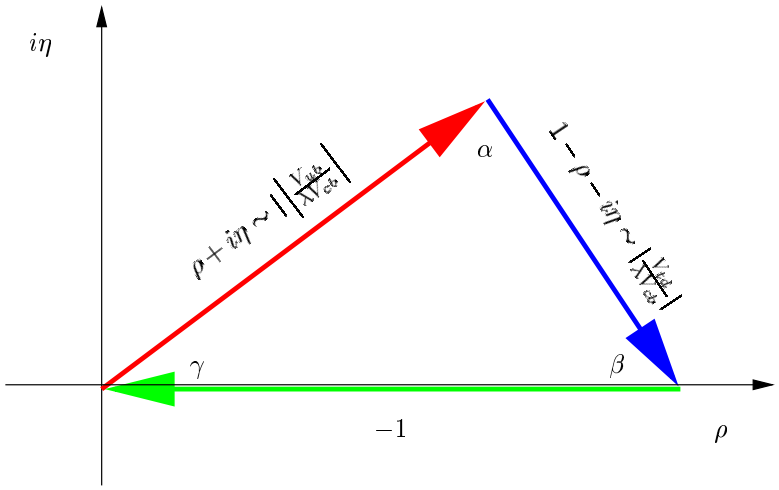


Figure 2.1: *The Wolfenstein parameterization and the $\rho$–$i\eta$ plane unitarity triangle. The Standard Model predicts that the CKM-matrix is unitary, resulting in a closed triangle. The side lengths and angles are individually measurable, and over-constraining the triangle provides a test of the Standard Model.*

The angles $\alpha$, $\beta$ and $\gamma$ of the $\rho$–$i\eta$ unitarity triangle are expected to be accessible through measurements of CP-violating amplitudes measured at the upcoming $b$-physics experiments [13] at DESY [14], KEK [15] and SLAC [16].

The $|\rho + i\eta|$-side is essentially $\left|\frac{V_{ub}}{\lambda V_{cb}}\right|$ and is measured in charm-less $b$-decays at ARGUS and CLEO, see [17].

The final triangle side $|1 - \rho - i\eta| = |\frac{V_{td}}{\lambda V_{cb}}|$ is accessible through the 2nd order weak interactions with virtual $t$ quarks as found in $B^0$–$\overline{B}^0$ mixing.

## 2.3 Phenomenology of $B^0$–$\overline{B}^0$ mixing

The description of $B^0$–$\overline{B}^0$ mixing is formally identical to the classical description of the $K^0$–$\overline{K}^0$ system [18] [19]. However, when it comes to observable quantities the experimental features of

the two systems differ grossly. Recall that $K_{\rm S}$ and $K_{\rm L}$ have dramatically different lifetimes and that the hitherto only observations of CP non-conservation have been done in the decay of neutral kaons [20]. In the $\rm B^0$ meson system, the lifetime difference between the two mass-eigenstates is yet to be observed, and the task of measuring CP-violation in the $b$ quark sector is left to future specialized experiments. On the other hand, the $\rm B^0$ meson mass difference is large compared to the total width, and the corresponding mixing, that is the $\Delta B = 2$ transitions, has been observed.

A general, time dependent, neutral $B$ meson state can be written in terms of the flavor eigenstates $|\rm B^0\rangle$ $(\overline{b}q)$ and $|\overline{\rm B}^0\rangle$ $(b\overline{q})$:

$$|\mathrm{B}(t)\rangle = B^0(t)|\mathrm{B}^0\rangle + \overline{B}^0(t)|\overline{\mathrm{B}}^0\rangle.$$

In this subspace, the Hamiltonian operator $\mathbf{H}$ takes the form of a $2 \times 2$ non-Hermitian matrix. Separating the Hermitian and anti-Hermitian parts, $\mathbf{M} = \frac{1}{2}\left(\mathbf{H} + \mathbf{H}^\dagger\right)$ and $\mathbf{\Gamma} = i\left(\mathbf{H} - \mathbf{H}^\dagger\right)$, the Scrödinger equation reads:

$$i\frac{\partial}{\partial t}\begin{pmatrix} B^0(t) \\ \overline{B}^0(t) \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} B^0(t) \\ \overline{B}^0(t) \end{pmatrix}. \tag{2.10}$$

The optical theorem relates the anti-Hermitian (dispersive) part of the transition matrix to the total width for decay into final states containing real (on-shell) particles. $\mathbf{\Gamma}$ thus represents forward scattering through on-shell intermediate states. Consequently, $\mathbf{M}$ describes scattering through virtual intermediate states.

From CPT-invariance, the on-diagonal elements satisfy $m \equiv M_{11} = M_{22}$ and $\Gamma \equiv \Gamma_{11} = \Gamma_{22}$. By definition, the off-diagonal matrix elements fulfill $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. CP-invariance would imply $M_{21} = M_{12}$ (real) and $\Gamma_{21} = \Gamma_{12}$ (real), because the flavor eigenstates transform into each other under CP: $\mathrm{CP}|\mathrm{B}^0\rangle = |\overline{\mathrm{B}}^0\rangle$ and $\mathrm{CP}|\overline{\mathrm{B}}^0\rangle = |\mathrm{B}^0\rangle$.

Defining the parameters

$$\begin{aligned}
\Delta m &= +2 \cdot \mathrm{Re}\sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \\
\Delta\Gamma &= -4 \cdot \mathrm{Im}\sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \\
\rho &= \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}},
\end{aligned}$$

the Hamiltonian can be diagonalized and the following eigenvalues and eigenvectors are found:

$$\begin{gathered}
m \pm \frac{1}{2}\Delta m - \frac{i}{2}\left(\Gamma \pm \frac{1}{2}\Delta\Gamma\right) \\
\frac{1}{\sqrt{1+|\rho|^2}}\left(|\mathrm{B}^0\rangle \mp \rho|\overline{\mathrm{B}}^0\rangle\right).
\end{gathered}$$

Ignoring the effect of CP-violation is equivalent to setting $\rho \equiv 1$. Doing so, the mass-eigenstates above will be CP-eigenstates as well, with eigenvalues $+1$ and $-1$.

The time-dependent probabilities of no mixing and mixing are found to be

$$P^{\mathrm{B}^0\to\mathrm{B}^0}(t) = \frac{1}{4}\Gamma\left(1-\left(\frac{\Delta\Gamma}{2\Gamma}\right)^2\right)e^{-\Gamma t}\left(e^{+\frac{1}{2}\Delta\Gamma t} + e^{-\frac{1}{2}\Delta\Gamma t} + 2\cos\Delta m t\right) \tag{2.11}$$

$$P^{\mathrm{B}^0\to\overline{\mathrm{B}}^0}(t) = \frac{1}{4}\Gamma\left(1-\left(\frac{\Delta\Gamma}{2\Gamma}\right)^2\right)e^{-\Gamma t}\left(e^{+\frac{1}{2}\Delta\Gamma t} + e^{-\frac{1}{2}\Delta\Gamma t} - 2\cos\Delta m t\right). \tag{2.12}$$

The total (or time integrated) mixing is quantified by the parameter $\chi$, defined as the probability of a $\Delta B = 2$ transition taking place:

$$\chi = \frac{N(\mathrm{B}\to\overline{\mathrm{B}})}{N(\mathrm{B}\to\overline{\mathrm{B}}) + N(\mathrm{B}\to\mathrm{B})}.$$

Integrating the time dependent probabilities, the total mixing probability can be related to the oscillation frequency $x = \frac{\Delta m}{\Gamma}$ and the width difference:

$$\chi = \frac{\left(\frac{\Delta\Gamma}{2\Gamma}\right)^2 + x^2}{2(1+x^2)},$$

which is frequently quoted omitting the $\Delta\Gamma$-dependence [21].

In the $B^0_d$-system, mixing is measured in time integrated measurements at the $\Upsilon(4S)$ resonance [22] [23] as well as in time dependent measurements at $e^+e^- \to Z^0 \to b\overline{b}$ [24] [25] [26] [27] [28] and in $p\overline{p}$-collisions [29]. The combined world average for the integrated mixing parameter using both kinds of measurements is [5]:

$$\chi_d = 0.172 \pm 0.010.$$

The mixing in the $B^0_s$-system mixing $x_s$ is expected to be larger than the $B^0_d$-mixing, simply because the $t$ quark couples stronger to the $s$ quark than to the $d$ quark. Combining available experimental information on CKM-matrix parameters with the Standard Model unitarity constraint, $x_s$ is expected to be of the order 15 [30]. The integrated mixing $\chi_d$ measured on a pure $B^0_d$-sample can be compared to the average mixing $\chi_B = f_d\chi_d + f_s\chi_s$ seen in the $B^0_d$, $B^0_s$ admixture found at high energies [5]:

$$\chi_B = 0.118 \pm 0.005$$

Using $f_d = (39.7^{+1.8}_{-2.2})\%$ and $f_s = (10.5^{+1.8}_{-1.7})\%$ [5], the integrated $B^0_s$ mixing $\chi_s$ is found to be close to saturation at 50%:

$$\chi_s = 0.473 \pm 0.108$$

(The quoted error is calculated using naive error propagation, not taking correlations into account.) The saturation of the total mixing indicates that time dependent measurements are necessary to access higher values of $x_s$.

## 2.4 $B^0$–$\overline{B}^0$ mixing in the Standard Model

In the Standard Model, $\Delta B = 2$ transitions between a neutral meson and the anti meson are possible by means of second order weak interactions. The mechanism at quark-level is shown in figure 2.2, exposing the $q\overline{q}$ annihilation channel and the double $W^\pm$ exchange channel.

Due to the large $t$ quark mass, $\Delta m_q$ is completely dominated by the $t\bar{t}$-diagrams. The mixing amplitude has been evaluated including QCD corrections (for references, see [21]):

$$\Delta m_q = \frac{G_F^2}{6\pi^2} |V_{tq}|^2 |V_{tb}|^2 M_W^2 m_{B_q} B_{B_q} f_{B_q}^2 \eta_B F\left(\frac{m_t^2}{M_W^2}\right) \quad (2.13)$$

$B_{B_q}$: Non-perturbative bag factor

$f_{B_q}$: B-decay constant

$\eta_B$: QCD-corrections

$F(x)$: Top-mass dependence

Hence, the mass differences of neutral B-systems probe directly the CKM-couplings of the $t$ quark to the $u$ and $s$ quarks. However, the hadronic factors $B_{B_q} f_{B_q}$ are known only from QCD lattice calculations and represent the main theoretical uncertainty relating $\Delta m_q$ to $V_{tq}$. Reference [17] quotes:

$$f_{B_d} = (180 \pm 50)\ \text{MeV}$$
$$B_{B_d} = 1.0 \pm 0.2$$
$$\sqrt{B_{B_s^0}} f_{B_s^0} = (210 \pm 50)\ \text{MeV}$$

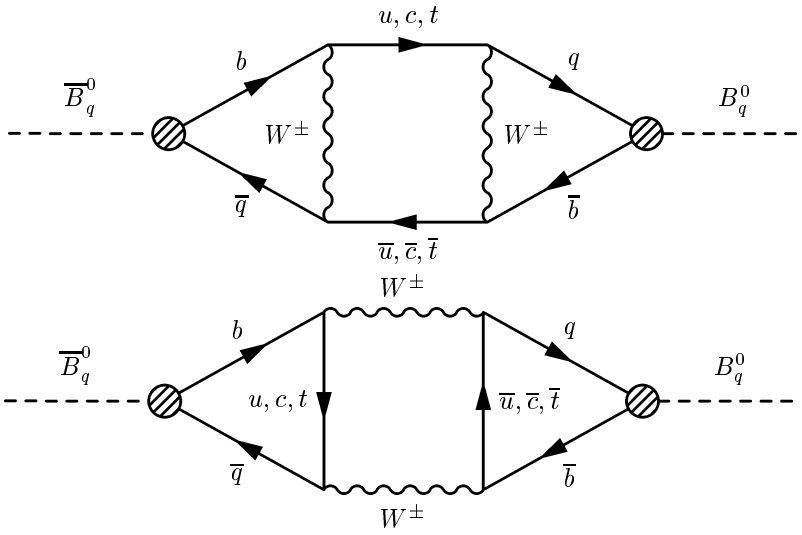


Figure 2.2: *Contributions to* $\mathrm{B}^0$*-mixing: W-exchange (top) and* $q\overline{q}$ *annihilation (bottom) diagrams. The dispersive part of the transition matrix, related to* $\Delta\Gamma_q$*, only involves on-shell intermediate states and is dominated by the c quark contribution. The non-dispersive part, which determines* $\Delta m_q$*, is dominated by virtual t quarks, due to its high mass.*

To reduce these theoretical uncertainties, it thus seems favorable to study the ratio

$$\frac{\Delta m_d}{\Delta m_s} = \left|\frac{V_{td}}{V_{ts}}\right|^2 \frac{B_{\mathrm{B}_d} f^2_{\mathrm{B}_d}}{B_{\mathrm{B}^0_s} f^2_{\mathrm{B}^0_s}},$$

as the uncertainty on the *ratio* of the hadronic factors is smaller than the uncertainty on each factor ([31], as quoted by [30]):

$$\frac{f_{\mathrm{B}^0_s}\sqrt{B_{\mathrm{B}^0_s}}}{f_{\mathrm{B}_d}\sqrt{B_{\mathrm{B}_d}}} = 1.17 \pm 0.06(\mathrm{Computational}) \pm 0.12(\mathrm{Model})$$

Recalling that only on-shell intermediate states contribute to $\mathbf{\Gamma}$, one concludes that $\Delta\Gamma_q$ is insensitive to the $t$ quark. Instead, $u\overline{u}$ or $c\overline{c}$ states dominate, as determined by the relevant CKM-matrix elements. In particular, the width difference in the $\mathrm{B}^0_s$-system is related to $V_{cs}$:

$$\Delta\Gamma_s \sim B f_B^2 |V_{cb}|^2 |V_{cs}|^2,$$

which is not too heavily CKM-suppressed. This represents a second route to handle the hadronic uncertainties that prevent extracting information about $V_{td}$ from $\Delta m_d$ measurements. The theoretical prediction [32] of the $\mathrm{B}^0_s$ width difference,

$$\frac{\Delta\Gamma_s}{\Gamma} = 0.16^{+0.11}_{-0.09},$$

indicates that the width difference could be experimentally accessible.

## 2.5 $\mathrm{B}^0_s/\mathrm{B}^0_d$ lifetime splitting.

Using the optical theorem, the total decay width of a hadron $H_b$ is given by the matrix element of the transition operator $\mathbf{T}$ [33]:

$$\Gamma_{\mathrm{B}_q} = \frac{1}{m_{\mathrm{B}_q}} \mathrm{Im}\langle B_q|\mathbf{T}|B_q\rangle$$

The transition operator is non-local,

$$\mathbf{T} = i \int \mathrm{d}^4 x T \mathcal{L}_{\mathrm{eff}}(x) \mathcal{L}_{\mathrm{eff}}(0).$$

The interaction $\mathcal{L}_{\mathrm{eff}}(x)$ is derived from the full electroweak Lagrangian by integrating out the $W$ degrees of freedom, and it contains local 4-fermion operators.The heaviness of the $b$ quark allows another Operator Product Expansion and the matrix element of $\mathbf{T}$ can be series-expanded with $1/m_b$ as parameter. Keeping only the first term (the $m_b \to \infty$ limit) gives the spectator model prediction that all $b$-hadrons have equal lifetime [34].

The corrections to the spectator model can be grouped in 3 terms:

$$\Gamma_{\mathrm{B}_q} = \Gamma_{\mathrm{spec}} + \Delta\Gamma^q_{\mathrm{kin}} + \Delta\Gamma^q_{\mathrm{mag}} + \Delta\Gamma^q_{\mathrm{WA}}$$

- $\Delta\Gamma^q_{\mathrm{kin}}$: time dilatation correction due to the movement of the $b$ quark inside the hadron.
- $\Delta\Gamma^q_{\mathrm{mag}}$: Chromo-magnetic interaction due to the spin of the $b$ quark.
- $\Delta\Gamma^q_{\mathrm{WA}}$: Weak annihilation diagrams.

A possible width difference between $\mathrm{B}_d^0$ and $\mathrm{B}_s^0$ can only come from $SU(3)_{\mathrm{flavor}}$-breaking of the already small correction terms. These terms have been evaluated in [32] and the upper bound is given as:

$$\left| \frac{\Gamma_{\mathrm{B}_s^0}}{\Gamma_{\mathrm{B}_d}} - 1 \right| < 1\%.$$

# Chapter 3

# The DELPHI experiment at LEP

DELPHI (Detector with Lepton, Photon and Hadron Identification) is a general purpose detector installed at the LEP (Large Electron Positron) $e^+e^-$-collider at CERN. It provides $4\pi$ charged particle detection and identification, electromagnetic and hadronic calorimetry, and muon identification. The general layout of the detector is shown in figure 3.1. DELPHI's main distinctions among the 4 LEP experiments are the use of RICH (Ring Imaging Cherenkov) detectors for charged particle identification, and its high granularity HPC (High density Projection Chamber) lead/gas sampling electromagnetic calorimeter capable of separating $\gamma$ from $\pi^0$ up to 25 GeV. A summary of LEP1 ($Z^0$ peak) data taking years is shown in table 3.1 The design, construction and operation of the complete DELPHI experiment are described in two papers: [35] and [36]. References to published work concerning various subdetector systems are given below when appropriate.

| Year | Hadronic $Z^0$'s ($\cdot 10^3$) | Off-peak (%) | RICH operation | Upgrades |
|---|---|---|---|---|
| 1989 | 17 | 23.5 | No RICH | |
| 1990 | 155 | 19.4 | No RICH | |
| 1991 | 327 | 15.6 | No RICH | 3 VD layers |
| 1992 | 751 | 0 | Barrel only | |
| 1993 | 998 | 24.3 | Barrel only | |
| 1994 | 1484 | 0 | Fully operational | VD $z$-layers |
| 1995 | 986 | 23.9 | Mostly operational | Extended ID |

Table 3.1: *LEP1 data taking periods at DELPHI. The numbers of hadronic $Z^0$ decays logged by the DELPHI detector are shown in the second column, the fraction of events collected during $Z^0$ lineshape scans are shown in the third column. In total, the barrel and forward RICHes were fully operational for 54% and 52% of the events respectively.*

The present analysis requires high efficiency charged track detection, high precision vertex resolution, charged particle identification separating $\pi^\pm$/K$^\pm$, electron and muon identification capabilities, and finally efficient electromagnetic energy reconstruction. The detector subsystems providing these capabilities are presented below, together with a brief description of the relevant reconstruction software. In particular, the vertex detector requirement confines the analysis to the DELPHI barrel region. For this reason, the forward tracking chambers, calorimeters and muon chambers are not described. By the same argument of relevance, the barrel hadron calorimeter is also omitted from the description below.

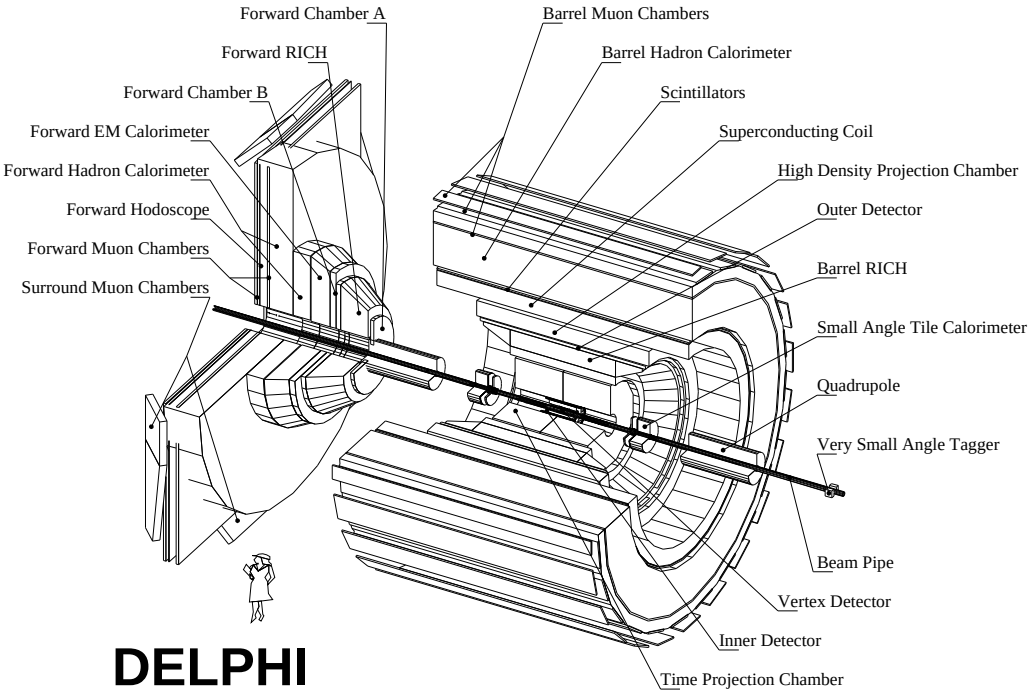


Figure 3.1: *Layout of the DELPHI detector at LEP [37]. Only one endcap is shown, in the retracted position that allows detector access during shut down periods. The vertex detector (hardly visible on this figure) is about* 22 *cm diameter, while the total length and diameter of the complete detector exceed* 10 *m.*

## 3.1 Track reconstruction

The barrel charged particle spectrometer consists of various tracking detectors placed inside a 5.2 m diameter and 7.4 m long superconducting solenoid providing a uniform 1.23 T axial magnetic field. The tracking system is highly modularized as up to 5 subdetectors can contribute to the reconstruction of a single charged track.

The DELPHI coordinate system is defined as a right-handed system with the origin at the center of the detector close to the interaction point, the $z$-axis pointing along the direction of the beam and the $x$-axis in the horizontal plane. In the barrel region, detector description and track reconstruction frequently employ cylindrical coordinates with polar and azimuthal angles $(\theta, \phi)$.

### 3.1.1 Central tracking subdetectors

#### Vertex Detector

The VD (Vertex Detector) [38] [39] consists of 3 cylindrical layers of silicon microstrip detectors arranged in 24 partially overlapping sectors at radii 6.3 cm, 9.0 cm and 10.9 cm. The full polar angle coverage is $44° \leq \theta \leq 139°$, as determined by the length of the outmost layer. Each cylinder provides $R\phi$ coordinate measurement with a readout pitch of 50 $\mu$m. In addition, the cylinders at 6.3 cm and 10.9 cm are double sided to provide $z$ measurements as well.

The VD has been aligned [40] using charged tracks and exploiting the overlap of adjacent

sectors. The alignment procedure parameterizes and takes into account complicated geometrical distortions like sag and torsion, resulting in an alignment precision at the level of 5 $\mu$m. The average single hit precision is 7.6 $\mu$m in the $R\phi$ direction and ranging from 9 $\mu$m to 30 $\mu$m in the $z$ direction, depending on track incidence angle.

#### Inner Detector

The ID (Inner Detector) [41] is really two subdetectors: The inner part is a 24-sector times 24 wire jet-chamber providing $R\phi$ measurements between 12 cm and 23 cm radius. The outer part consists of 5 cylindrical trigger layers with a coarse $R\phi$-resolution that also helps resolving left/right ambiguities in the jet-chambers.

Up to and including 1994, the trigger layer used MWPCs (Multi-wire Proportional Chambers) with cathode readout providing $z$ coordinate information as well. As part of the LEP-2 upgrade, a longer ID was installed before the 1995 data taking period. The newer detector uses 5 layers of straw tubes replacing the MWPCs, and no longer provides $z$ information.

The $R\phi$ track element resolution of the jet-chamber is 50 $\mu$m for the older detector, improving to 40 $\mu$m after the upgrade. The angular coverage improves from $23^\circ \leq \theta \leq 157^\circ$ to $15^\circ \leq \theta \leq 165^\circ$.

#### Time Projection Chamber

The TPC (Time Projection Chamber) [42] is the main tracking subdetector of DELPHI. Being divided in 6 sectors, its hexagonal geometry approximates a hollow cylinder of $2 \times 1.5$ m length and 1.2 m radius surrounding the ID (Inner Detector). The central plane at $z = 0$ is held at high negative voltage to provide a uniform axial drift field of 187 V/cm, which corresponds to a drift velocity of 7 cm/$\mu$s. The drift gas is 80% Ar 20% $CH_4$ at 1 atm and 29°C.

At each end plane, circular pads rows are read out to provide up to 16 3-dimensional space points for tracks inside a polar angle of $39^\circ \leq \theta \leq 141^\circ$. The sector walls contribute dead zones making up 4% of the subdetector $R\phi$ area. The geometric resolution for tracks is limited by distortions to 150 $\mu$m in $R\phi$ and 600 $\mu$m in $z$.

#### Outer Detector

The OD (Outer Detector) [43] consists of 5 cylinders of drift chambers at a radius of 197 cm to 206 cm covering a polar angle of $42^\circ \leq \theta \leq 138^\circ$. Three of the detector cylinders are capable of providing $z$ information using signal timing at the two detector ends. The single hit precision is 110 $\mu$m in $R\phi$ and 3.5 cm in $z$.

### 3.1.2 Offline track reconstruction and global alignment

The offline reconstruction program DELANA [44] respects the modularity of the DELPHI central tracking system [45] in the sense that a number of software modules representing individual subdetectors perform local pattern recognition and track fit. At the global level, the track elements output from the local modules are merged to track candidates. After resolving ambiguities in the combination of track elements, only successfully reconstructed tracks are retained. The final step consists of associating VD hits and finally refitting the track parameters.

Traditionally, track merging consisted of extrapolating TPC track elements inwards and outwards through the other subdetectors to build up strings of track elements. The weakness of this procedure is that tracks crossing a TPC sector wall or hitting the central plane at $\theta = 90^\circ$ are not seen by the TPC and will not be reconstructed even though they might be well measured by other subdetectors. To remedy this, an improved version of the DELANA track search was introduced. The new approach reconstructs tracks that were not recognized in the TPC or were split in two short track elements due to a TPC sector wall.

The global alignment of the various subdetectors is crucial to obtain the best possible tracking precision. Starting with optical and mechanical survey data, the global alignment procedure relies

on muon pairs from $Z^0 \to \mu^+\mu^-$ decays. Assuming perfectly collinear $e^+e^-$ beams, the two tracks of a muon pair will form a single helix. In practice, the minor beam acollinearity of LEP gives a kink at the interaction point which is corrected for in the alignment procedure.

The initial reference is the OD, whose wire position is known from surveying to a precision of 30 $\mu$m. Taking advantage of the large lever arm, the VD is aligned with respect to the OD using muon pairs. Interpolating reference tracks between the VD and the OD trough the ID, the TPC and the RICH, the position of all the barrel tracking detectors are measured.

The final charged track momentum precision in the barrel region is reported as [35]

$$\sigma_p/p = 0.15\% \cdot p/\mathrm{GeV}$$

## 3.2 Hadron identification

### 3.2.1 Track ionization measurement of the TPC

The specific ionization $\mathrm{d}E/\mathrm{d}x$ of charged tracks crossing the TPC is measured by sense wires at the end planes, just above the pad rows that provide geometric space points. The ionization of a single track is sampled with up to 192 wires. This number decreases as tracks below $\theta \leq 39^\circ$ and above $\theta \geq 141^\circ$ leave the TPC through the end planes.

Ionization clusters are associated to reconstructed tracks using the wire number and time of arrival. To handle the long upper tail of the Landau-distribution, the largest 20% of the samples are excluded, calculating the 80% truncated mean. Comparing with the calibrated expectation for $\mathrm{d}E/\mathrm{d}x$ versus particle momentum and mass, K and $\pi$ with momentum above 2 GeV/c can be distinguished at the $1\sigma$ level, see figure 3.2

As the sense wire does not resolve the $R\phi$-direction, the ionizations of close tracks cannot be disentangled if the $z$-direction difference is less than 2 cm. This somewhat reduces the usefulness of the $\mathrm{d}E/\mathrm{d}x$ measurements for particle identification in hadronic jets.

### 3.2.2 Cherenkov radiation and the RICH

When relativistic particles traverse an optically dense medium (refraction index $n > 1$), photons are emitted at an angle depending on $n$ and the particle velocity $\beta$ [4]:

$$\cos\theta_\mathrm{C} = \frac{1}{n\beta}. \tag{3.1}$$

The threshold for light emission is the local speed of light $1/n$, and the Cherenkov radiation is the electromagnetic equivalent of supersonic shock waves. The number of photons emitted in a given part of the spectrum for a given path length is proportional to $\sin^2\theta_\mathrm{C}$.

The DELPHI BRICH (Barrel Ring Imaging Cherenkov detector) [46] is situated between the TPC and the OD. Figure 3.3 illustrates the principle of operation.

Tracks first cross the liquid radiator ($C_6F_{14}$ with $n = 1.2718$ enclosed in UV-transparent quartz), giving on average 12 photons per track. By proximity focusing, the photons are projected in a ring around the track in the photon detector.

Outside the photon detector, the tracks traverse the gas radiator ($C_5F_{12}$, $n = 1.00194$). The photons emitted here are reflected by spherical mirrors and focused onto the photon detector. On average, 8 photons are emitted per track.

The photon detector is a time projection chamber with UV-transparent quartz windows. The drift gas contain 0.1% TMAE photosensitive gas that converts UV-photons to photoelectrons. The photoelectrons (and the electrons created by track ionization) drift towards the end of the chamber, where they are detected by a multi wire proportional chamber. Using the drift time, the 3-dimensional coordinate of the photo conversion is reconstructed.

In principle, DELPHI RICH particle identification is done looking for rings of converted photons. After rejecting clusters that are most probably due to track ionization or detector noise,

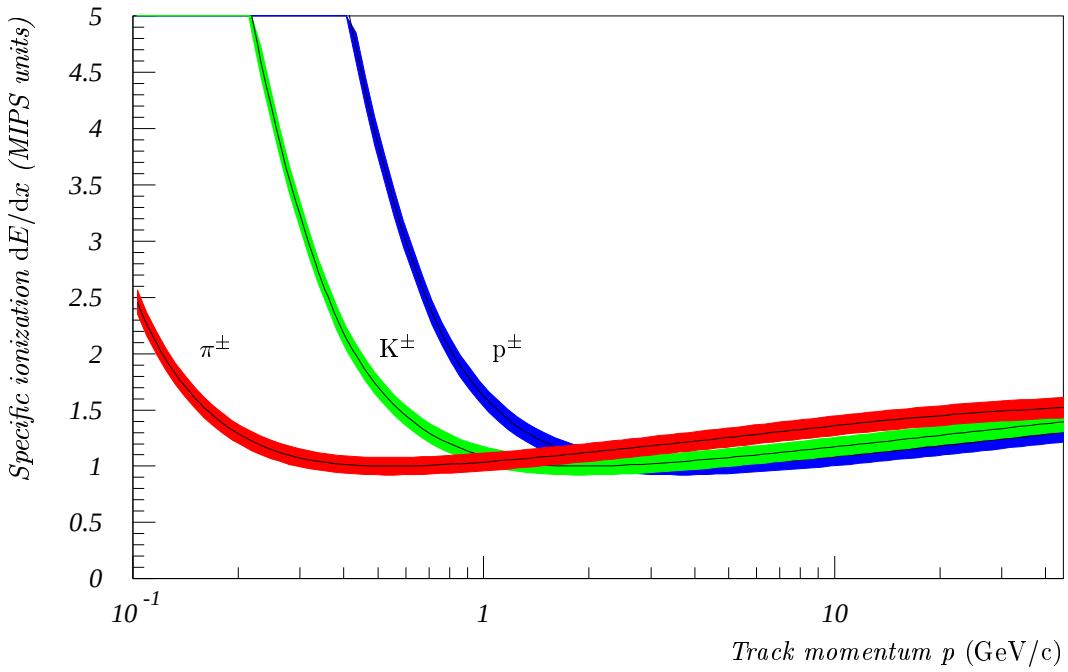


Figure 3.2: *Particle identification capabilities of the TPC* $\mathrm{d}E/\mathrm{d}x$ *measurement. The shaded bands indicate minimal* $\pm 1\sigma$ *errors on* $\mathrm{d}E/\mathrm{d}x$ *for a truncated mean of 192 sense wires. In hadronic jets, tracks are often so close in space that it is not possible to separate charge clusters of different tracks, resulting in fewer than 192 measurements per track.*

photoelectron candidates are assigned to close charged tracks. The radiation point is estimated using the track extrapolation through the RICH and eventually the single photon Cherenkov angle is calculated.

Having associated photons to charged tracks and reconstructed the individual Cherenkov angles, three different approaches to particle identification [47] have been used in the analysis of DELPHI data. All three methods use the number of reconstructed photons as well as their Cherenkov angles:

- `HADSIGN` Fitting a flat background level, the likelihood of each mass hypothesis ($\pi$, K, p) is calculated using the expected Cherenkov angle and expected number of photons.

- `RINGSCAN` The likelihood ratio of the signal+background hypothesis versus the background hypothesis is maximized with respect to a continuously varying Cherenkov angle.

- `RIBMEAN` Photons close in $\theta_{\mathrm{C}}$ are clustered and the mean Cherenkov angle is computed using the reconstruction errors and some quality weight.

For analyses that require good efficiency instead of maximum background rejection, the `HADSIGN` approach has proven useful. Expected Cherenkov angles with typical errors as can be reconstructed using the `RIBMEAN` method is shown in figure 3.4.

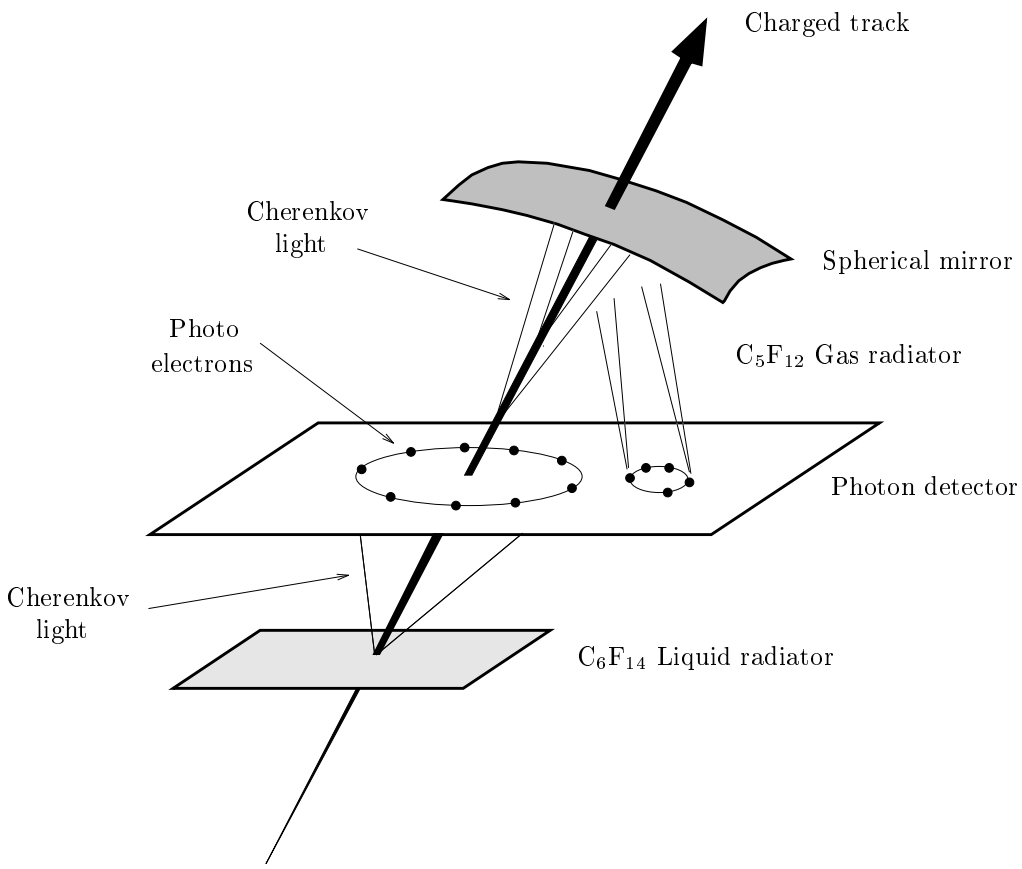


Figure 3.3: *DELPHI RICH principle of operation [37]. A single photon detector is used with both radiators. By proximity focus, Cherenkov photons emitted in the liquid radiator end up on a ring in the photon detector. Gas radiator photons are reflected and focused by spherical mirrors.*

## 3.3 Electromagnetic calorimetry

The HPC (High density Projection Chamber) [48] is a lead/gas sampling electromagnetic calorimeter of 208 cm inner radius and 260 cm outer radius. The lead conversion material constitutes $18./\sin\theta$ radiation lengths. The ionization charge created by electromagnetic showers is drifted in the $z$-direction and detected by MWPCs at the end planes, exploiting parallel $\vec{E}$ and $\vec{B}$ fields to minimize transverse dispersion. The resulting granularity is 2 mrad $\times$ 20 mrad in the $\theta$ and $R\phi$ directions. The energy precision can be parameterized as

$$\sigma_E/E = 0.043 \oplus 0.32/\sqrt{E(\mathrm{GeV})}$$

## 3.4 Lepton Identification

### 3.4.1 Electron identification

Several DELPHI subdetectors can contribute to identifying electrons. The most obvious indication is a charged track pointing to an electromagnetic shower in the HPC. Secondly, as the electrons are ultra-relativistic, the specific ionization measured by the TPC and the Cherenkov angle measured by the RICH are expected to be saturated at the high-velocity limit. Finally, tracks radiating

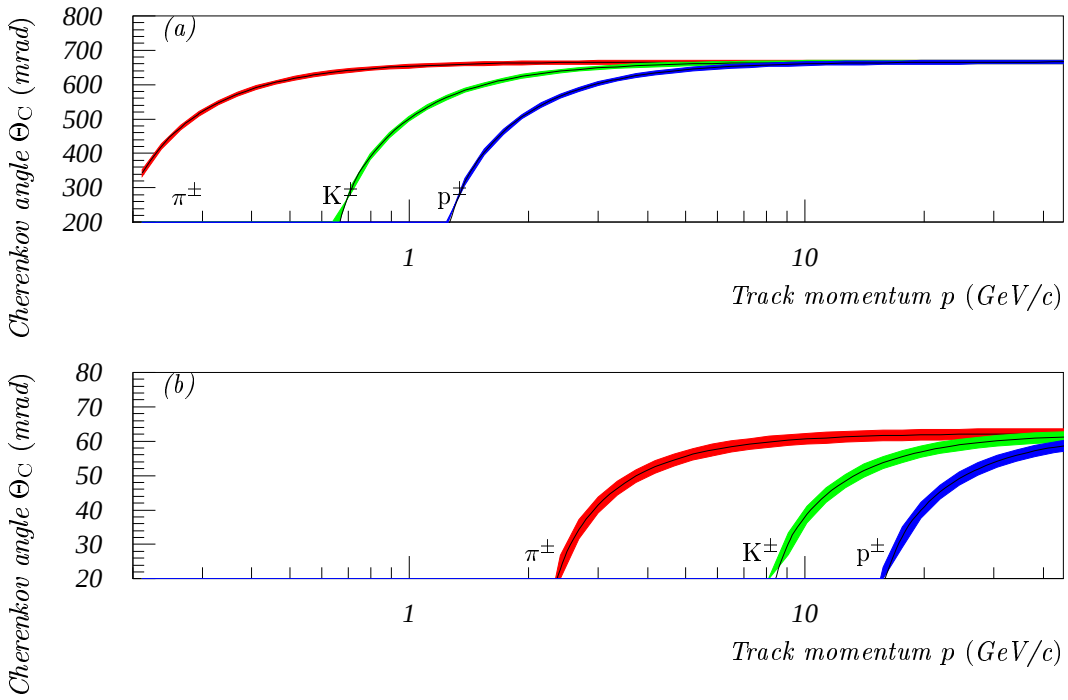


Figure 3.4: *Expected Cherenkov angles in the Barrel RICH. The shaded bands indicate typical $\pm 1\sigma$ errors on the measured Cherenkov angle, extrapolating from the Cherenkov angle error at saturation, measured in $Z^0 \to \mu^+\mu^0$ events [36]. In multihadronic events, the errors will generally be larger.*

hard photons as found by the secondary interactions reconstruction program `ELEPHANT` [49] are strong electron candidates. The different pieces of information are combined using neural network techniques [50] to provide an optimal discriminant.

### 3.4.2 Muon identification

The solenoid flux return yoke acts as a muon filter stopping most hadrons from reaching the MUB (MUon chambers, Barrel). Muon candidate tracks are extrapolated through the non-constant magnetic field of the yoke looking for matching hits in the muon chambers. After dealing with ambiguous hit assignments, a discrete muon tag variable is computed [51]. The "loose" tag has a $(94.8 \pm 0.1)\%$-efficiency for 45 GeV muons coming from $Z^0 \to \mu^+\mu^-$, and a misidentification probability of $(1.5 \pm 0.1)\%$ evaluated on $\tau \to 3\pi\nu_\tau$-events.

## 3.5 Trigger subsystem

The two lowest levels of the DELPHI trigger system [52] [53], T1 and T2, operate synchronously to the bunch crossings, while the two higher levels, T3 and T4, are implemented as asynchronous software triggers. In the normal $4+4$ bunch operation mode of LEP, there is $11\mu$s between bunch crossings. T1 is available $3.5\mu$s after the bunch crossing, and T2 takes $39\mu$s to decide. The event read-out triggered by a T2 yes decision takes 3ms on average. The total dead-time introduced by triggers is 3%.

The input to T1 is supplied by fast tracking detectors (ID trigger layer, OD, FCA and FCB),

and by scintillators in the barrel and in the forward region (TOF, HPC scintillators and HOF) and by barrel muon chambers (MUB). At the T2 level, this information is supplemented by information from detectors that take longer to read out: TPC, HPC and MUB. The purpose of the software triggers T3 and T4 is to reduce the data logging rate. T3 use logic similar to T2 but with more detailed information, and T4 is in fact a special version of the DELPHI reconstruction program DELANA [44].

The charged track trigger is based on track elements from TPC, FCA/FCB, ID, OD and TOF, and is the only component of importance for triggering hadronic $Z^0$ decays. Overlapping acceptance of different trigger components allows evaluation of the efficiency. Due to the large charge track multiplicity of hadronic $Z^0$ decays, [36] claims that the trigger efficiency is hardly distinguishable from 1.

## 3.6 Data flow

The physics data flow in the DELPHI experiment is sketched in figure 3.5. Following a positive

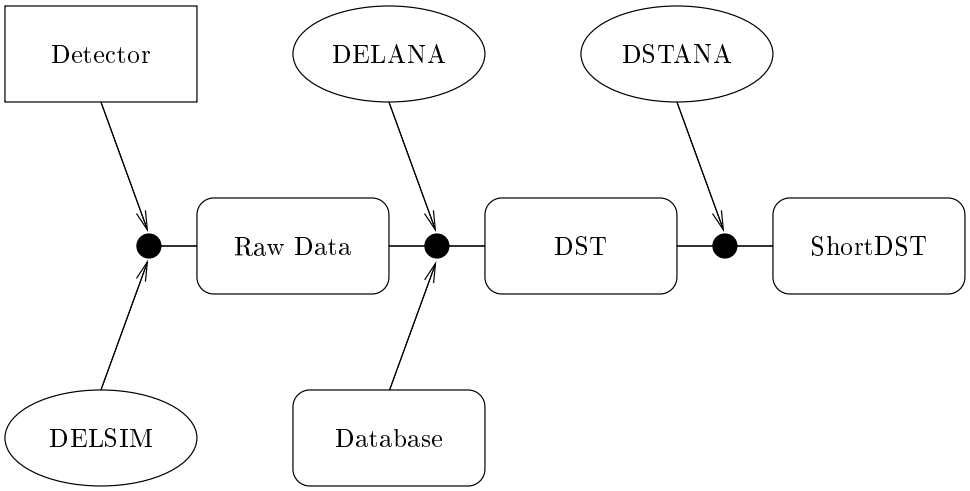


Figure 3.5: *Physics data flow in DELPHI. Ovals represent programs, rectangles represent data and black dots represent processing. As detector calibration and alignment improve with time, 2–3 DELANA reprocessings of the data from each year have been necessary to reach the optimum data analysis quality.*

trigger decision, the detector data collected by the distributed front end system is assembled into a global event structure and logged on the main raw data stream in the ZEBRA [54] format. DELANA [44], the offline reconstruction program, combines information from the geometrical and calibration databases [55] based on the CARGO [56] system with the raw detector data. DELANA performs pattern recognition and reconstructs charged tracks, calorimeter energy deposits, hits in the muon chambers and RICH photoelectrons. The reconstructed events are written to Data Summary Tapes (DST) [57]. Foreseeing the need for detector studies, realignment, and more elaborate particle identification software, without having to go back to the raw data, the DSTs contain information at a rather detailed level: single detector track elements, details on calorimeter clusters and individual RICH photoelectrons. For this reason the complete DST data sets are quite bulky and normally not used for physics analyses.

Mainly due to the complexity of the DELPHI detector system, it has proven necessary to maintain a second stage of centrally managed analysis programs. DSTANA is a collection of computer codes that do detector realignment and response correction, HPC calibration, $V^0$ and secondary interaction reconstruction and RICH particle identification 3.2.2. The output format

of DSTANA is derived from the DELPHI full DST format but tailored to specific analysis needs: ShortDST [58] is the preferred format for analyses of hadronic final states. The LongDST format contains the full DST information in addition to the ShortDST banks and is used for detector studies and leptonic final states analyses. Various MiniDST formats are even more abbreviated than ShortDST and contain only information needed for a few specific physics analyses.

Monte Carlo simulated events are analyzed similarly to data. The different kinds of $Z$ final states are generated by various event generators. For the present work, $Z \rightarrow q\overline{q}$ events are generated by JETSET [59]. The generated particles are tracked through the detector, simulating detector response using the full simulation program DELSIM [60], producing simulation raw data with simulation truth information added. The simulated events are reconstructed using the same DELANA analysis program as with real data, producing events in the DST format. The association between generated and reconstructed tracks is performed looking for matching track direction, charge and momentum. Finally, the simulated events are run through DSTANA. In addition to the reconstruction tasks performed on data, the simulated events receive resolution and efficiency corrections to optimize the Monte Carlo versus data correspondence.

# Chapter 4

# Inclusive $\mathrm{B}_s^0 \to \phi\ell^+$ reconstruction

## 4.1 Rationale for analyzing the $\mathrm{B}_s^0 \to \phi\ell^+$ channel

Most analyses involving $\mathrm{B}_s^0$[1] mesons at LEP are based on the semileptonic decay $\mathrm{B}_s^0 \to \mathrm{D}_s^-\ell^+\nu_\ell$, $\ell^+$ meaning $e^+$ or $\mu^+$., with $\mathrm{D}_s^-$ completely reconstructed. This has been studied by OPAL [26], ALEPH [61] and DELPHI [62]. (The figures of merit as quoted below are taken from the DELPHI publication.) In addition, inclusive hadronic decays ($\mathrm{B}_s^0 \to \mathrm{D}_s^-\pi^+X$) [63] [64] and even exclusive hadronic modes [65] have been used. From a resolution point of view, completely reconstructed decay channels are very desirable. They do, however, suffer from tiny branching fractions. The only measured all-charged mode so far [5], $Br(\mathrm{B}_s^0 \to J/\psi(1\mathrm{S})\phi) = (9.3 \pm 3.3) \times 10^{-4}$, is not even useful for mixing analyses as this final state does not contain information about the $\mathrm{B}_s^0$ decay flavor.

Requiring a lepton with a high transverse momentum reduces background from $\mathrm{Z}^0 \to \mathrm{c\overline{c}}$ and $b \to c \to l$, typically to a level of 10% in semileptonic analyses. The rather large opening angle of the decay products also gives a good handle on the flight length reconstruction. A drawback of the semileptonic mode is that the unobserved neutrino carries away an unknown fraction of the $\mathrm{B}_s^0$ energy. This energy is of course necessary to reconstruct the $\mathrm{B}_s^0$ decay time. The problem is not as severe as it might seem. The $b$-fragmentation function is rather peaked. The average fraction of the beam energy carried by the $b$-hadron is [66]

$$\langle x_E \rangle = 0.716 \pm 0.0006(stat) \pm 0.007(syst),$$

with the maximum of the distribution at about 0.85. Truncating the lower tail at $x_E^{\mathrm{cut}} = 0.44$, the relative spread of the b hadron energy is down at 18%, calculated from the differential cross section given in the previous reference. Using the $e^+e^-$ beam constraint, it is possible to improve the energy resolution further.

In existing LEP analyses, the secondary charm meson is normally completely reconstructed, one notable exception being the $\mathrm{D}_s^+ \to \phi\ell^+\nu_\ell$ channel. In analyses with a fully reconstructed $\mathrm{D}_s^+$, cutting on the invariant mass allows for good background suppression. These analyses operate at a typical combinatorial background of $f_{\mathrm{Bkg}} = (33 \pm 5)\%$ ($\phi\pi^+$ mode). The contamination from $\mathrm{B}_d^0$ mass reflection, caused by a pion misidentified as kaon, is also low. A typical number is $f_{\mathrm{refl}}/f_{\mathrm{B}_s^0} = (6.9 \pm 2.3)\%$ for the $\mathrm{K}^{*0}\mathrm{K}^+$ mode. However, the all-charged decay branching ratio for $\mathrm{D}_s^+$ is quite small, see table 4.1, and most analyses require additional subresonances like $\phi$ or $\mathrm{K}^{*0}$ in the final state to reduce background. The DELPHI analysis cited above observes $138 \pm 14$ all-charged signal events on 1991–1994 data.

The low all-charged branching ration of the $\mathrm{D}_s^-$ motivates the use of inclusive final states correlated to a high transverse momentum lepton. A particularly feasible class of inclusive decays is $\mathrm{D}_s^- \to \phi h^- \mathrm{X}^0$, its branching ratio is $\mathrm{Br}(\mathrm{D}_s^- \to \phi + \text{anything}) = (18^{+15}_{-10})\%$ [5]. The unidentified particle $h^-$, in most cases a pion, carries the charge of the $\mathrm{D}_s^-$ and is thus of opposite charge to the

[1] Charge conjucate states are always implied

| $D_s^+$ final state | Branching ratio (%) |
|---|---|
| $K^+K^-\pi^+$ | $4.4 \pm 1.2$ |
| $K^+K^-\pi^+\pi^+\pi-$ | $0.83 \pm 0.33$ |
| $\pi^+\pi^+\pi^-$ | $1.0 \pm 0.4$ |
| $\pi^+\pi^+\pi^+\pi^-\pi^-$ | $0.69 \pm 0.30$ |
| $K^+\pi^+\pi^-$ | $1.0 \pm 0.4$ |

Table 4.1: *All-charged decay modes of the $D_s^+$ meson as listed in the 1998 Review of Particle Physics [5], see which for information on subresonance final states.*

lepton. The $h^-$-track may or may not be reconstructed, but if observed, it will improve flight length resolution as well as energy resolution. The unobserved decay fragments $X^0$ can be any totally neutral combination of photons and charged as well as neutral pions and kaons, as permitted by the $D_s^-$ invariant mass. The narrow width of the $\phi$ allows for combinatorial background suppression, but the background cannot be expected to be as low as in the all-charged decay modes: There is no $D_s^-$ invariant mass that can distinguish $D_s^- \to \phi\pi^- X^0$ from $D^- \to \phi\pi^- X^0$ and $\overline{D}^0 \to \phi X^0$. A significant level of $B^+$, $B^0$ background will of course dilute $B_s^0$ measurements, and the uncertainty of this level will give an important contribution to the systematic errors. As several final state particles remain unobserved, the energy resolution will also suffer. To summarize, it can be expected that the inclusive $B_s^0 \to \phi h^-\ell^+$ analysis will give a sizeable number of events, at the expense of larger background and worse resolution than analyses of completely reconstructed $D_s^+$ final states. Nevertheless, the severeness of these problems cannot be known unless a complete analysis is done.

## 4.2 Experimental data and Monte Carlo simulation.

The analysis is done on data collected during the years 1994 and 1995, a total of 2470K hadronic $Z^0$ decays, see table 3.1. Using the standard value [5] for $R_b = (21.69 \pm 0.12)\%$, this corresponds to $1072K$ $b$ quarks. The datasets from the years 1991–1993 data are not used; the final improved DELANA [44] reprocessings were only available from the beginning of 1998 and the differences in detector setup require a completely separate treatment of the older data.

All Monte Carlo simulated events are generated by JETSET [59] and passed through DELSIM [60], the full detector simulation program. For background description, the standard DELPHI $Z^0 \to q\overline{q}$ Monte Carlo is used, the number of events corresponds to twice the data statistics. To study signal efficiency and resolution, a sample of simulated signal events is obtained, partially from the standard DELPHI $Z^0 \to b\overline{b}$ simulations, and partially from a dedicated sample, selecting $B_s^0$ decaying with a lepton and a $\phi$ in the final state. The number of simulated signal events exceeds 20 times the signal expected in data. All simulated events are subject to the same selection cuts and reconstruction procedure as data.

## 4.3 Event selection

Pairs of tracks with opposite charges are considered as candidates for coming from a $\phi \to K^+K^-$-decay. Minimum cuts on the track momentum $p_{K^\pm}$ as well as the combined momentum $p_{K^+K^-}$ are applied to suppress combinatorial background. These cuts are determined using the procedure outlined in section 4.4.1:

$$p_{K^+K^-} \geq 3.0 \text{ GeV/c} \tag{4.1}$$

$$p_{K^\pm} \geq \frac{p_{K^+K^-}}{2} - \frac{\sqrt{(p^2_{K^+K^-} + m^2_{\text{crit}})(m^2_{\text{crit}} - 4m^2_{K^\pm})}}{2m_{\text{crit}}} \tag{4.2}$$

$$m_{\mathrm{crit}} = 1.04\ \mathrm{GeV/c}$$

In order to further reduce combinatorial background from $\pi^+\pi^-$-pairs, the combined hadron identification capabilities of the RICH and the specific ionization measurements of the TPC are used. In this analysis, tags from the `HADSIGN` package (see section 3.2.2) are used in the following way: A candidate pair is kept only if none of the tracks are identified as a pion, *and* at least one of the tracks is tagged as a kaon with the `very loose` tag.

Lepton candidates are subject to cuts on momentum $p_\ell$ and transverse momentum $p_-^{\mathrm{out}}$:

$$p_\ell \geq 3.0\ \mathrm{GeV/c} \quad (4.3)$$
$$p_-^{\mathrm{out}} \geq 1.0\ \mathrm{GeV/c} \quad (4.4)$$

The procedure used to optimize these cuts is exposed in section 4.4.2. Muon candidates are considered if their muon tag (see section 3.4.2) is `loose` or better. For electron candidates, a neural network (section 3.4.1) output above 0.7 is required. Electron candidates that are tagged as coming from photon conversion ($\gamma \to \mathrm{e^+e^-}$) are discarded.

As the hypothesised $\phi \to \mathrm{K^+K^-}$ is supposed to come from a $\mathrm{D_s^+} \to \phi X^+$ decay, there is always at least one more charged track coming from the $\mathrm{D}_s$ decay vertex. Looking at tracks with charge opposite to the identified lepton, the track with the highest momentum projected along the $\phi \to \mathrm{K^+K^-}$ direction is associated to the $\phi \to \mathrm{K^+K^-}$ pair. As these tracks will be used for $\mathrm{D_s^+} \to \phi(\mathrm{K^+K^-})\mathrm{h}^+X^0$ pseudotrack reconstruction, it is required that they all come from a common vertex. The momentum distribution of associated tracks is shown in figure 4.1. A lower momentum cut of 0.4 GeV/c is imposed by the DELPHI tracking acceptance. Thus, the first bin of the histogram in figure 4.1 represents events with no associated hadron. The overall efficiency for associating a charged hadron in data is $(88.7 \pm 1.1)\%$, compared to $(88.2 \pm 0.8)\%$ for Monte Carlo simulated events. In the sample of simulated signal events, $(94.9 \pm 0.3)\%$ of the reconstructed events have an associated hadron and according to the simulation truth information, $(81.1 \pm 0.6)\%$ of the associations are correct.

## 4.4 Cut optimization

In order to get the best possible measurement from the available data, the kinematical cuts of the analysis have to be optimized. A few general viewpoints on cut optimization are presented in appendix B. It is argued that cuts should be chosen to maximize the number of signal events divided by the square root of the total number of events, because ratio is related to the statistical significance of the sample. Naming the number of signal and background events $S$ and $B$, respectively, this is equivalent to optimizing $\frac{S^2}{S+B}$. The procedure used to choose the optimum kinematic cuts for the $\phi$ candidate momentum and the lepton momentum and transverse momentum is shown in detail below. The cut on the reconstructed energy is optimized using a similar procedure, its treatment is postponed until section 4.5.2.

### 4.4.1 Kinematic cuts on the $\phi \to \mathrm{K^+K^-}$ candidates

The combinatorial background to the $\phi$ is dominated by low-momentum fragmentation tracks. Thus, introducing a lower cut on the momentum of the $\phi \to \mathrm{K^+K^-}$ candidate will reduce the background. Such a cut is even more important as the $\phi$'s from $\mathrm{B_s^0}$ decays peak at a higher momentum than $\phi$'s coming from fragmentation. Thus, only B-decay $\phi$'s are counted as signal. The background is modelled by combinatorial $\mathrm{K^+K^-}$-pairs having an invariant mass less than 1.15 GeV/$c^2$. The momentum distributions and the corresponding cut optimization are plotted in figure 4.2. Maximizing signal significance gives a lower momentum cut of $p_{\mathrm{K^+K^-}} > 3.0$ GeV/c.

Given the $\phi \to \mathrm{K^+K^-}$-candidate momentum and invariant mass, the smallest possible momentum carried by the softest kaon can be computed using 4-momentum conservation and assuming

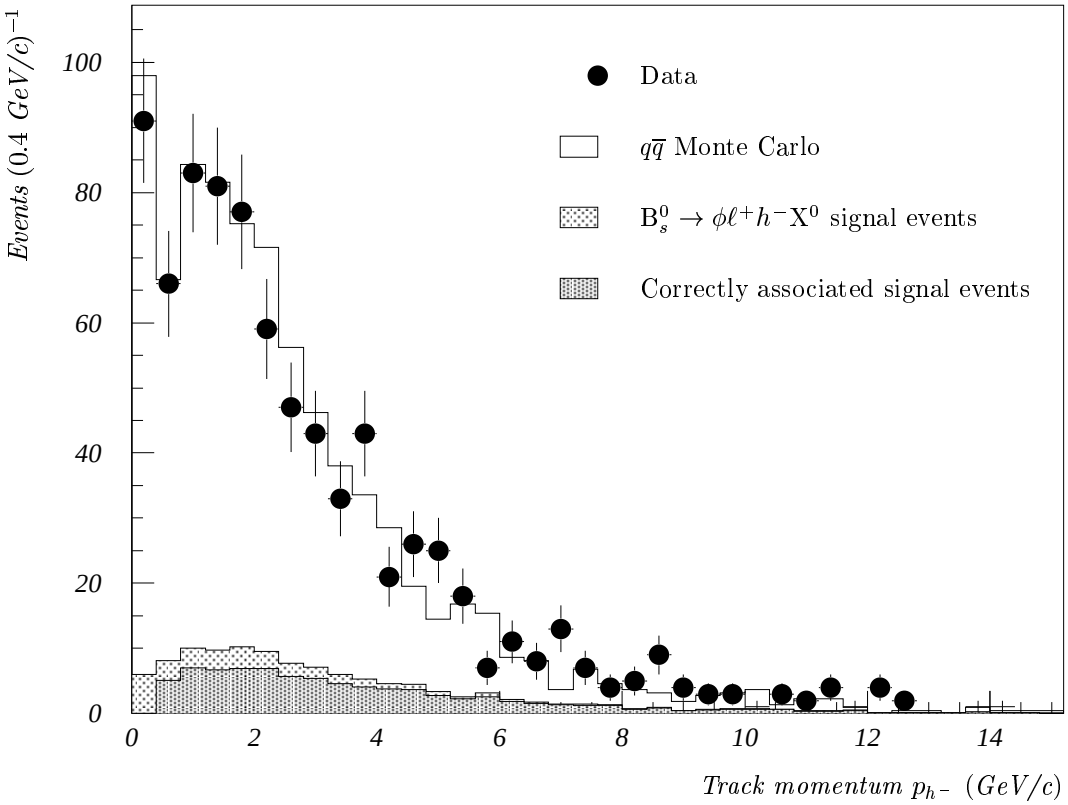


Figure 4.1: *Momentum spectrum of opposite charge associated track. The first bin (0...0.4 GeV/c) represents events where no association could be made. The full Monte Carlo histogram (open) is normalized to the number of events in data. The signal histograms (shaded) are evaluated on a dedicated Monte Carlo sample and normalized to the number of signal events expected in data.*

collinear kaons:

$$p_{\mathrm{K}}^{\mathrm{min}} = \frac{p_{\mathrm{K^+K^-}}}{2} - \frac{\sqrt{(m^2_{\mathrm{K^+K^-}} - 4m^2_{\mathrm{K}})(p^2_{\mathrm{K^+K^-}} + m^2_{\mathrm{K^+K^-}})}}{2m_{\mathrm{K^+K^-}}} \tag{4.5}$$

For a fixed $m_{\mathrm{K^+K^-}}$, requiring the $\mathrm{K}^\pm$ momentum to be above $p_{\mathrm{K}}^{\mathrm{min}}$ will only affect $\phi \to \mathrm{K^+K^-}$-candidates with an invariant above the given $m_{\mathrm{K^+K^-}}$. Choosing $m_{\mathrm{K^+K^-}}$ somewhat above the nominal $m_\phi + \Gamma_\phi$ gives a minimum $\mathrm{K}^\pm$ momentum cut that damps the upper sideband of the $\phi \to \mathrm{K^+K^-}$-candidate invariant mass spectrum. Substituting the lower $\phi$ momentum cut for $p_{\mathrm{K^+K^-}}$ gives a fixed cut. Alternatively, keeping $p_{\mathrm{K^+K^-}}$ gives a running cut for the minimum $\mathrm{K}^\pm$ momentum. Figure 4.3 shows the resulting effect on the $\mathrm{K^+K^-}$-pair invariant mass spectrum. Evidently, this cut does not improve signal-to-background conditions. Rather, it is included to make sure that the kinematics of the events in the sidebands are similar to the combinatorial events under the $\phi$ peak. This is important because the sideband events will be used in the likelihood fit to parameterize the combinatorial background.

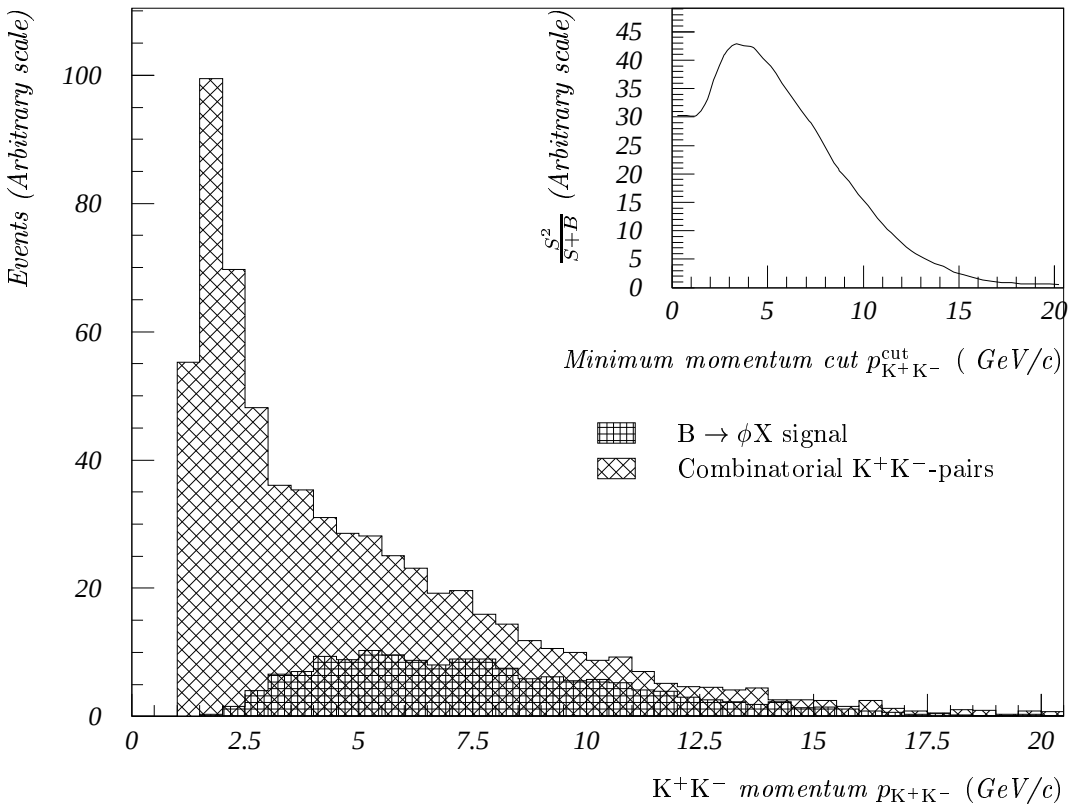


Figure 4.2: *Kaon pair momentum distribution and cut optimization. The background distribution is evaluated on combinatorial charged track combinations in $q\overline{q}$ Monte Carlo while the signal distribution comes from true* $B \to \phi M$ *events in a dedicated Monte Carlo sample, normalized to the full Monte Carlo statistics. The inset shows the signal significance* $\frac{S^2}{S+B}$ *as a function of the minimum momentum cut. A broad peaks around* 3 *GeV/c is observed.*

### 4.4.2 Kinematic cuts on the lepton candidates

The hard fragmentation of the $b$ quark [49] means that a $b$ meson carries a relatively high momentum which is subsequently shared among its decay products. The semileptonic decay is 3-body and the momentum of the direct lepton will be relatively high. Further down the decay chain, the initial energy will be shared among more tracks, and the cascade leptons produced in $c$ meson decays will have a notably lower momentum.

The scale of energy available in semileptonic $b$ meson decays is determined by the large $b$-$c$ quark mass difference. The $B_s^0$ mass is measured [4] to be 5.396 GeV/$c^2$ whereas the $c$ mesons and resonances range from 1.968 GeV/$c^2$ ($D_s$) to 2.5 GeV/$c^2$ ($D^{**}$). This energy scale determines the magnitude of the direct lepton transverse momentum, computed with respect to its containing jet's axis. The transverse momentum of fragmentation tracks is at a scale set by the fragmentation process. The transverse momentum of cascade decay tracks is low as well, as the kinetic energy released in $c$ quark decays is small and shared among several tracks.

The lepton transverse momentum $p_\perp^{out}$ is computed by projecting the lepton momentum into the transverse plane of its jet. The lepton is not included when the direction of the jet is determined.

The lepton candidate momentum distributions for signal and background events are shown in

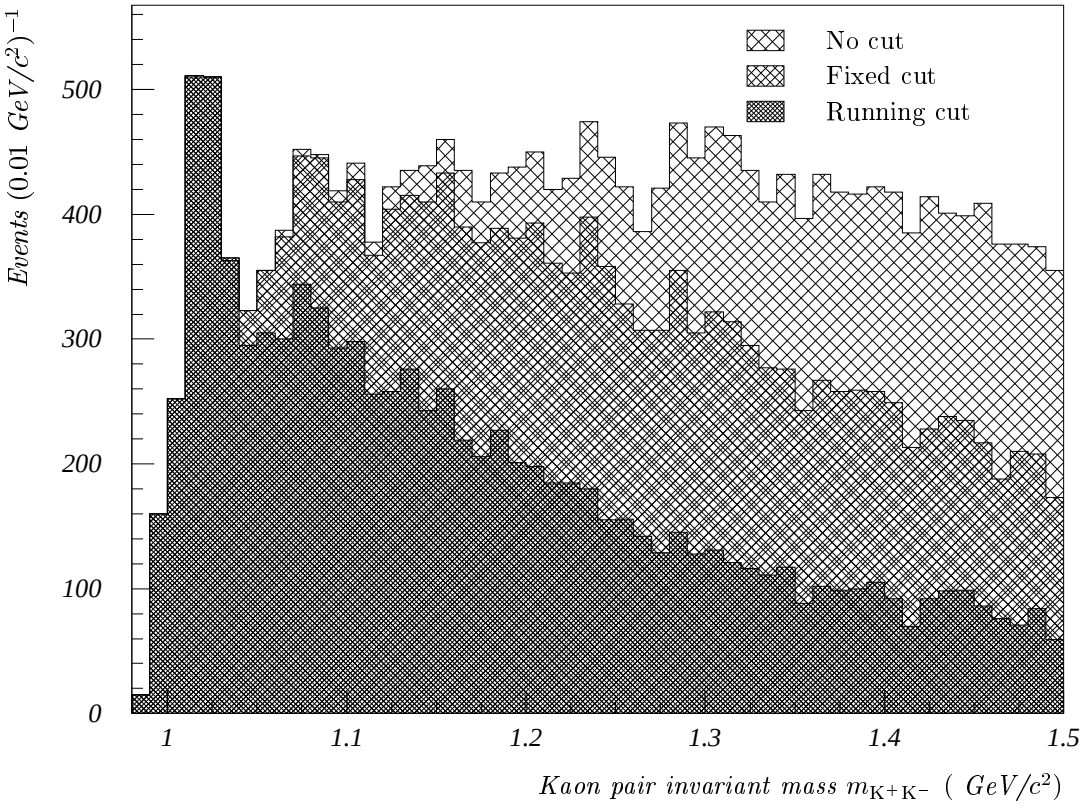


Figure 4.3: *Invariant mass* $m\mathrm{K}^+\mathrm{K}^-$ *distributions after minimum* $\mathrm{K}^\pm$ *momentum cuts. The histograms show the distributions of* $q\overline{q}$ *Monte Carlo events after applying the various* $p_{\mathrm{K}^+\mathrm{K}^-}$ *cut described in the text.*

figure 4.4. Figure 4.5 shows the corresponding distributions of the lepton candidate transverse momentum. The lepton candidate momentum and transverse momentum are strongly correlated. Optimizing the cuts in one variable at the time would thus require iteration to obtain the global maximum. It is more convenient to optimize these two cuts simultaneously. A contour plot of the signal significance $\frac{S^2}{S+B}$ as a function of the cut values is shown in figure 4.6.

To be conservative, the cuts shown in equation 4.3 are a bit looser than what seems to be the maximum in figure 4.6.

## 4.5 Decay time reconstruction

To measure the decay time in the rest frame of $\mathrm{B}_s^0$, it is necessary to measure the flight length in the laboratory frame as well as the $\mathrm{B}_s^0$ energy:

$$t = \frac{l}{\beta\gamma} = l\frac{m_{\mathrm{B}_s^0}}{p_{\mathrm{B}_s^0}} \tag{4.6}$$

The flight length resolution is of course determined by the resolution of the vertex detector whereas the energy resolution is limited by physics: In all semileptonic decays the neutrino carries away an unknown fraction of the total energy.

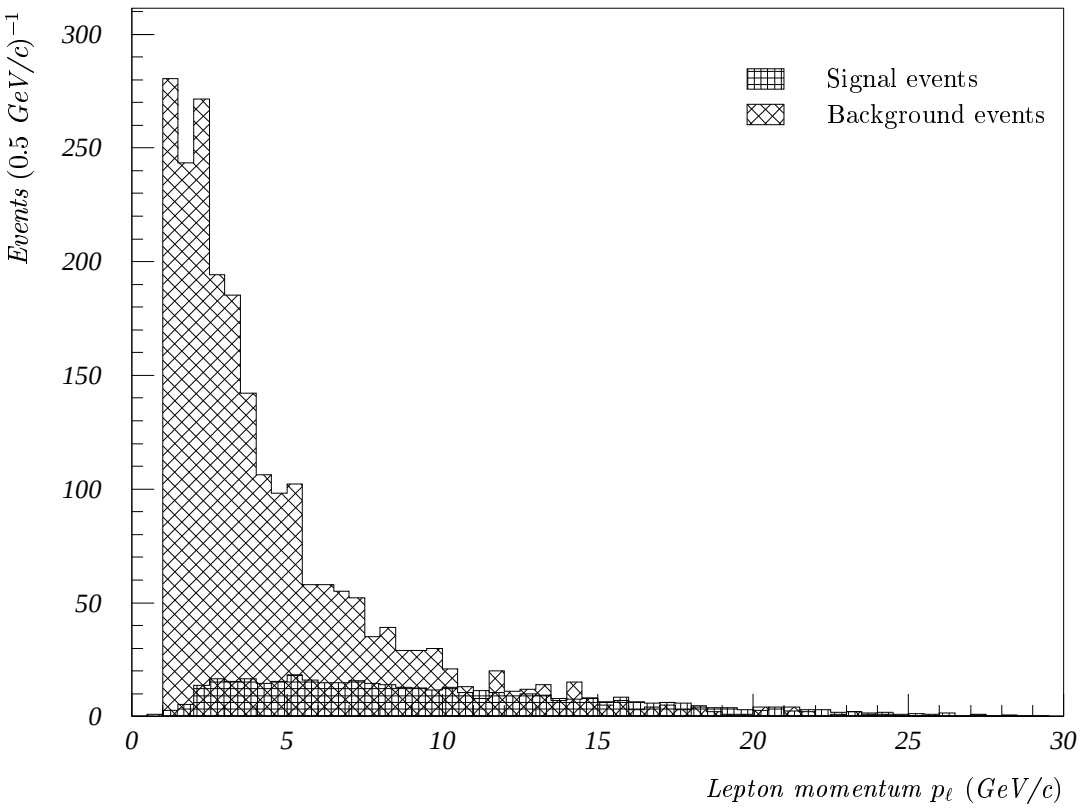


Figure 4.4: *Lepton candidate momentum distributions. The background distribution is evaluated on $q\overline{q}$ simulations using events where the lepton candidate (fake or true) does not come from a* B. *The signal distribution is taken from a dedicated sample of semileptonic b decays, normalized to the statistics of the full simulation.*

The two following subsections describe how the flight length and the $B_s^0$ meson energy are measured.

### 4.5.1 Geometric reconstruction

A complete geometric reconstruction of the b $\rightarrow$ c $\rightarrow$ (uds) decay chain means that the c meson decay vertex as well as its flight direction is needed. Figure 4.7 sketches the principle of the geometric event reconstruction.

In events where the associated hadron $h^-$ is reconstructed, the $K^+K^-h^-$ tracks are combined in a full three dimensional vertex fit. The $\chi^2$-probability of the vertex fit is required to be above $10^{-10}$. This very modest cut is a compromise between suppressing random combination of background tracks and keeping as many well measured signal events as possible. For the same reason, the 3 tracks are only required to have at least one associated vertex detector hit in total. The $\chi^2$-probability distributions of data events and simulated events are shown in figure 4.8. The inset of figure 4.8 shows the difference between the reconstructed and the true $D_s^+$-decay vertex projected along the $D_s^+$ flight direction. The distribution is well fitted by a sum of two Gaussian distributions of negligible means and variances $\sigma_1 = 244\mu$m and $\sigma_2 = 767\mu$m, accounting for 52.7% and 47.3% of the sample, respectively.

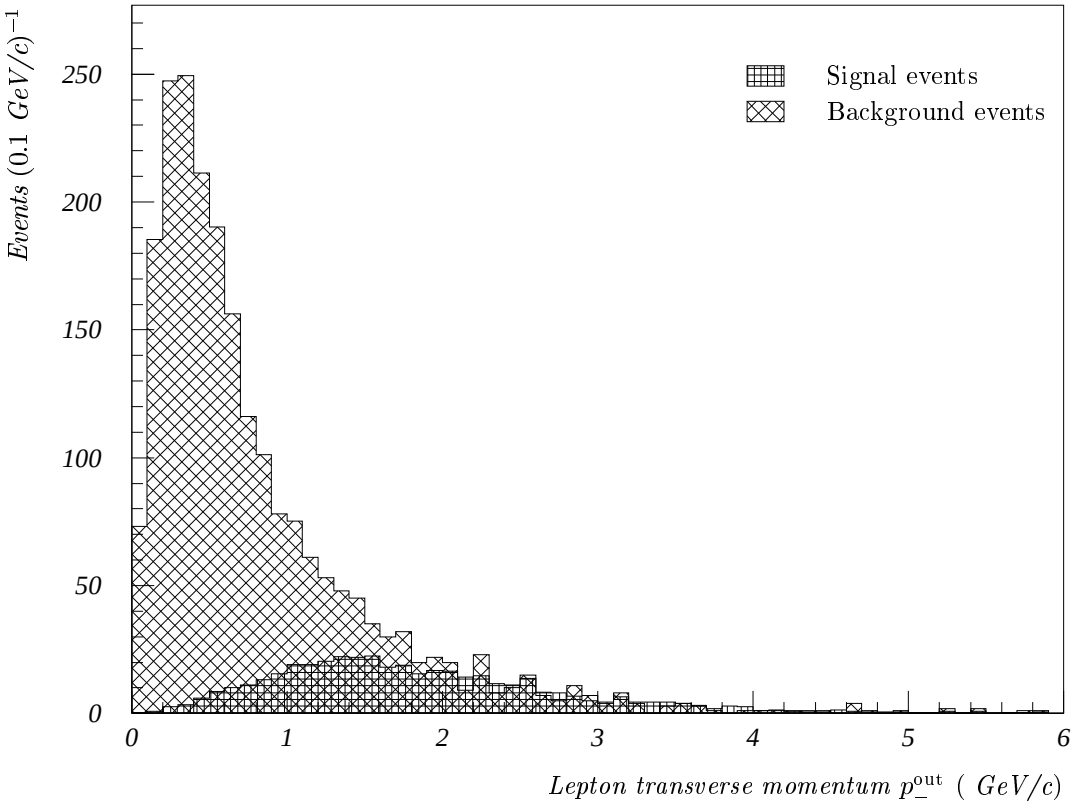


Figure 4.5: *Lepton candidate transverse momentum distribution. The background distribution is evaluated on $q\overline{q}$ simulations using events where the lepton candidate (fake or true) does not come from a* B*. The signal distribution is taken from a dedicated sample of semileptonic b decays, normalized to the statistics of the full simulation.*

Using the re-fitted momenta of the tracks, a $D_s^+$ pseudotrack is reconstructed and combined with the identified lepton to form the $B_s^0$ decay vertex. The lepton is required to have at least one associated vertex detector hit, and the $\chi^2$-probability of the $B_s^0$ decay vertex fit is again required to be above $10^{-10}$. Figures 4.9 and 4.10 show the $\chi^2$-probability and longitudinal flight length resolution of reconstructed $B_s^0$-vertices. The vertex resolution is fitted with a sum of two Gaussians. The narrow one has a mean $\mu_1 = (10.6 \pm 4.7)\mu$m and variance $\sigma_1 = 174\mu$m (72% of the sample), and the parameters of the broad Gaussian are $\mu_2 = (60.7 \pm 22.1)\mu$m and $\sigma_2 = 589\mu$m (28% of the sample). The means of the fitted distributions not being zero is a signal of a possible bias towards too long flight lengths. However, on the available statistics of Monte Carlo simulated events, the observed effect is only significant at the level of $2\sigma$ and is corrected for in the final analysis. Comparing the inset of figure 4.8 and figure 4.10 shows that the $B_s^0$-decay vertex is reconstructed with a better resolution than the $D_s^+$-vertex, and a larger fraction of the events is contained in the narrow Gaussian. This is attributed to the larger opening angle of semileptonic B-decays

Projecting the spatial difference between the reconstructed $D_s^+$ and $B_s^0$ vertices onto the direction of the partially reconstructed $D_s^+$ momentum gives an estimate of the $D_s^+$ meson flight length. The observed distribution is shown in figure 4.11 together with the expectations from Monte Carlo simulations. The comparison between data and simulated events is satisfactory. The signal $D_s^+$

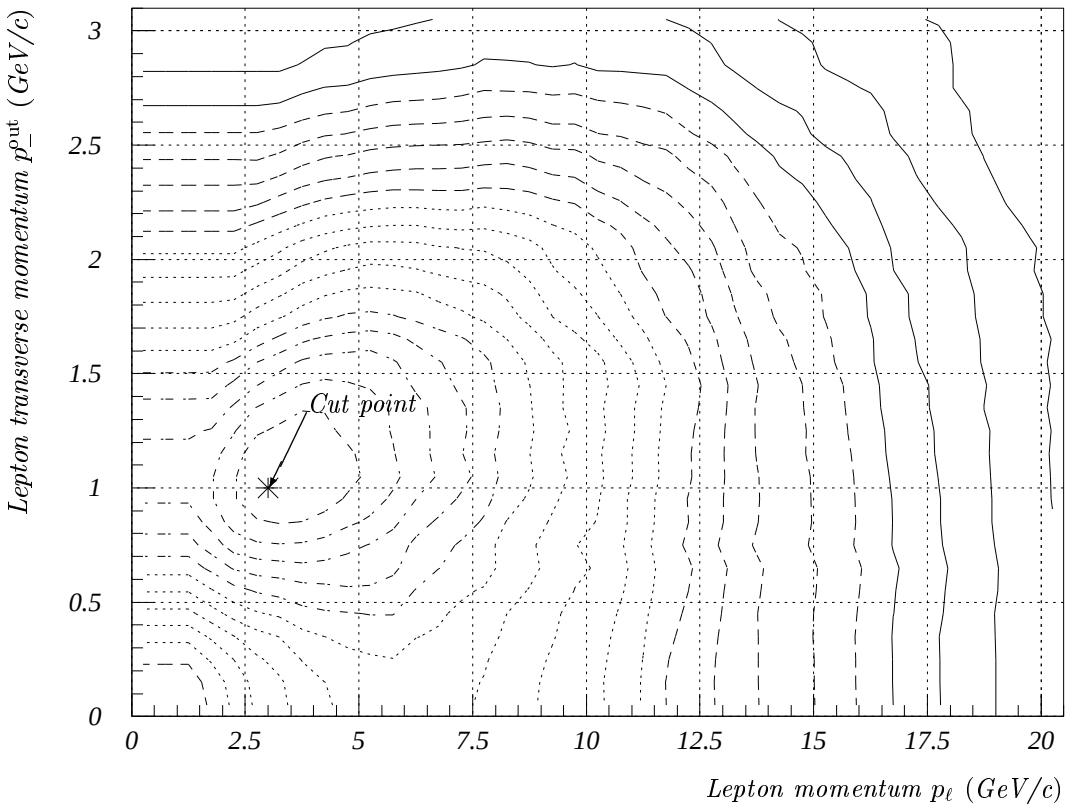


Figure 4.6: *Lepton candidate momentum and transverse momentum cut optimization. Simulated events have been used as described in the caption of figures 4.4 and 4.5. The two-dimensional significance contour is computed integrating the number of signal and background events above a given pair of cuts.*

decay length distribution is fitted with a sum of a smeared exponential and two Gaussians of mean zero. The smeared exponential accounts for 78% of the events and has a decay length of $l = 1200\mu$m and the variance of the Gaussian smearing function is $\sigma = 430\mu$m. Only a slight excess of positive flight length events is observed in the sample of reconstructed events. For this reason, a cut on $D_s^+$ flight length is hardly useful, and is omitted.

Charged tracks not used for the $B_s^0$ reconstruction are candidates for coming from the primary vertex. Using the measured track parameters and the beamspot information, a full three dimensional fit of the event primary vertex is done. The standard DELPHI procedure used to select primary vertex tracks is described as pseudo code in figure 4.12. The upper right corner inset of figure 4.13 shows the distribution of the number of tracks selected for the primary vertex fit using the algorithm described above. Data shows an excess of events with less than 7 tracks attached to the primary vertex. This is probably explained by the tracking efficiency being overestimated in the Monte Carlo simulations and the alignment of the detector not quite reaching the precision assumed in the simulations [67].

The main histogram of figure 4.13 shows the $\chi^2$-probability distribution of the primary vertex fit. As the prospective $\chi^2$-contribution of individual tracks is used to select tracks for the primary vertex, the $\chi^2$-probability distribution cannot be expected to be flat. Indeed, the distribution shown in figure 4.13 is highly distorted. However, it is reassuring that the distributions of data

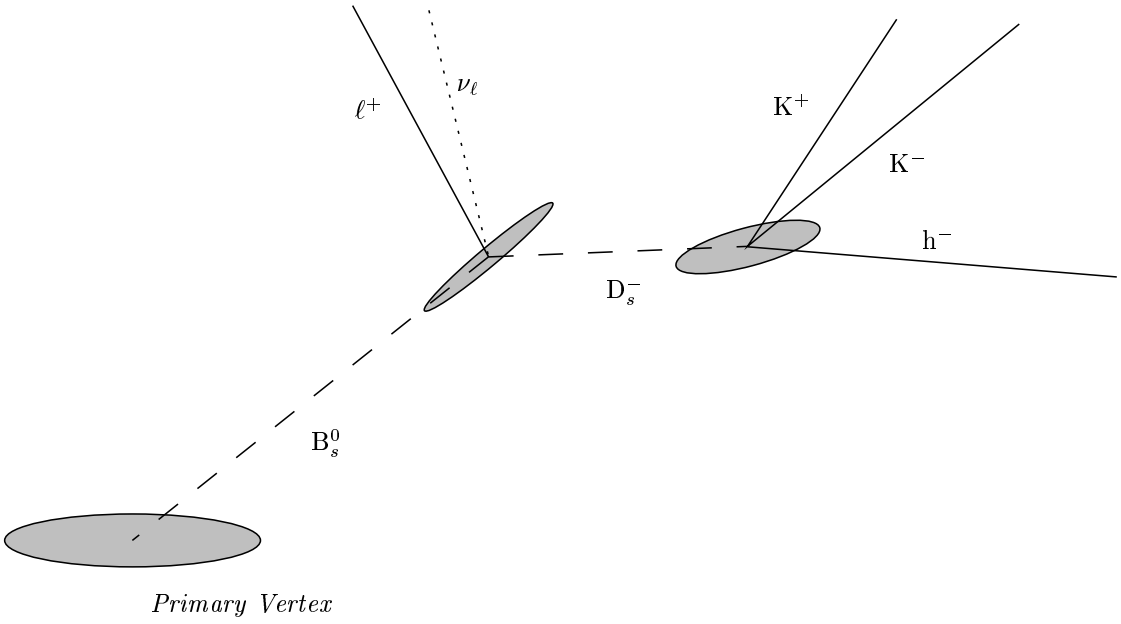


Figure 4.7: *Schematic picture of* $B_s^0$ *decay reconstruction. Charged tracks are shown as solid lines, reconstructed vertices as shaded ellipses and indirectly reconstructed particles (pseudotracks) are shown as dashed lines. The neutrino (dotted line) escapes undetected.*

| Axis | Narrow Gaussian variance ($\mu$m) | Narrow Gaussian fraction | Typical Beam Spot size ($\mu$m) |
|---|---|---|---|
| $x$ | 34.2 | 57.1% | 100 |
| $y$ | 10.2 | 98.9% | 10 |
| $z$ | 54.0 | 31.4% | 1000 |
| $r$ | 18.7 | 53.3% | - |

Table 4.2: *Primary vertex resolution components. Each projection shown in figure 4.14 is fitted with a sum of two Gaussians.*

and Monte Carlo simulations are distorted in more or less the same way.

Using simulated events, the primary vertex resolution projected along the coordinate axes as well as along the $B_s^0$ flight direction is shown in figures 4.14 a-d. The resolution along the $y$-axis is seen to be dominated by the small $y$-size of the LEP beam spot [68] whereas the $x$- and $z$-direction vertex precision get a sizeable contribution from the attached tracks. Comparing 4.14d with figure 4.10, one concludes that the flight length uncertainty will be totally dominated by the secondary vertex resolution.

Projecting the three dimensional distance between the fitted primary and secondary vertices onto the reconstructed $B_s^0$ meson flight direction gives the distribution of flight lengths shown in figure 4.15

### 4.5.2 Energy reconstruction

The energy of semileptonically decaying $B_s^0$ is shared between the identified lepton, the undetected neutrino and the recoiling hadronic system. In principle, the neutrino energy can be reconstructed using the beam constraint. Energy hermeticity is however a weak point of the DELPHI detector, and the possibility of having another neutrino in the opposite hemisphere further degrades the

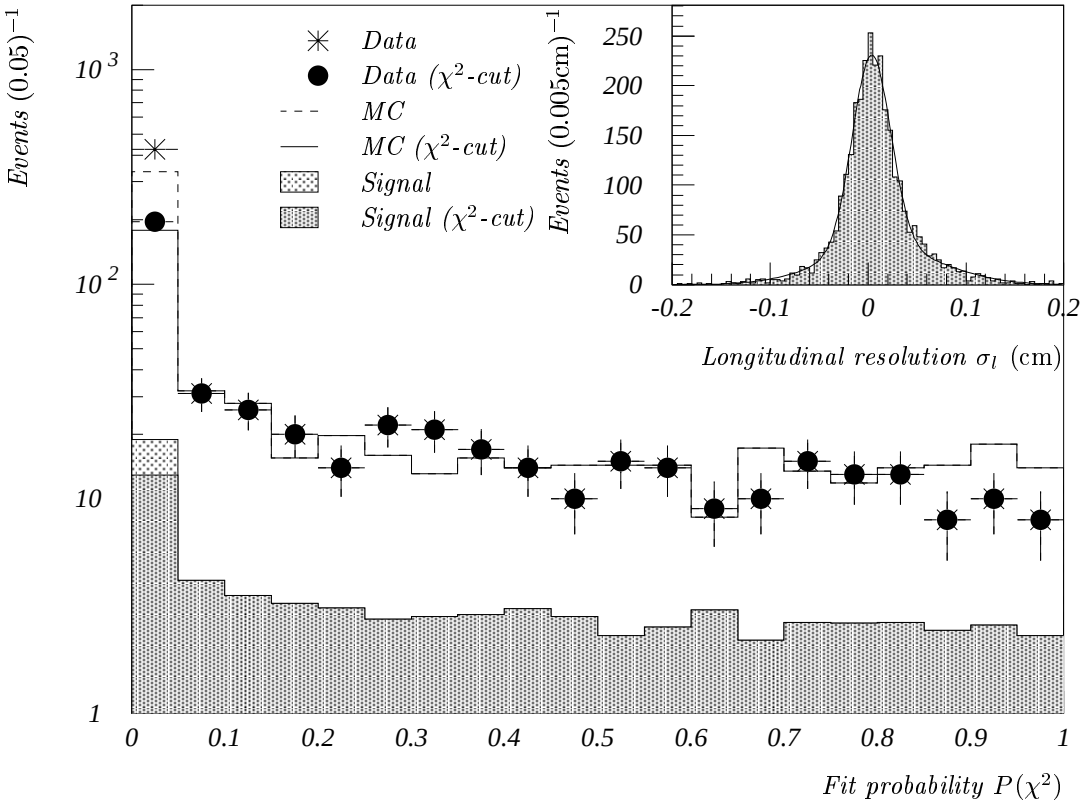


Figure 4.8: $K^+K^-h+$ *vertex probability of fit. Data events are selected as described in section 4.3, the first bin shows the difference when applying the* $\chi^2$*-probability cut. The histograms of simulated events passing the same cuts as data are normalized to correspond to the number of data events after cut. The signal histograms are normalized to the number of signal events expected in data. Shown as an inset is the longitudinal (flight direction) resolution of the* $D_s^+$ *decay vertex, as evaluated on a dedicated signal simulation sample.*

$B_s^0$ energy resolution. Having reconstructed all the $B_s^0$ decay products except the neutrino, one might combine this with the information about the flight direction from the decay vertex reconstruction. The $B_s^0$ energy and momentum represent 4 unmeasured quantities. Energy-momentum conservation, the mass of the neutrino, the mass of $B_s^0$ and the flight direction give 4 constraints. In principle, this gives a zero-constraint system that could be solved to yield the $B_s^0$ energy and momentum. Multiple difficulties do however make this impractical:

- Any undetected decay products would add to the recoiling "neutrino" system, giving it non-zero mass.
- The flight direction resolution from the geometric decay vertex reconstruction is poor.
- The zero-constraint system is second order and may give none, one or two solutions. Hence, one would have to specify ad-hoc procedures for choosing between two solutions and for what to do in the case of no solutions.

The approach chosen in the present analysis is to improve the beam constraint energy using hemisphere masses (see below) followed by a correction depending on the visible $B_s^0$ decay energy.

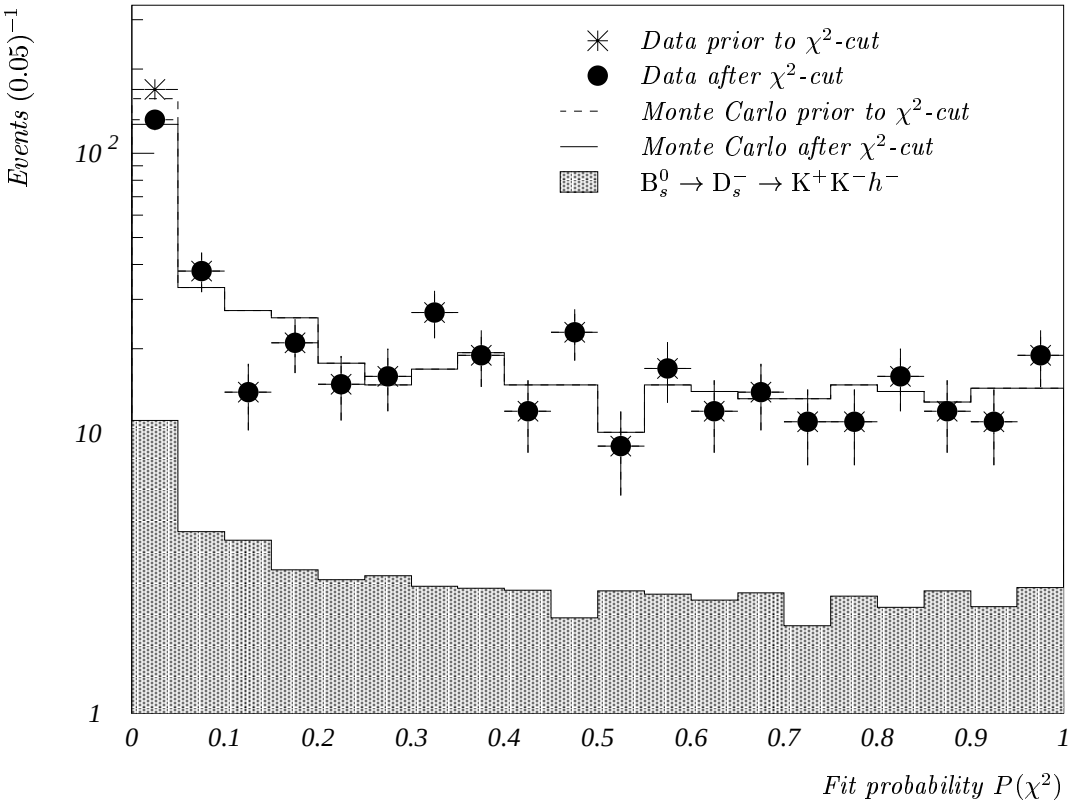


Figure 4.9: $B_s^0$-*decay vertex probability of fit. Data events are selected as described in section 4.3, the first bin shows the difference when applying the* $\chi^2$-*probability cut. The histograms of simulated events passing the same cuts as data are normalized to correspond to the number of data events after the* $\chi^2$ *cut. The signal histograms are normalized to the number of signal events expected in data.*

In semileptonic $B_s^0$ decays, the excited states $D_s^{-*}$, $D_{1s}^-$ and $D_{js}^{-*}$ ($j = 0 \cdots 2$) and the ground state $D_s^-$ are expected [60] to occur at a ratio of about 5 : 1. For convenience, the possibly excited hadronic system is labelled $D_s^{-(*)}$. The excited states decay immediately, emitting photons or $\pi^0$'s. The neutral electromagnetic energy thus created adds to the energy of gammas and $\pi^0$'s created in the subsequent $D_s^0$ decay. To reconstruct the $D_s^{-(*)}$ energy as completely as possible, neutral electromagnetic energy found in the direction of the $K^+K^-h^-$ system is added to the charged energy using the following procedure: All neutral electromagnetic showers in the hemisphere of the $B_s^0$ candidate are sorted using invariant mass of the combined charged+neutral system as the sorting key. Starting with the shower that gives the smallest increase in the invariant mass, neutral electromagnetic energy is added until the mass reaches a cut-off value of 2.11 GeV. This happens to be the mass of the $D_s^*$ vector meson [4]. However, as heavier resonances are also present, the cut-off is somewhat arbitrarily chosen. Figure 4.16 shows a comparison of energy distributions of data and Monte Carlo simulations; the total energy of the three-track system is shown in 4.16a whereas 4.16b shows the energy including associated neutral electromagnetic energy. A shift towards higher energies is clearly seen as electromagnetic energy is added. Figure 4.17 shows the difference between the $D_s^{-(*)}$ energy estimates and the true generated value. Including neutral electromagnetic energy reduces the width as well as the lower tails of the distributions.

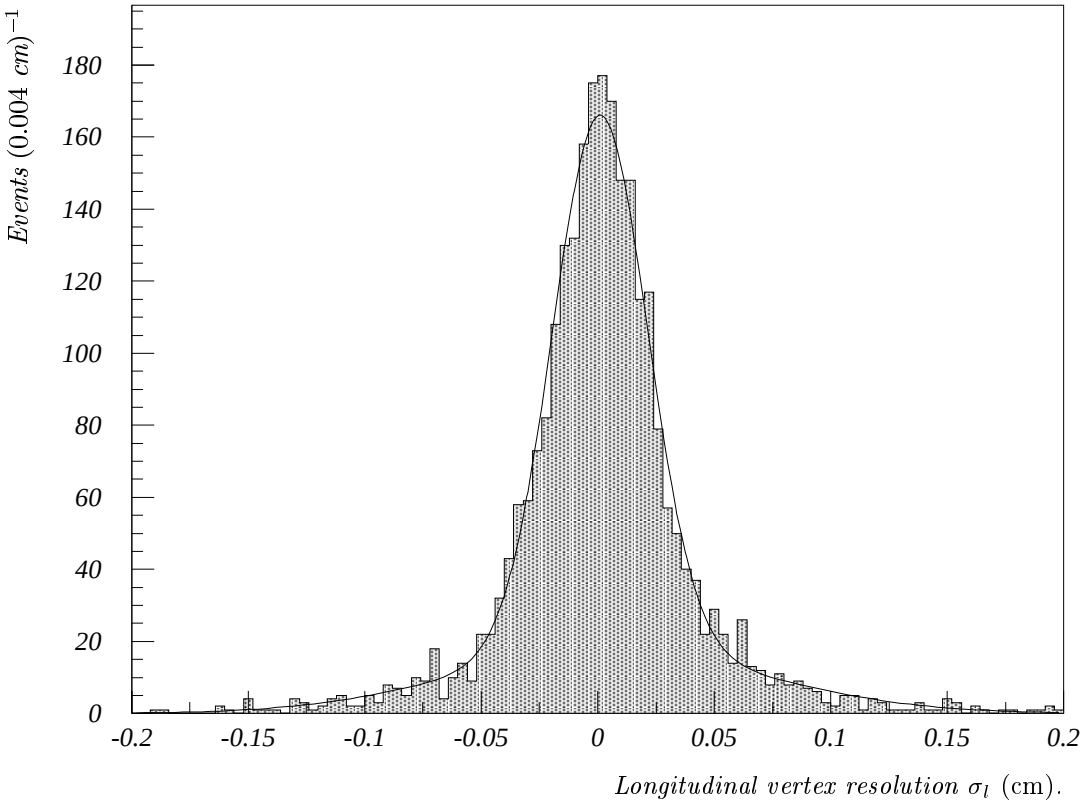


Figure 4.10: $B_s^0$-*decay vertex longitudinal resolution, evaluated on a dedicated signal simulation sample using the same selection criteria as data (section 4.3). The fitted function is a sum of two Gaussians, the narrow one has a variance of* $\sigma_l = 174\mu$m *and accounts for* 72% *of the sample.*

Employing the beam energy-momentum constraint turns out to be a successful way of determining the $B_s^0$ energy in semileptonic decays. Dividing the event in two hemispheres using the thrust axis, energy-momentum conservation gives:

$$(E_1, \vec{p}_1) + (E_2, \vec{p}_2) = (\sqrt{s}, \vec{0}) \tag{4.7}$$

Rearranging and squaring give

$$m_2^2 = s - 2\sqrt{s}E_1 + m_1^2 \tag{4.8}$$

The hemisphere energy can then be written:

$$E_1 = \frac{1}{2}\sqrt{s} + \frac{m_1^2 - m_2^2}{2\sqrt{s}} \tag{4.9}$$

Of course, the true hemisphere masses cannot be known as some energy and momentum is always carried by invisible neutrinos. However, using the visible hemisphere mass of both hemispheres, there is a tendency that the underestimation of $m_1$ is canceled by an underestimation of $m_2$. Subtracting the visible energy in the signal hemisphere gives an estimate of the neutrino energy. Alternatively, subtracting the visible fragmentation energy from the estimated hemisphere energy gives an estimate of the $B_s^0$ energy, $\hat{E}^{\mathrm{hem}}_{B_s^0}$. The distribution of the reconstructed decay energy $\hat{E}^{\mathrm{rec}}_{B_s^0}$ is shown in figure 4.18a, and figure 4.18b shows the hemisphere based $B_s^0$ energy estimate

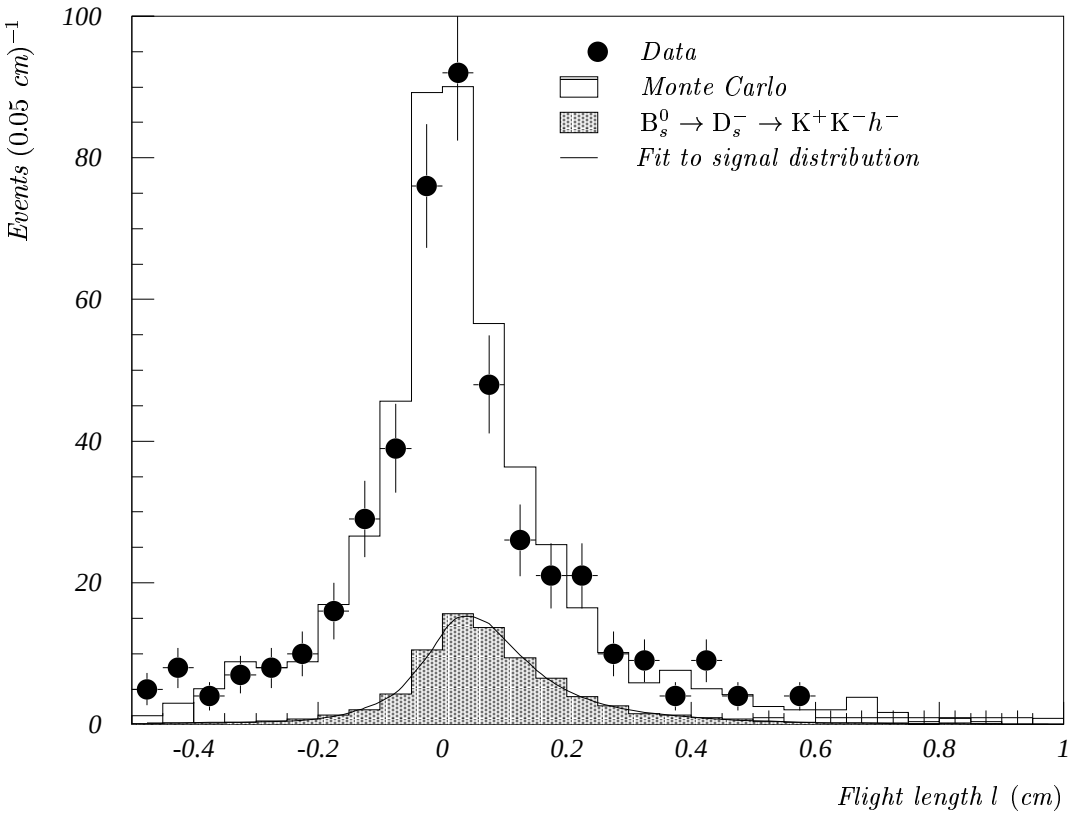


Figure 4.11: $D_s^-$ *meson candidate flight length, measured as the distance between the reconstructed* $B_s^0$ *and* $D_s^0$ *decay vertices. The excess of positive flight length events is attributed to the charm lifetime.*

$\hat{E}^{\rm hem}_{B_s^0}$. Comparing these two estimates with the simulation truth energy $E^{\rm true}_{B_s^0}$ , the resolution plots in figure 4.19 are obtained.

Noting that fluctuations of $\hat{E}^{\rm hem}_{B_s^0}$ and $\hat{E}^{\rm rec}_{B_s^0}$ are only weakly correlated through the hemisphere mass, the two estimators can be successfully combined allowing for a correction term:

$$\hat{E}_{B_s^0} = \hat{E}^{\rm hem}_{B_s^0} + a + b \cdot \hat{E}^{\rm rec}_{B_s^0} \tag{4.10}$$

The parameterization constants $a$ and $b$ are determined using the least squares method on the Monte Carlo simulated signal sample:

$$\begin{aligned} a &= (-6.65 \pm 0.22)\ \text{GeV/c} \\ b &= 0.165 \pm 0.008, \end{aligned}$$

with a correlation coefficient of $-0.965$. This is essentially the same as fitting a straight line to the profile plot of the deviation of the hemisphere estimator versus reconstructed decay energy, as shown in figure 4.20. The distribution of the corrected $B_s^0$ energy estimator is shown in figure 4.21a. Figures 4.21b and 4.21c show the resulting absolute and relative resolutions of the estimator. Table 4.3 compares the mean of the energy distributions on data and Monte Carlo simulations. The errors shown are approximated using the RMS of the distribution and the number of events in the histograms.

```
      FUNCTION PVFIT(available_tracks)
*     Initial impact parameter cut
      DO I=1,NTRACK
         compute_impact_parameters_wrt_beamspot
         IF (IMP(I).GT.0.25.OR.IMZ(I).GT.5.0) THEN
            USE_TRACK(I)=.FALSE.
         ENDIF
      ENDDO
*     Preliminary vertex fit
      CALL beamspot_constrained_vertex_fit
*     Second impact parameter cut
      DO I=1,NTRACK
         compute_impact_parameters_wrt_fitted_vertex
         IF (IMP(I).GT.0.05.OR.IMZ(I).GT.1.0) THEN
            USE_TRACK(I)=.FALSE.
         ENDIF
      ENDDO
*     Iterate until good fit
      DO WHILE tracks_left
         I=find_track_with_largest_chi2_contribution
         IF (CHI2(I).LT.5.0) RETURN good_vertex
         USE_TRACK(I)=.FALSE.
         CALL beamspot_constrained_vertex_fit
      ENDDO
      RETURN bad_vertex
      END
```

Figure 4.12: *Simplified pseudo code for primary vertex fit. Pseudo code is not a listing of compilable FORTRAN code but uses simplified programming constructs (upper case words) as a terse language for algorithm description.*

| Quantity | Data average ( GeV) | Monte Carlo average ( GeV) |
|---|---|---|
| $\langle E_{\mathrm{D}_s^+} \rangle$ | $11.5 \pm 0.3$ | $11.1 \pm 0.2$ |
| $\langle E_{\mathrm{D}_s^{-(*)}} \rangle$ | $13.5 \pm 0.3$ | $13.0 \pm 0.2$ |
| $\langle \hat{E}^{\mathrm{rec}}_{\mathrm{B}_s^0} \rangle$ | $23.3 \pm 0.3$ | $22.7 \pm 0.2$ |
| $\langle \hat{E}^{\mathrm{hem}}_{\mathrm{B}_s^0} \rangle$ | $33.6 \pm 0.3$ | $33.4 \pm 0.2$ |
| $\langle \hat{E}_{\mathrm{B}_s^0} \rangle$ | $30.8 \pm 0.3$ | $30.5 \pm 0.2$ |

Table 4.3: *Comparison of data and simulated energy distributions.*

## 4.6 Production charge tag

As $b$ quarks are always produced in $b\overline{b}$ pairs, a $\overline{b}$ quark fragmenting to a $\mathrm{B}_s^0$ will be accompanied by a $b$ quark in the opposite hemisphere. Unfortunately, requiring an unambiguously reconstructed charged $b$-hadron for tagging would ruin statistics completely. To improve the situation, several event observables correlated with the charge of the initial $b$ quark are combined to provide an optimal tag. Observables from the tagging hemisphere give most of the information, but it should be noted that fragmentation tracks in the signal hemisphere contribute too. A summary of the observables are shown in table 4.4 and the distribution of each tag is shown in figure 4.22

Charged tracks not used in the $\mathrm{B}_s^0$ reconstruction are available for the charge tagging algorithm. For observables that summarize the charge of several tracks, a momentum weighted algorithm is

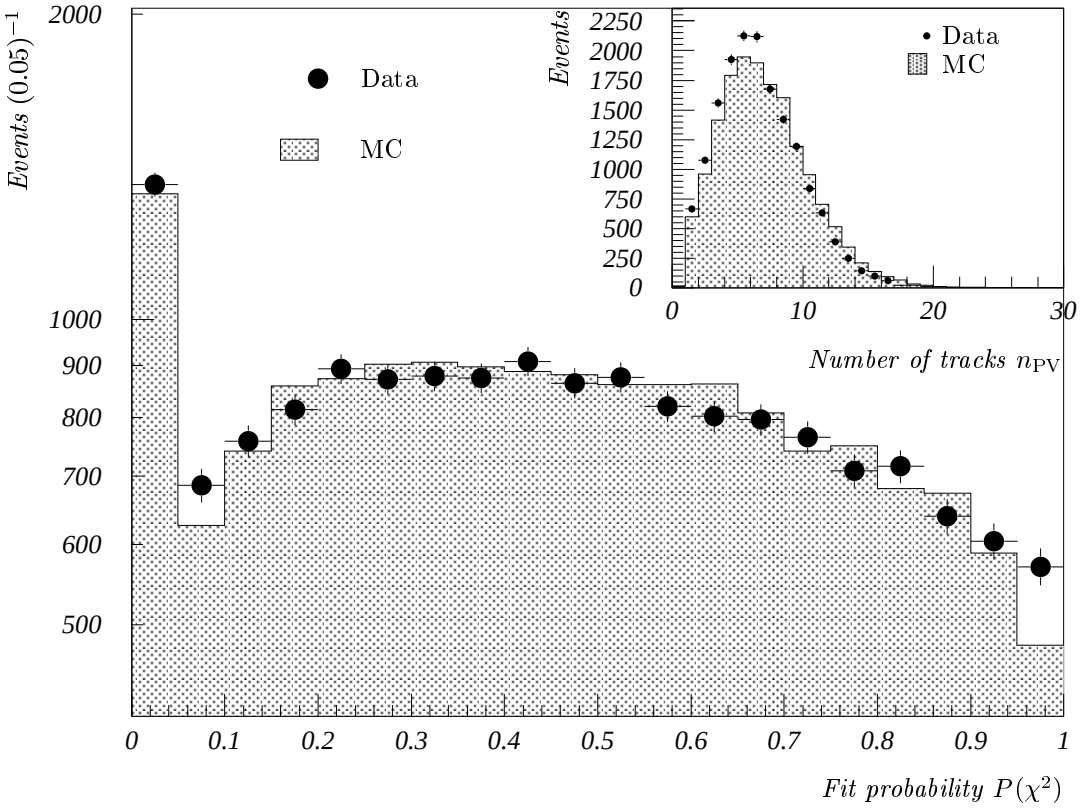


Figure 4.13: *Primary vertex fit probability. The simulated histogram is normalized to the number of events found in data. As the $\chi^2$ of the fit is used to reject outliers in the primary vertex reconstruction algorithm, the distribution of the fit probability cannot be expected to be flat. The inset shows the distribution of the number of tracks assigned to the primary vertex.*

used [25]:

$$Q_{\mathrm{jet}} = \frac{\sum_{\mathrm{jet}} q_i p_i^{0.6}}{\sum_{\mathrm{jet}} p_i^{0.6}} \tag{4.11}$$

For single-track observables, the observable taken is the track rapidity with respect to the thrust axis multiplied with its charge. Identified lepton is an exception: In this case, the product of the charge and the transverse momentum is used as the observable.

The 12 observables $x_i$ are combined using a linear Fisher discriminant [69] [70]:

$$f = \sum_i a_i x_i \tag{4.12}$$

The coefficients $a_i$ are determined using the means and covariances of the observables:

$$a_i = \sum_j \left( \langle x_i x_j \rangle_{\overline{b}} + \langle x_i x_j \rangle_b \right)^{-1} \left( \langle x_j \rangle_{\overline{b}} - \langle x_j \rangle_b \right) \tag{4.13}$$

The subscripts $\overline{b}$ and $b$ mean that the averages the $\langle \cdots \rangle$ are taken over events with the tagging hemisphere containing the corresponding quark. The sign given in equation 4.13 implies that the

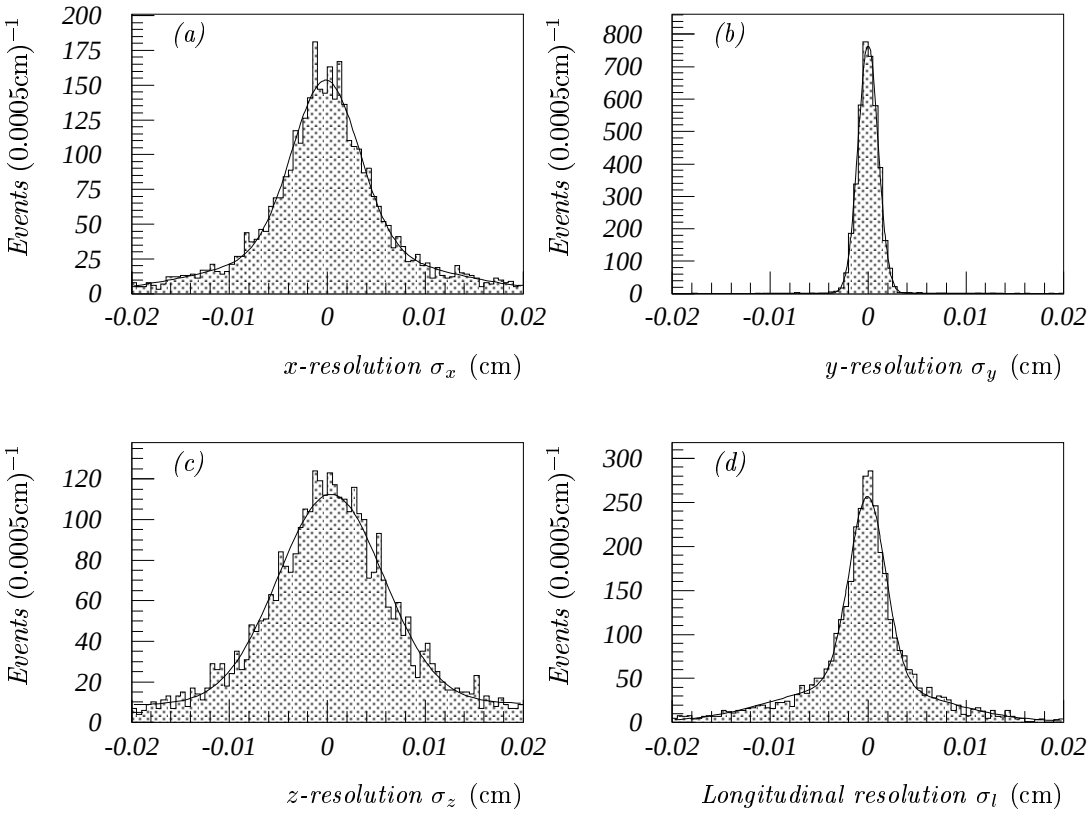


Figure 4.14: *Primary vertex resolution, projected along each coordinate axis as well as along the* $B_s^0$ *flight direction. The y-direction resolution is dominated by the small vertical size of the LEP beamspot. The fitted functions are sums of two Gaussians, the fit results are summarized in table 4.2*

sign of the Fisher discriminant is expected to be the same as the sign of the $b$ quark in the tagging hemisphere.

In order to control the overfitting problem, the Monte Carlo simulated signal sample is split in two parts. One half is used to determine the coefficients, the other to evaluate the performance of the discriminant.

Due to the inclusive nature of the present analysis, some $B_s^0$ decay products might be missed in the reconstruction and then contribute to the tag observables. If this happens, the tag performance might depend on whether the $B_s^0$ has mixed or not. Thus, it is necessary to study the charge correlation, i.e. the product of the Fisher discriminant and the charge of the identified lepton. For events with no $B_s^0$ mixing, a value less than 0 is expected, and a value greater than 0 is expected for events where the $B_s^0$ meson mix before it decays. A summary of the tagging performance is given in table 4.5. The distribution of the charge tag and of the charge correlation is shown in figure 4.23

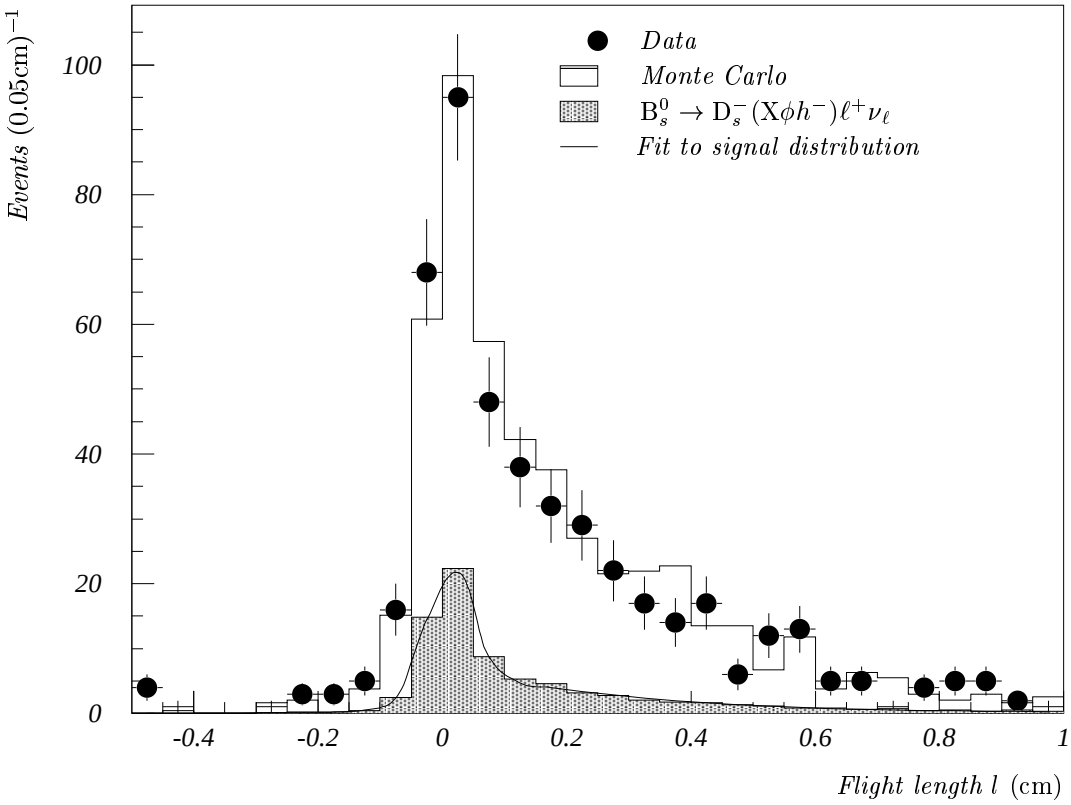


Figure 4.15: $B_s^0$ *meson candidate flight length, reconstructed as the three dimensional distance between the primary vertex and the* $B_s^0$ *decay vertex projected onto the direction of the reconstructed decay tracks.*

## 4.7 A non-linear multidimensional discriminant for background suppression

The purpose of using a multidimensional discriminant is to enhance the signal to background separation. The idea is to combine a few features of the events into a single variable that discriminates between signal and background. To be specific, suppose the $n$ observables $x_i$, $i = 1 \ldots n$ are distributed according to $p(x_1 \ldots x_n)$ in the signal and $q(x_i \ldots x_n)$ in the background. Using these distributions, a discriminating statistic can be built using the likelihood ratio [71] $\frac{p(x_1 \ldots x_n)}{q(x_1 \ldots x_n)}$ (or $\log \frac{p(x_1 \ldots x_n)}{q(x_1 \ldots x_n)}$). When cutting in the discriminant to suppress background, the likelihood ratio is optimal in the following sense: For a given signal efficiency, the background acceptance is minimal and vice versa. Alternatively, $p(x_1 \ldots x_n)$ and $q(x_1 \ldots x_n)$ can be included in the full likelihood function of the analysis. Conceptually, this looks like using a per-event signal fraction. In the present analysis, there is no way of having a priori knowledge of the multidimensional distributions $p$ and $q$, they can only be sampled using Monte Carlo simulations. With more than 2 observables, the number of simulated events needed makes mapping out the full multidimensional space a forbidding task. Instead, the correlations between observables are discarded and a discriminant $X$ is built from the one dimensional distributions of $p_i(x_i)$ and $q_i(x_i)$: $X = \prod_i \frac{p_i(x_i)}{q_i(x_i)}$. Carefully selecting observables $x_i$ with minimal correlations, this gives an approximation of the

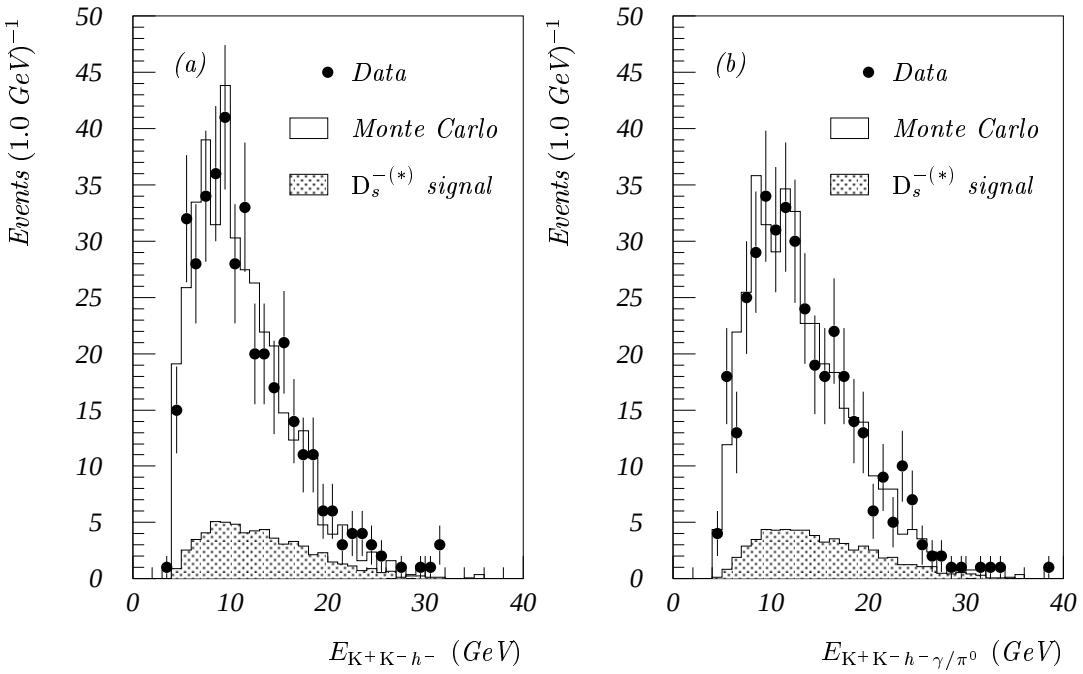


Figure 4.16: *Reconstructed* $D_s^{-(*)}$ *energy distributions. Panel (a) shows the reconstructed charged track energy, panel (b) shows charged+neutral energy. The algorithm for associating neutral electromagnetic clusters is described in section 4.5.2.*

| Hemisphere | | Mnemonic name | Efficiency (%) | Purity (%) |
|---|---|---|---|---|
| Reconstruction hemisphere | All tracks | samejch | $89.8 \pm 0.5$ | $62.4 \pm 0.8$ |
| | PV-tracks | samepch | $85.7 \pm 0.6$ | $61.4 \pm 0.8$ |
| | Identified proton | samepst | $16.6 \pm 0.6$ | $54.2 \pm 2.0$ |
| | Identified kaon | samekst | $27.5 \pm 0.7$ | $71.3 \pm 1.4$ |
| Both | Thrust axis polar angle | oppocth | 100 | $54.2 \pm 0.8$ |
| Tagging hemisphere | All tracks | oppojch | 100 | $65.4 \pm 0.8$ |
| | PV-tracks | oppopch | $98.5 \pm 0.2$ | $57.6 \pm 0.8$ |
| | SV-tracks | opposch | $92.3 \pm 0.4$ | $57.7 \pm 0.8$ |
| | Identified proton | oppopst | $47.2 \pm 0.8$ | $58.3 \pm 1.2$ |
| | Identified kaon | oppokst | $60.5 \pm 0.8$ | $60.1 \pm 1.0$ |
| | Identified lepton | oppolpt | $29.8 \pm 0.7$ | $60.3 \pm 1.4$ |
| | Identified lambda | oppolam | $4.1 \pm 0.3$ | $52.2 \pm 4.0$ |

Table 4.4: Charge tag variables

full likelihood ratio. However, the quasi likelihoods $\prod_i p_i(x_i)$ and $\prod_i q_i(x_i)$ cannot be directly included in the likelihood analysis, as this would not give a true likelihood function. This problem is circumvented by studying the quasi likelihood ratio $X$ and using its distributions on signal and background, $P(X)$ and $Q(X)$, in the full likelihood.

The following variables are selected for the purpose of building the discriminant:

- The transverse momentum of the identified lepton

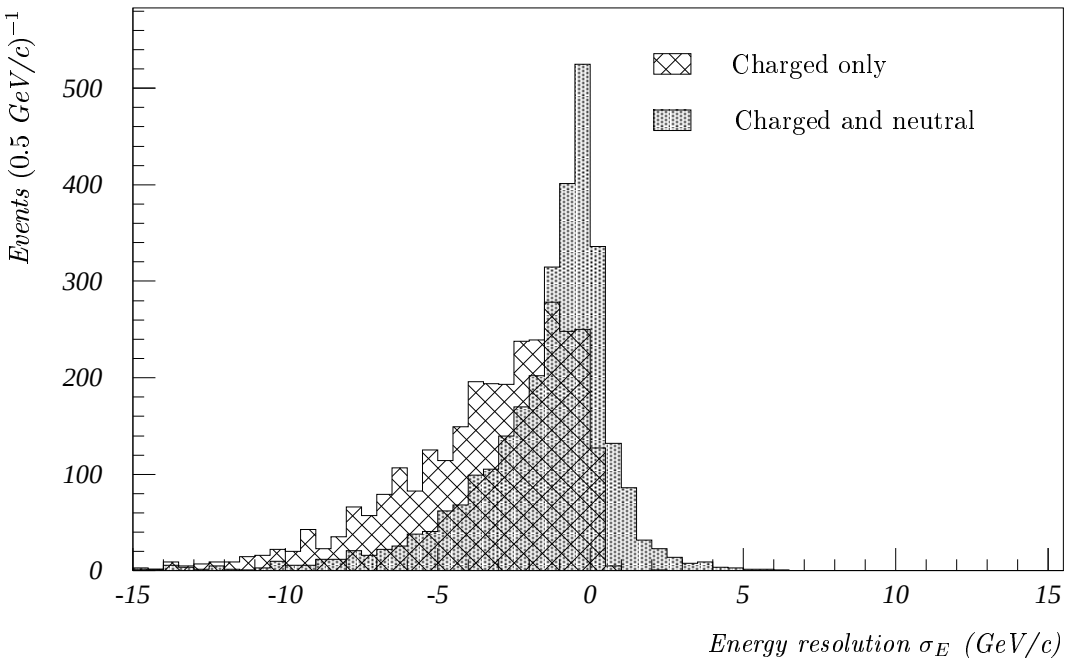


Figure 4.17: *Reconstructed* $D_s^{-(*)}$ *energy resolution. Adding neutral electromagnetic clusters to the reconstructed charged track energy clearly improves* $D_s^{-(*)}$ *energy resolution.*

| Events | Purity (%) |
|---|---|
| All | $70.1 \pm 0.7$ |
| Mixed | $67.8 \pm 1.5$ |
| Unmixed | $73.0 \pm 1.4$ |

Table 4.5: Tagging performance

- The b-tag of the opposite (non-signal) hemisphere
- The invariant mass of the $K^+K^-$-pair
- The invariant mass of the $K^+K^-h^-$ vertex

With the exception of the $K^+K^-$-mass, simulated signal and background events are histogrammed and the distributions are parameterized by means of empirical functions as shown in figure 4.24. For the mass of the $K^+K^-$-pair, the fit of signal and background to data shown in figure 5.2 is used. Figure 4.25 shows the distribution of the combined discriminant for the selected data events, compared to the distribution found in Monte Carlo simulated events. Also shown is the expected distribution of the $B_s^0$-signal.

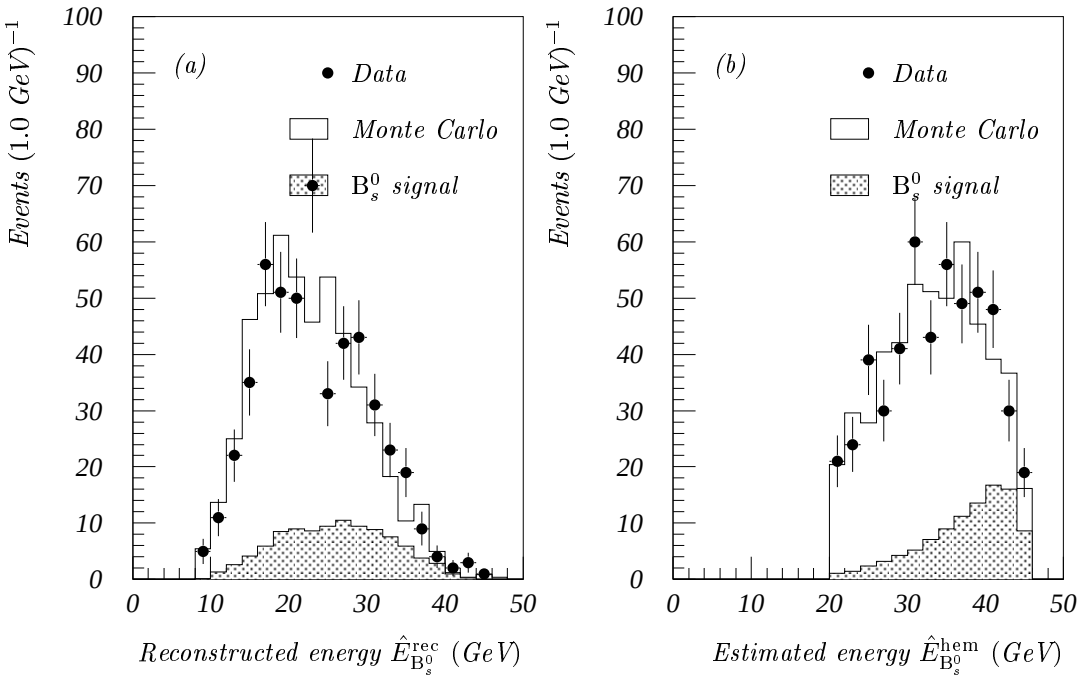


Figure 4.18: *Preliminary* $B_s^0$ *energy estimators. Panel (a) shows* $\hat{E}^{rec}_{B_s^0}$*, the visible decay energy. Panel (b) shows* $\hat{E}^{hem}_{B_s^0}$ *(GeV) the b meson energy estimated using the beam constraint and hemisphere masses as described in the text.*

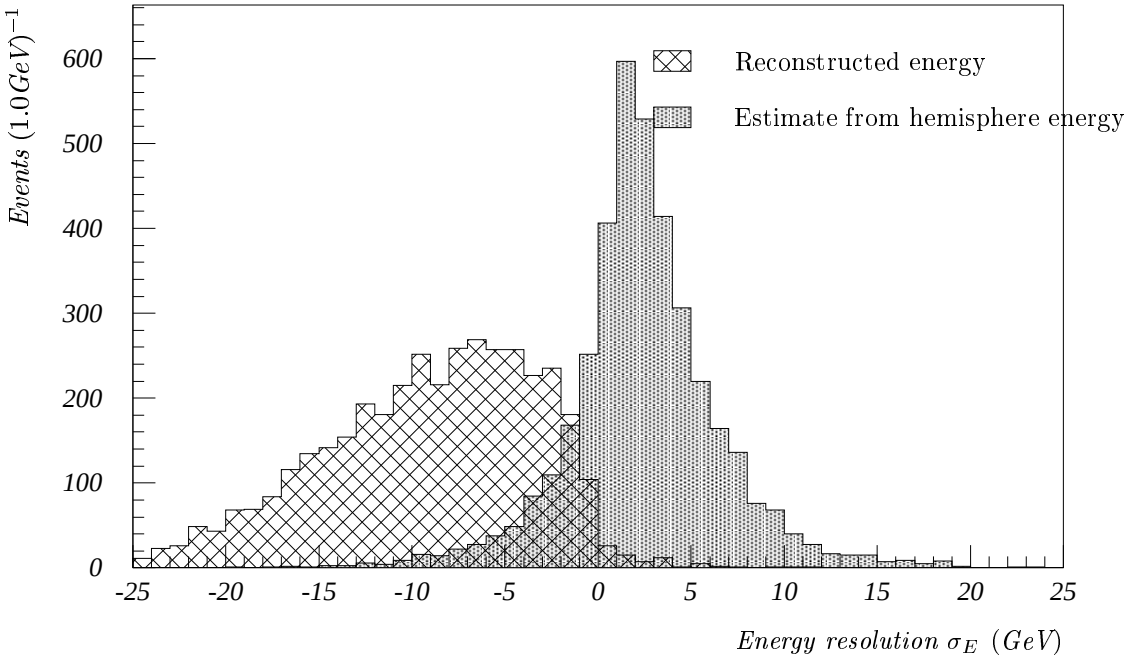


Figure 4.19: *Resolutions of the $B_s^0$ energy estimators. While the hemisphere based estimate is still biased, its variance clearly is much smaller than the variance of the reconstructed energy.*

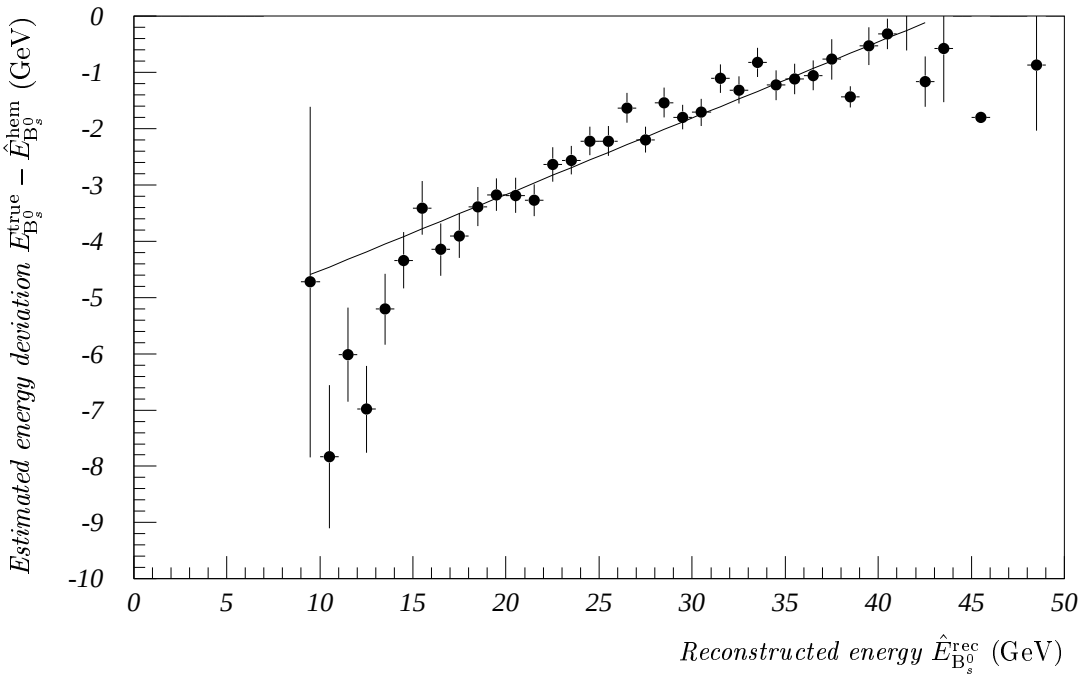


Figure 4.20: $B_s^0$ *energy estimator parameterization. The fitted straight line gives a correction to the hemisphere based energy estimate that depends linearly on the reconstructed energy*

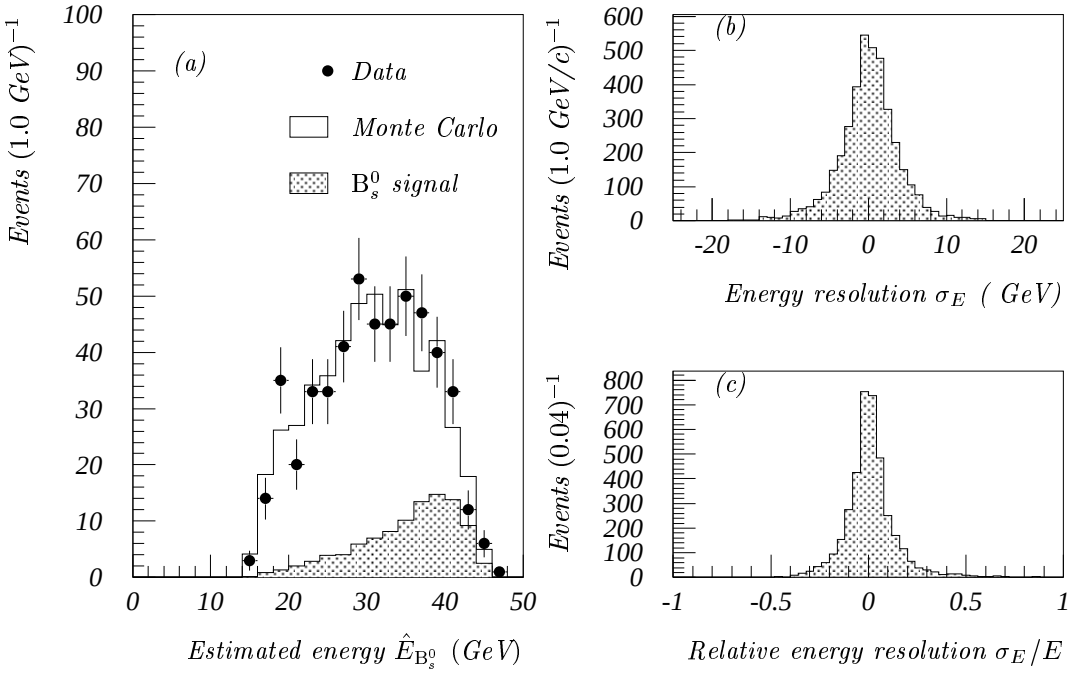


Figure 4.21: *Corrected* $B_s^0$ *energy estimator. Panel (a) shows the data versus simulation comparison together with the signal expectation. Panel (b) shows the absolute energy resolution, computed as the difference between the estimated and the true energy. The relative energy resolution shown in (c) is directly related to the time dependent part of the decay time resolution introduced in section 5.1.3*

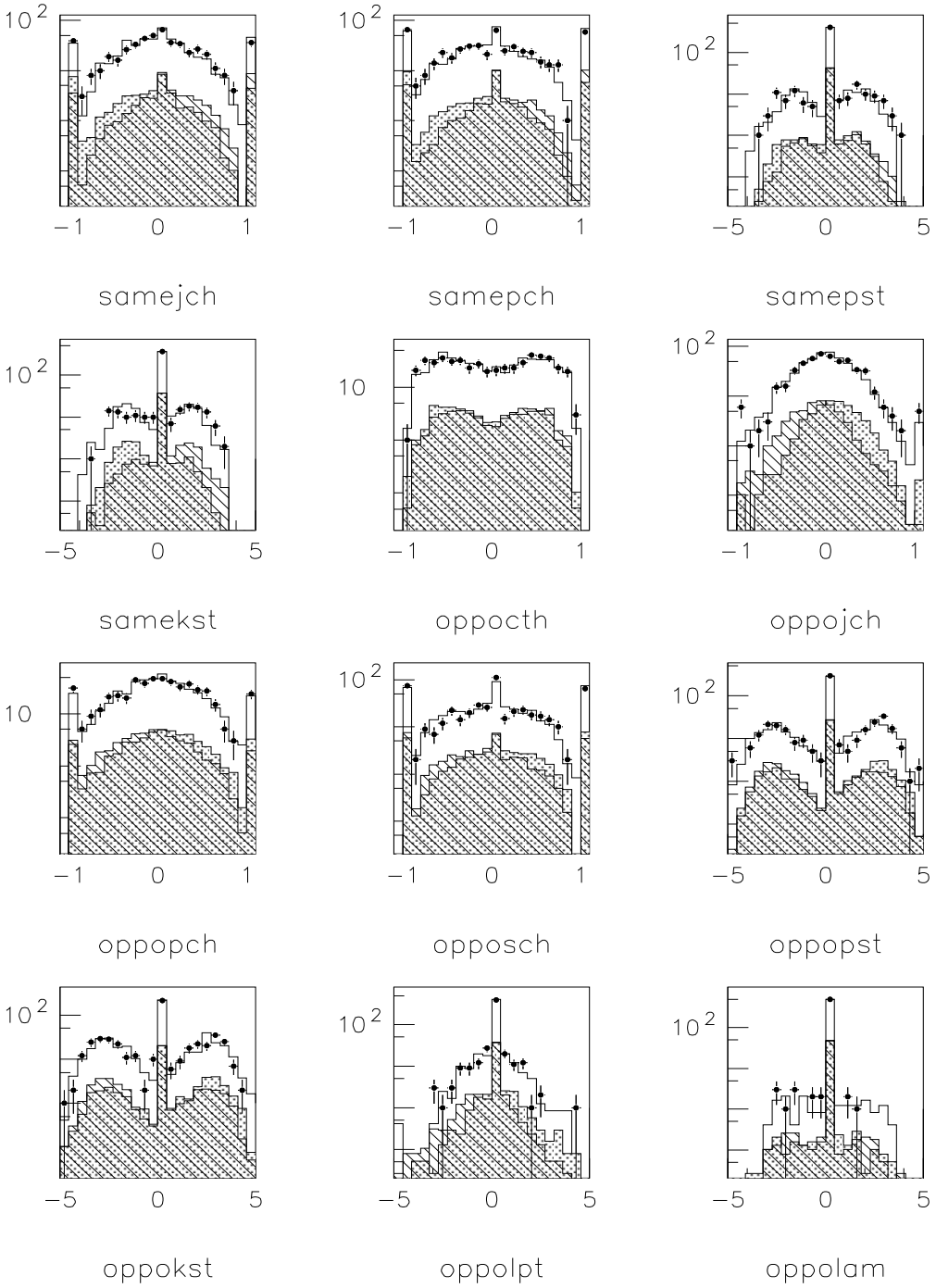


Figure 4.22: *Charge tag variables. The meaning of the mnemonic labels are explained in table 4.4, for a legend of the symbols used see figure 4.23. Each histogram shows the distribution found in data compared to simulated events, as well as the asymmetry of each tag variable depending on the production flavor of the b quark.*

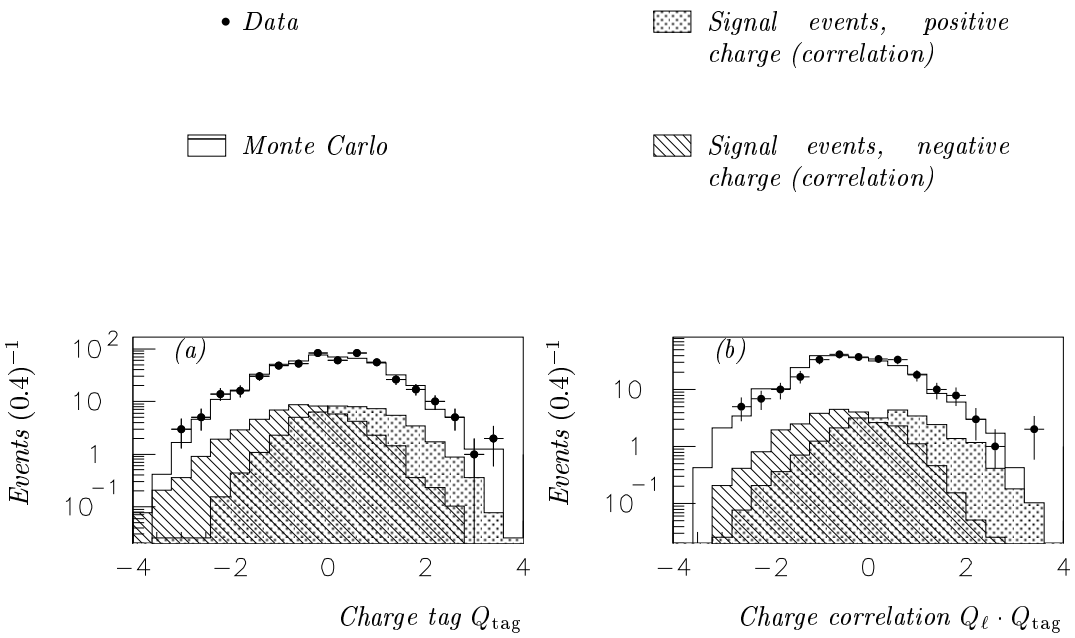


Figure 4.23: *Charge tag and charge correlation distribution. Panel (a) shows the charge tag which by construction positively correlated to the charge of the b quark opposite to the reconstructed* $B_s^0$. *Panel (b) shows the charge correlation, that is the product of the charge tag and the charge of the lepton candidate. The charge correlation is positive for mixed events and negative for unmixed events.*

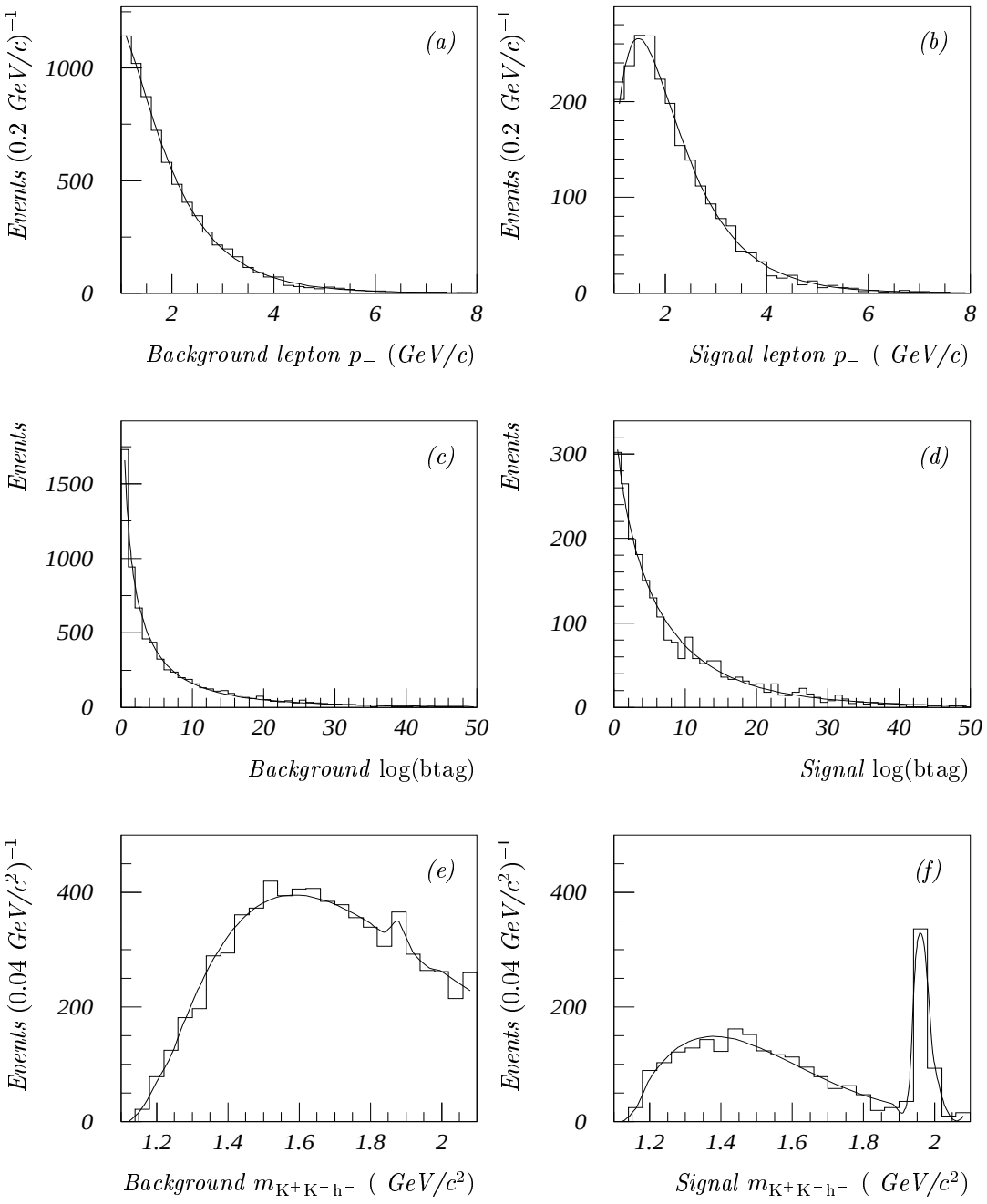


Figure 4.24: *Parameterizations for the multidimensional discriminant. The background distributions (left column) are generated using $q\overline{q}$ simulated events, retaining only non-$B_s^0$ events. The signal distributions (right column) come from a dedicated simulation sample. Both samples are subject to the signal selection cuts, except that the $\phi$ invariant mass cut is not applied. The fitted functions are similar to the background description used in section 5.1.1, with Gaussians added to describe the peaks in panels (e) and (f).*

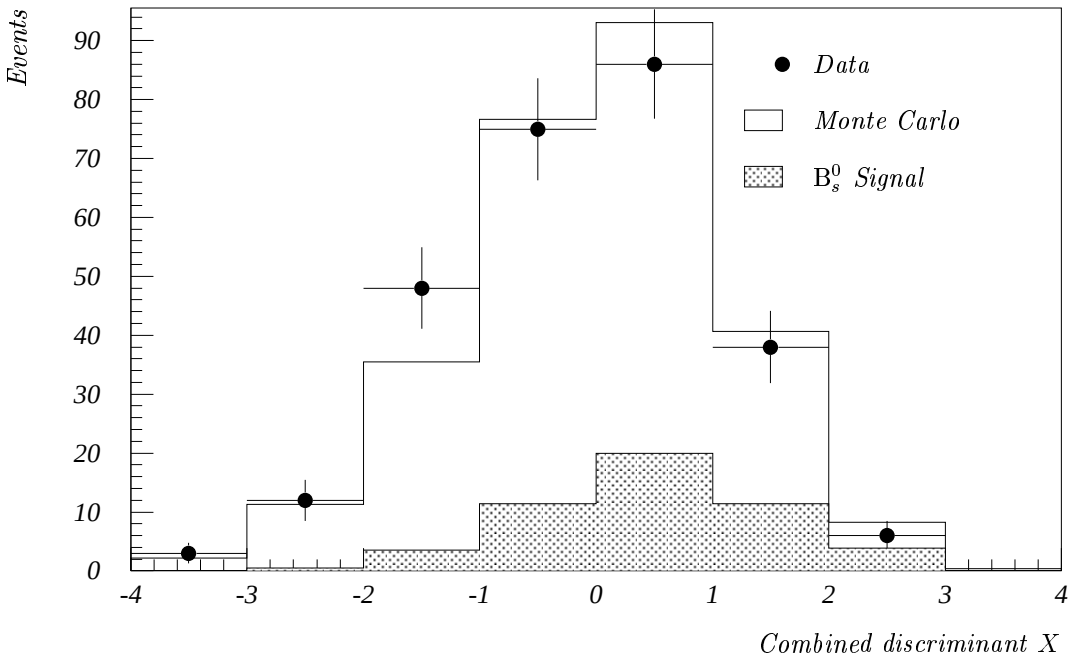


Figure 4.25: *Combined multidimensional discriminant. The distribution found in data is compared to the $q\overline{q}$ simulation distribution. Also shown is the expected signal distribution, generated from a dedicated signal simulation sample.*

# Chapter 5

# Statistical analysis: Extracting parameters and setting limits.

## 5.1 Likelihood function composition

The starting point for building the analysis' likelihood function is the theoretical decay time distribution for unmixed and mixed signal events:

$$p(t;\tau,\Delta\Gamma,\Delta m) = \frac{1-(\tau\Delta\Gamma)^2}{4\tau}e^{-t/\tau}\left(e^{+\Delta\Gamma t}+e^{-\Delta\Gamma t}\pm 2\cos\Delta m t\right) \tag{5.1}$$

In order to extract physical parameters from the measured decay times and charge correlations, accurate descriptions of the sample composition as well as the expected distribution of each background constituent are needed. For the purpose of building the likelihood function, the event sample is classified according to the following criteria:

- True $\phi$-events versus combinatorial background
- True leptons versus hadrons misidentified to be leptons
- B-decay lepton candidates versus lepton candidates from fragmentation and non-$b\overline{b}$-events
- Leptons from semileptonic $b$-decays versus leptons produced in $\overline{b}\to c\overline{c}\to\ell^{\pm}$ and $\overline{b}\to\tau\to\ell^{+}$ cascade decays.
- The flavor of the $b$-hadron ($\mathrm{B}^0$, $\mathrm{B}^+$, $\mathrm{B}_s^0$, $\Lambda_b$)

Building the likelihood-description in full detail, allowing all possible combinations of true or fake, direct or cascade, $\phi$ or combinatorial etc, would require a large number of fractions to describe the signal composition and a large number of likelihood functions to describe the expected distribution of each event class. To obtain a simple organization of the possible combinations, a hierarchical classification is needed. The hierarchy used for this purpose is sketched in figure 5.1.

In addition to the classification based on the physical contents of the selected sample, the events are split into two parts based on each event's expected resolution: The fit error on the $b$-hadron flight length projected along the flight direction, $\sigma_l$, and the combined $\chi^2$-probability of the primary and secondary vertex fits, $p_{\chi^2}$, are computed. Events that satisfy

$$\sigma_l \quad \leq \quad 300\mu\mathrm{m} \tag{5.2}$$

and

$$p_{\chi^2} \quad > \quad 2\% \tag{5.3}$$

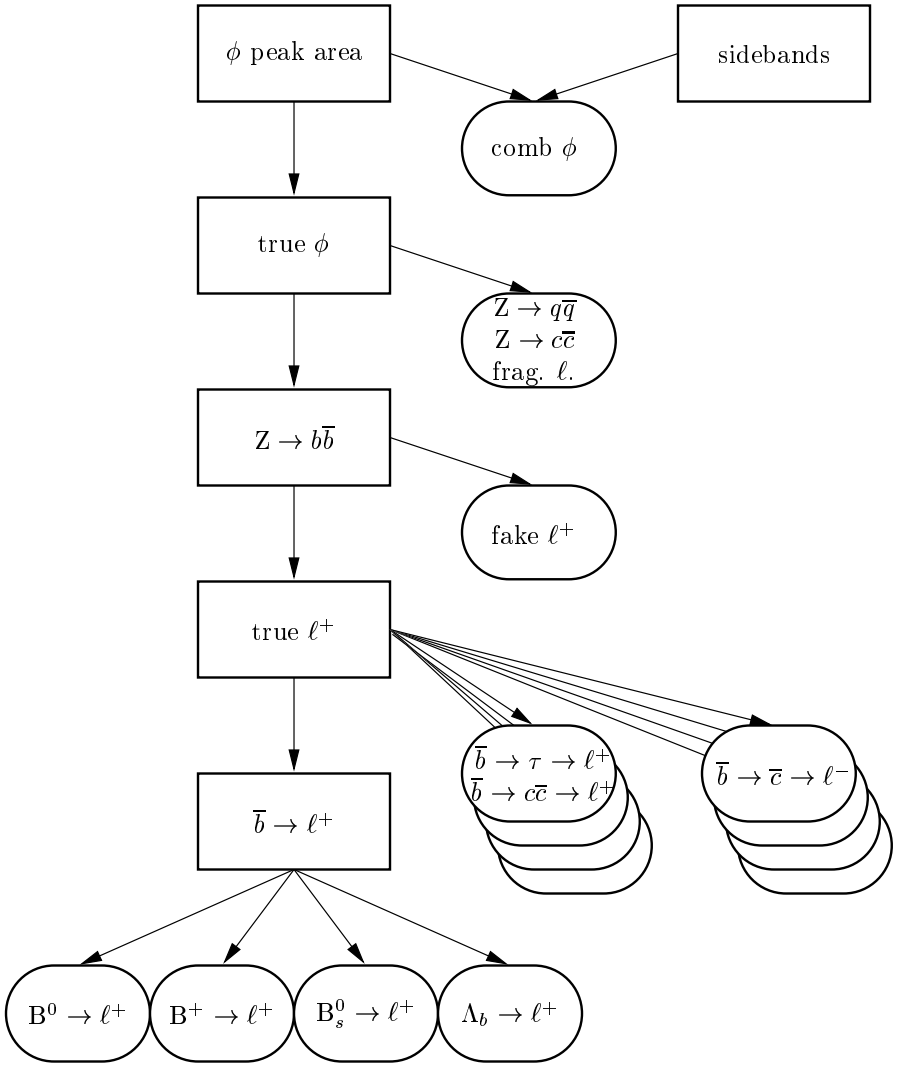


Figure 5.1: *Likelihood description hierarchy. Each oval represents a likelihood function to be parameterized or inferred from the theoretical distribution, and each arrow represents an event fraction to be computed. The stacks of ovals seen under the ones labelled* $\overline{b} \to c \to \ell^-$ *and* $\overline{b} \to \tau \to \ell^+$ *represent the possibilities of having cascade decays starting with* $\mathrm{B}^0$, $\mathrm{B}^-$, $\mathrm{B}_s^0$ *and* $\Lambda_b$.

are labelled as "good", the rest "bad". As will be shown below, the "good" events do indeed have a better expected resolution and a higher signal fraction than the "bad" events. At the top level, the likelihood function can thus be written as a product over "good" and "bad" events:

$$\begin{aligned} \mathcal{L}(\tau, \Delta\Gamma, \Delta m) &= \mathcal{L}^{\mathrm{good}}(\tau, \Delta\Gamma, \Delta m) \times \mathcal{L}^{\mathrm{bad}}(\tau, \Delta\Gamma, \Delta m) \\ &= \prod_{\mathrm{good}} p^{\mathrm{good}}(t_i, c_i; \tau, \Delta\Gamma, \Delta m) \times \prod_{\mathrm{bad}} p^{\mathrm{bad}}(t_i, c_i; \tau, \Delta\Gamma, \Delta m) \end{aligned} \tag{5.4}$$

(For clarity, physical parameters $(\tau, \Delta\Gamma, \Delta m)$ as well as probability function arguments $(t_i, c_i)$ might be omitted from here on.)

### 5.1.1 True $\phi \to \mathrm{K}^+K^-$ and combinatorial background

Using the measured momenta of the $\phi \to \mathrm{K}^+\mathrm{K}^-$ candidate tracks and assuming $\mathrm{K}^\pm$ masses (0.4937 GeV), the invariant mass of the pair is calculated. The resulting spectrum is shown in figure 5.2. This invariant mass spectrum is fitted using the Breit-Wigner distribution for the $\phi$

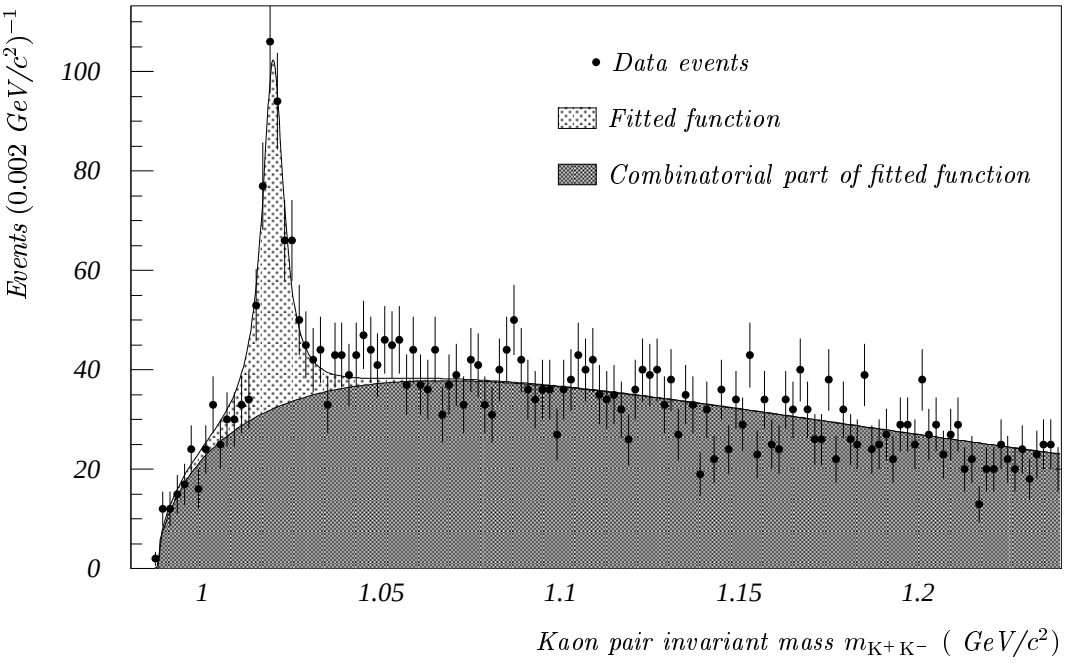


Figure 5.2: *$\phi$ candidate invariant mass. The signal peak is described using a Breit-Wigner function whereas the combinatorial background is parameterized with an empirical function (equation 5.5). For fit results, see table 5.1*

peak and a 3-parameter empirical function for the background spectrum:

$$f(m;\ldots) = n_\phi \frac{\Gamma_\phi/2\pi}{(m-m_\phi)^2 + (\Gamma_\phi/2)^2} + n_\not{\phi}(m - 2m_{\mathrm{K}^\pm})^a \cdot e^{-b(m-2m_{\mathrm{K}^\pm})^c} \tag{5.5}$$

The $\phi$ signal area is defined using the result of the fit: $\phi \to \mathrm{K}^+\mathrm{K}^-$ candidates having an invariant mass in the interval $m_\phi \pm \Gamma_\phi$ are kept. Using the integral [72] of the background description function, the number of combinatorial background events in the $\phi$ signal area is estimated:

$$\int_0^x t^a \cdot e^{-bt^c}\mathrm{d}t = \frac{1}{c \cdot b^{\frac{a+1}{c}}} \cdot \gamma(\frac{a+1}{c}, bx^c) \tag{5.6}$$

The fit result for the "good" and "bad" resolution samples is shown in table 5.1

The motivation for selecting true $\phi$ versus combinatorial background as the first classification criterion is that events outside the $\phi$ signal area (the sidebands or wings) will be used to describe the distribution of the combinatorial background under the $\phi$ peak. Keeping the top level "good" and "bad" separation in mind, the likelihood is written:

$$\mathcal{L}^{\{\mathrm{good,bad}\}} = \prod_{\mathrm{peak}} (f_\phi p^\phi + f_\not{\phi} p^\not{\phi}) \times \prod_{\mathrm{wing}} p^\not{\phi} \tag{5.7}$$

| Sample | Peak normalization | $m_\phi$( GeV/$c^2$) | $\Gamma_\phi$( GeV/$c^2$) | Events within cuts | $f_{\rm com}$(%) |
|---|---|---|---|---|---|
| "good" | $179.0 \pm 21.5$ | $1.019 \pm 0.001$ | $0.0067 \pm 0.0011$ | 180 | $33.6 \pm 3.5$ |
| "bad" | $195.7 \pm 30.2$ | $1.020 \pm 0.001$ | $0.0063 \pm 0.0016$ | 317 | $48.4 \pm 2.8$ |
| all | $381.4 \pm 37.2$ | $1.020 \pm 0.001$ | $0.0068 \pm 0.0009$ | 497 | $44.8 \pm 2.2$ |

Table 5.1: *$\phi$ invariant mass spectrum fit results*

Ignoring the possibility that the charge correlation distribution could be time dependent, $p^\phi(t,c) = T^\phi(t) \cdot C^\phi(c)$ is a simple product of the decay time distribution $T^\phi$ and the charge correlation distribution $C^\phi$. A single Gaussian is used to fit the charge correlation and a combination of two Gaussians and two smeared exponential decay distributions is used to describe the decay time. A single set of resolution parameters is used to describe "good" as well as "bad" events but the fractions described by each resolution are left as independent in the two resolution classes. Projections of the distributions with the fitted functions are shown in figure 5.5.

The use of sideband events to describe the distribution of combinatorial background in equation 5.7 is just a formal way of performing background subtraction. It is thus essential to control that the sideband events and the combinatorial background under the peak have the same distributions. The idea of using same sign $K^\pm$ pairs for checking is rejected because such events are produced by different physics processes than for opposite sign $K^\pm$-pairs. Instead, the sideband events are divided in bins of different invariant mass and the parameters describing the charge correlation and apparent lifetime are checked in each bin. The evolutions of these parameters are shown in figure 5.3, and no systematic trends are observed.

### 5.1.2 Light ($udsc$) quark and fake lepton events

The next level of classification in figure 5.1 is with respect to the origin of the lepton candidate.

$$p^\phi = f_{\rm udsc} p^{\rm udsc} + f_{\rm fake} p^{\rm fake} + f_{\rm signal} p^{\rm signal} \tag{5.8}$$

Neither the distribution nor the fraction of events with lepton candidates not coming from a $b$ quark decay chain ($udsc$-events) and of events with hadrons incorrectly identified as leptons (fake events) can be determined from the data alone. The compositions are evaluated on $q\overline{q}$ Monte Carlo events applying the same cuts as on data, see table 5.2. The time and charge distribution are

| Sample | $f_{\rm udsc}$(%) | $f_{\rm fake}$(%) |
|---|---|---|
| "good" | $16.0 \pm 2.0$ | $2.3 \pm 0.8$ |
| "bad" | $17.6 \pm 1.9$ | $2.7 \pm 0.8$ |

Table 5.2: *Estimated background contributions to $\phi$ signal*

parameterized using functions resembling the ones used to describe the combinatorial background. A sum of two Gaussians of zero mean is fit to the $udsc$-class decay time distribution, and the fake-distribution is fit with a single smeared exponential. In each class, the charge correlation distribution is fit with a Gaussian. One single set of functions is used to describe "good" and "bad" events. Plots of the distributions observed on simulated events and the fitted functions are shown in figure 5.5

### 5.1.3 Cascade and direct semileptonic $b$ quark decays

After splitting off the combinatorial $\phi$, light quark events and fake leptons, what remains of the likelihood function is the part that depends explicitly on the physical parameters. Leptons from

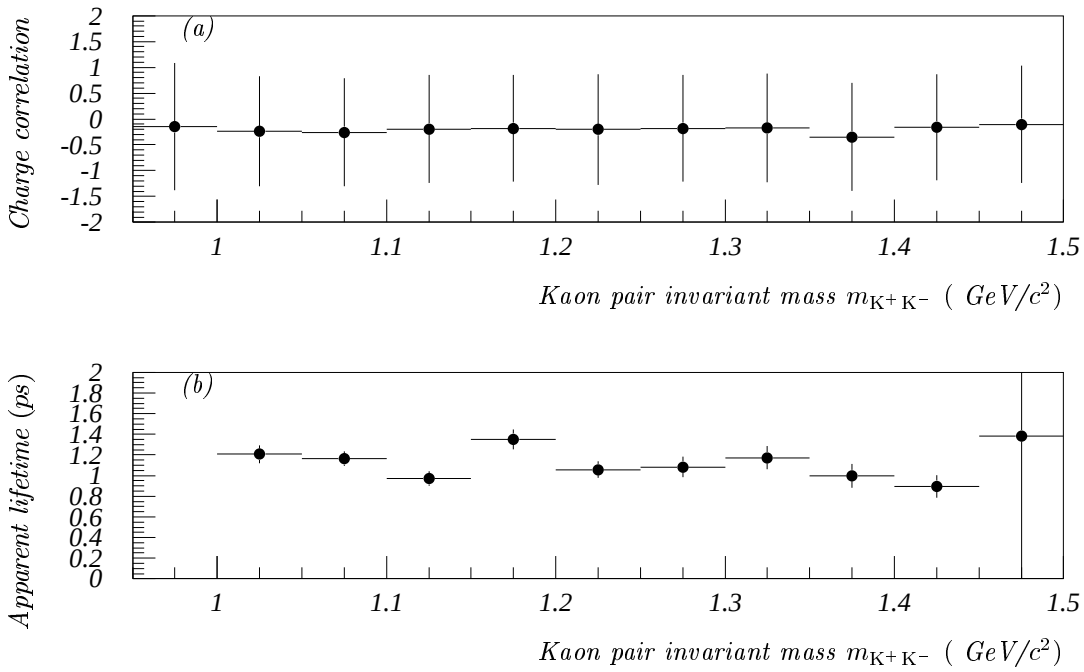


Figure 5.3: *Charge correlation and apparent lifetime for combinatorial background. Panel (a) shows the mean (dot) and variance (vertical error bar) of the charge correlation variable. Panel (b) is generated fitting the decay time distribution with two Gaussians and one smeared exponential, leaving the apparent lifetime free to vary from bin to bin.*

semileptonic decays ($\overline{b} \to \ell^+$), wrong sign cascade decays ($\overline{b} \to c \to \ell^-$) and right sign cascade decays ($\overline{b} \to c\overline{c} \to \ell^+$ and $\overline{b} \to \tau^+ \to \ell^+$) are all treated as signal but with different resolution functions:

$$p^{\text{signal}} = \sum_{\text{i}\in\text{resolutions}} \sum_{\text{j}\in\text{flavors}} f_i f_j R^{ij} \otimes p^j(t, c, x; \tau_j, \Delta m_j, \Delta\Gamma_j) \tag{5.9}$$

In principle, the $\otimes$-symbol denotes convolution in two dimensions: time and charge. Fortunately, the second convolution is just a sum over the $-1$ (unmixed) and $+1$ (mixed) possibilities. In addition, the charge and time resolution functions are assumed to factorize, i.e. the charge resolution is assumed to be independent of the decay time and vice versa. Naming the true decay time and true decay mixing $\tilde{t}$ and $\tilde{c}$ respectively, the convolution is written:

$$p^{ij}_{\text{obs}}(t, c) = \int_0^\infty \mathrm{d}\tilde{t} \sum_{\tilde{c}=\{-1,+1\}} R^{ij}_t(t|\tilde{t}) R^{ij}_c(c|\tilde{c}) p^j_{\text{true}}(\tilde{t}, \tilde{c}; \tau, \Delta\Gamma, \Delta m) \tag{5.10}$$

The signal fractions denoted $f_i$ and $f_j$ are computed using reconstruction efficiencies from Monte Carlo simulations combined with updated branching fractions from [4]. The fractions of direct lepton and cascade lepton events are shown in table 5.3. The contributions of the different $b$-hadron species to the direct lepton sample are shown in table 5.4. Using the numbers quoted in tables 5.1, 5.2, 5.3 and 5.4, the selected signal of $\mathrm{B}^0_s \to \phi\ell^+$ decays is estimated to be $127.9 \pm 9.7$ (binomial error).

Using simple error propagation on equation 4.6, two contributions to the decay time error are

| Sample | $\overline{b} \to \ell^+(\%)$ | $\overline{b} \to \overline{c}c \to \ell^+(\%)$ | $\overline{b} \to c \to \ell^-(\%)$ |
|---|---|---|---|
| "good" | $90.6 \pm 8.5$ | $3.7 \pm 0.6$ | $5.6 \pm 0.9$ |
| "bad" | $87.2 \pm 7.8$ | $5.6 \pm 0.8$ | $7.3 \pm 1.1$ |
| total | $89.1 \pm 5.9$ | $4.5 \pm 0.5$ | $6.4 \pm 0.7$ |

Table 5.3: *Lepton sources*

| Sample | $B^0$ | $B^+$ | $B_s^0$ | $\Lambda_b$ |
|---|---|---|---|---|
| "good" | $18.1 \pm 5.7$ | $12.0 \pm 3.8$ | $69.8 \pm 19.0$ | $0.1 \pm 0.1$ |
| "bad" | $19.0 \pm 5.5$ | $21.4 \pm 6.1$ | $57.6 \pm 21.2$ | $2.0 \pm 2.4$ |
| total | $18.5 \pm 4.0$ | $16.1 \pm 3.4$ | $64.4 \pm 14.4$ | $1.0 \pm 1.1$ |

Table 5.4: *Contributions from the different b-hadron species*

found ($\oplus$ means geometric sum):

$$\sigma_t = \frac{\partial t}{\partial l}\sigma_l \oplus \frac{\partial t}{\partial p}\sigma_p = \frac{m}{p}\sigma_l \oplus t\frac{\sigma_p}{p} \tag{5.11}$$

This motivates the following parameterization of the decay time resolution: The measurement error as well as a possible measurement bias are allowed to depend on the true time:

$$\mu(\tilde{t}) = u + v\tilde{t} \tag{5.12}$$

$$\sigma(\tilde{t}) = \sqrt{r^2 + s^2\tilde{t}^2} \tag{5.13}$$

Inserting $\sigma(\tilde{t})$ and $\mu(\tilde{t})$ in a standard Gaussian, the time dependent resolution function is obtained:

$$R_t(t|\tilde{t};\ldots) = \frac{1}{\sqrt{2\pi}\sigma(\tilde{t})} e^{-\frac{1}{2}\left(\frac{t-\mu(\tilde{t})}{\sigma(\tilde{t})}\right)^2} \tag{5.14}$$

For a fixed true decay time $\tilde{t}$, $R_t$ is normalized to 1, just as probability density functions should be:

$$\int_{-\infty}^{+\infty} R_t(t|\tilde{t})\mathrm{d}t = 1$$

In order to have an accurate description of the time measurement resolution including the tails, a sum of two resolution functions like equation 5.14 is used. For convenience, the function $S(t-\tilde{t},\tilde{t} = R(t,\tilde{t})$ is fitted to a two dimensional histogram of the measurement error $t-\tilde{t}$ versus the true time $\tilde{t}$ for Monte Carlo signal events. The resulting resolution functions for the "good" and "bad" signal samples are shown in figure 5.4. The parameters describing the central Gaussian of the resolution in each class are summarized as follows:

"good": $\sigma_t = 0.115$ ps $\oplus\ 0.079 \cdot t$ (75% of sample)
"bad": $\sigma_t = 0.287$ ps $\oplus\ 0.108 \cdot t$ (83% of sample)

The charge correlation is parameterized for each signal class using a single Gaussian fitted to the simulated charge correlation distributions for unmixed and mixed events. Projections of the multidimensional distribution together with the fitted likelihood function appear in figure 5.5

## 5.2 Measuring the $B_s^0$ lifetime $\tau_{B_s^0}$

To measure the lifetime of the $B_s^0$ meson, the charge correlation information is ignored and the life time difference is fixed at $\Delta\Gamma_s = 0$. The parameters describing the background distributions are

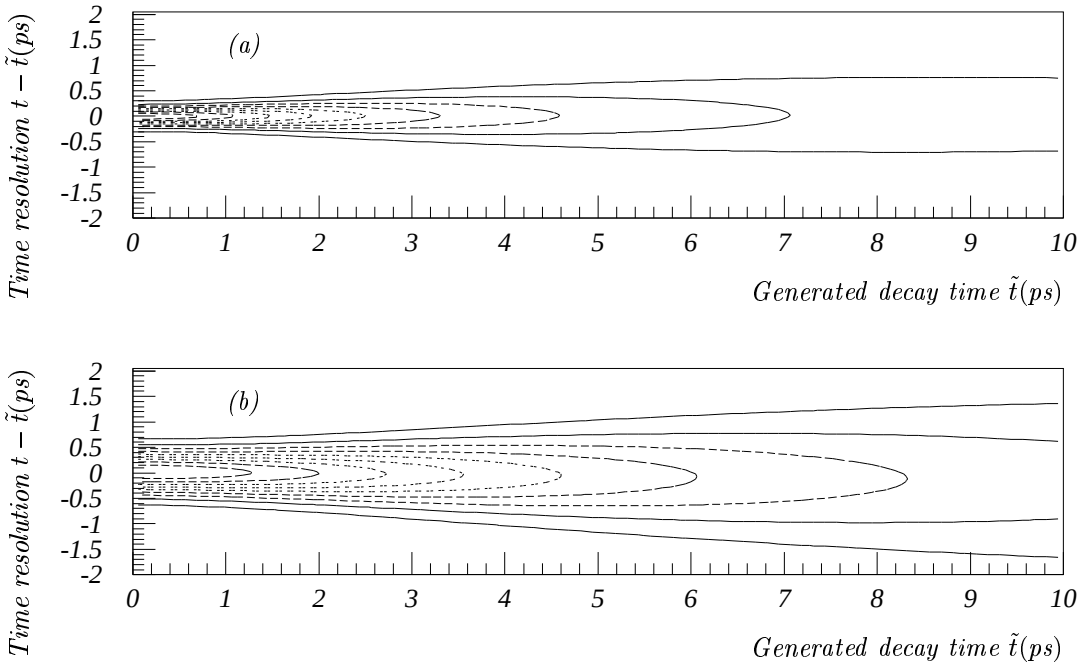


Figure 5.4: *Contour plot of time resolution functions. The simulated time is shown on the horizontal axis and the measurement error on the vertical axis. Panel (a) shows the resolution of the "good" sample and the lower Panel (b) the "bad" sample.*

allowed to vary freely in the fit. Thus, the statistical error on the fitted value $\tau_{B_s^0} = (1.393^{+0.310}_{-0.272})$ ps includes systematic errors coming from the shape of the background distribution. The contributions to the systematic errors from various other sources are evaluated explicitly by changing each parameter and observing the shift in the fitted lifetime:

- The lifetime of each of the other $b$-hadron species is varied within its experimental error ([4]).
- The fraction of signal events is displaced according to the error on $f_s$ ([4]).
- The level of the combinatorial background is changed within the error on the fitted value (figure 5.2).

Comparing the average estimated $B_s^0$ energy in data with Monte Carlo simulations (figure 4.21, table 4.3) indicates that the $B_s^0$ meson energy is controllable at the level of

$$\frac{30.8 \text{ GeV} - 30.5 \text{ GeV}}{30.5 \text{ GeV}} = 1.0\%.$$

This is translated directly into a 1.0% systematic error on the fitted lifetime.

To check for a possible bias, the fitting procedure is applied to a pure $B_s^0$ signal sample generated with a lifetime of $\tau_{B_s^0} = 1.6$ ps. The corresponding fitted lifetime is $(1.591 \pm 0.028)$ ps, indicating a systematic shift of

$$\Delta^{\text{shift}} = (-0.009 \pm 0.028) \text{ ps}. \tag{5.15}$$

This shift is subtracted from the fitted value and its uncertainty is included as a systematic error[1].

[1] For consistency, this correction done even when the systematic shift is smaller than its uncertainty

All the contributions to the systematic error on the $B_s^0$ lifetime are listed in table 5.5. Summing the contributions in quadrature and applying the correction from equation 5.15, the lifetime of the $B_s^0$ meson measured in the $\phi\ell^+$ channel is:

$$\tau_{B_s^0} = \left(1.40^{+0.31}_{-0.27}\ (stat) \pm 0.08\ (syst)\right)\ \text{ps}, \tag{5.16}$$

| Parameter | Central value | Variation | Contribution |
|---|---|---|---|
| $\tau_{B^0}$ | 1.56 ps | $\pm$0.06 ps | 0.020 ps |
| $\tau_{B^+}$ | 1.62 ps | $\pm$0.06 ps | 0.017 ps |
| $\tau_{\Lambda_b}$ | 1.14 ps | $\pm$0.08 ps | 0.002 ps |
| $f_{B_s^0} \cdot \mathrm{Br}(B_s^0 \to D_s^- \ell^+ \nu_\ell X)$ | 0.016 | $^{-0.0043}_{+0.038}$ | 0.052 ps |
| $f_\phi$ | 0.539 | $\pm$10% | 0.049 ps |
| Resolution | – | $\pm$10% | 0.017 ps |
| $\langle E \rangle$ | – | $\pm$1.0% | 0.014 ps |
| $\Delta^{\text{shift}}$ | – | – | 0.028 ps |
| Total | | | 0.084 ps |

Table 5.5: *Contributions to the systematic error on* $\tau_{B_s^0}$

## 5.3 An upper limit on the decay width splitting $\Delta\Gamma_s$

To extract an upper limit on the width difference between the wide and narrow $B_s^0$-states, $\tau_{B_s^0}$ is kept at its maximum likelihood value while scanning the parameter $\Delta\Gamma_s/\Gamma_s$ of equation 5.1. The resulting log-likelihood function shown in figure 5.6 has a broad minimum at the border of the physical region, $\Delta\Gamma_s/\Gamma_s = 0$. Using $\Delta \ln \mathcal{L} = -\frac{1}{2}(1.96)^2$, as conventional for a 95% C.L., an upper limit on the relative width difference is found:

$$\Delta\Gamma_s/\Gamma_s =\leq 0.91\ (95\%\ \text{C.L.}). \tag{5.17}$$

As this hardly constrains $\Delta\Gamma_s$ at all, the contributions from systematic errors to this limit are not considered.

## 5.4 Excluded regions for the mass difference $\Delta m_s$

Keeping all other parameters fixed, $\Delta \ln \mathcal{L}$ can be studied as a function of $\Delta m_s$, see figure 5.7. However, the classical likelihood ratio method for extracting confidence intervals has several complications when applied to the analysis of $\Delta m_s$. Firstly, the statistical distribution is far from Gaussian, so the standard value of $\Delta \ln \mathcal{L} = -\frac{1}{2}(1.96)^2$ cannot be used to extract the 95% confidence limit. Instead, one would have to perform Monte Carlo simulations to compute the correct $\Delta \ln \mathcal{L}$. Secondly, including systematic errors in the quoted limit is at best cumbersome; it would probably involve another series of Monte Carlo simulations. And finally, there is no prescription on how to account for common systematics when combining the results from several analyses and experiments.

The amplitude method [73] was invented to overcome the problems mentioned above. The method proceeds by introducing an arbitrary oscillation amplitude through the substitution

$$\cos \Delta m t \mapsto A \cos \Delta m t$$

in equation 5.1. For a given value of $\Delta m_s$, the amplitude $A$ is fitted, measuring the Fourier component $A$ of the mixing at the frequency $\Delta m_s$, with some measurement error $\sigma_A$. If the true

frequency of the time dependent mixing is far away from $\Delta m_s$, the expectation value of $A$ is 0. On the other hand, if the selected data sample indeed has a mixing frequency of $\Delta m_s$, the expectation value of $A$ is 1. Thus, excluding $A = 1$ at a given confidence level is equivalent to excluding a mixing frequency of $\Delta m_s$.

For an alternative interpretation of the amplitude method, consider the convolution of the oscillatory term in equation 5.1 with a finite width resolution function as the oscillation frequency grows without limit:

$$\lim_{\Delta m \to \infty} \int_0^\infty \mathrm{d}\tilde{t} R(t, \tilde{t}) \cdot e^{-\bar{t}/\tau} \cos \Delta m \tilde{t}.$$

For fixed $t$, this is recognized as the complex Fourier transform of the resolution, which should vanish as $\Delta m \to \infty$ for any physical function. From this, it is concluded that setting the oscillation amplitude $A = 0$ is equivalent to evaluating the likelihood function at $\Delta m \to \infty$, which again is equivalent to time independent (instantaneous) mixing. On the other hand, unity amplitude is equivalent to the physical situation of having oscillations at the given frequency. All other values of $A$ are unphysical. In this respect, the amplitude method is a continuation of the likelihood function into an unphysical region.

According to [73], the contributions to the systematic error on the amplitude have to take into account the correlation between the fitted amplitude and its error:

$$\sigma_A^{\mathrm{syst}} = \Delta A + (1 - A)\frac{\Delta \sigma_A}{\sigma_A} \tag{5.18}$$

The following four sources of systematic errors were considered:

- The level of combinatorial background has an uncertainty of 10%, coming from the fit to the $\mathrm{K^+K^-}$ invariant mass spectrum.
- The uncertainty on the fraction of $\mathrm{B_s^0}$-events in the signal is changed according to [74]:

$$f_{\mathrm{B_s^0}} \cdot \mathrm{Br}(\mathrm{B_s^0} \to \mathrm{D_s^-} \ell^+ \nu_\ell \mathrm{X}) = (1.60^{+0.43}_{-0.38})\%$$

- The time resolution parameters $\sigma_l$ and $\sigma_p$ of the signal were simultaneously changed by $\pm 10\%$
- The mean and width describing the resolution of the charge correlation tag were changed to the effect of a $\pm 3\%$ change in the total tag efficiency.

Figure 5.8 shows the fitted amplitude $A$ as a function of the mass difference parameter $\Delta m_s$. The 95% upper limit on the amplitude, computed as $A + 1.645 \cdot \sigma_A$, is found below the shaded band. The shaded band shows the upper limit when including systematic errors. Regions of $\Delta m_s$ where the upper limit on the amplitude is below 1 is excluded. Using that the expectation of $A$ vanishes for $\Delta m_s$ hypotheses far away from the true value, 50% of an ensemble of equal experiments will exclude amplitudes above $1.645\sigma_A$ as indicated with a dashed line. Thus, the point where $1.645\sigma_A$ reaches 1 is a measure of the sensitivity of the experiment. For the present analysis, the sensitivity reach $\Delta m_s < 1.5$ ps$^{-1}$. The amplitude method gives several excluded regions for $\Delta m_s$:

$$\Delta m_s \notin \langle 0.9 \text{ ps}^{-1} \ldots 1.8 \text{ ps}^{-1} \rangle \cup \langle 2.2 \text{ ps}^{-1} \ldots 3.3 \text{ ps}^{-1} \rangle \cup \langle 8.0 \text{ ps}^{-1} \ldots 9.0 \text{ ps}^{-1} \rangle \text{ (95\% C. L.)},$$

of which the first two are confirmed by the likelihood scan (figure 5.7). For the third region, a corresponding decrease in the likelihood function is seen, but not below the canonical $-\frac{1}{2}(1.96)^2 = 1.92$ level. This is hardly surprising, as the simple likelihood method is sensitive to the likelihood ratio at $A = 1$ and $A = 0$ for a given $\Delta m_s$, while the amplitude method in principle compares the likelihood at $A = 1$ to the maximum likelihood when leaving $A$ free.

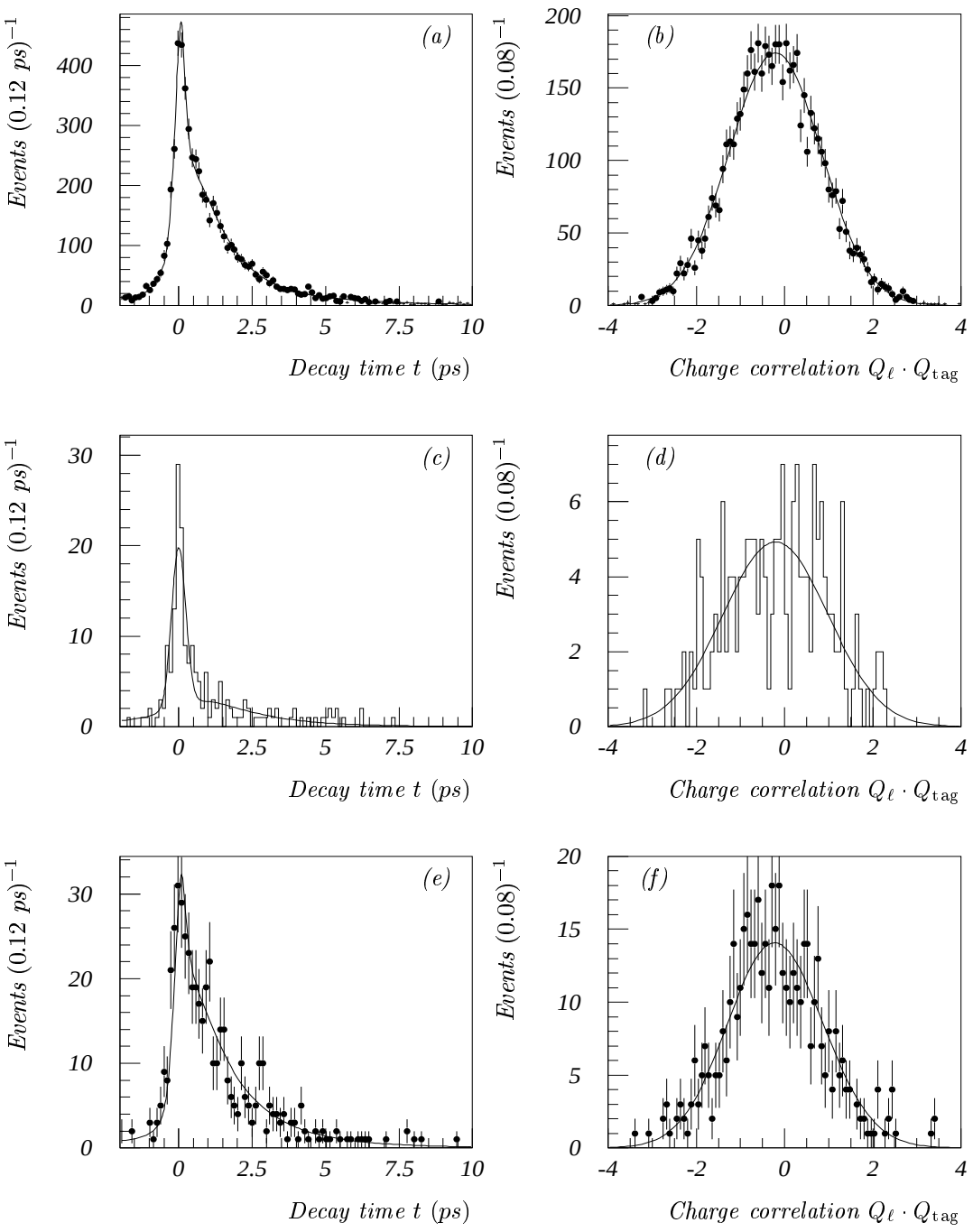


Figure 5.5: *Decay time and charge correlation distributions with fitted likelihood functions. Panels (a) and (b) show the distributions of sideband events, used to describe combinatorial background under the $\phi$ peak. Panels (c) and (d) show the distributions of udsc and fake lepton events added as found in $q\overline{q}$ simulation. Panels (e) and (f) show the distributions of signal events.*

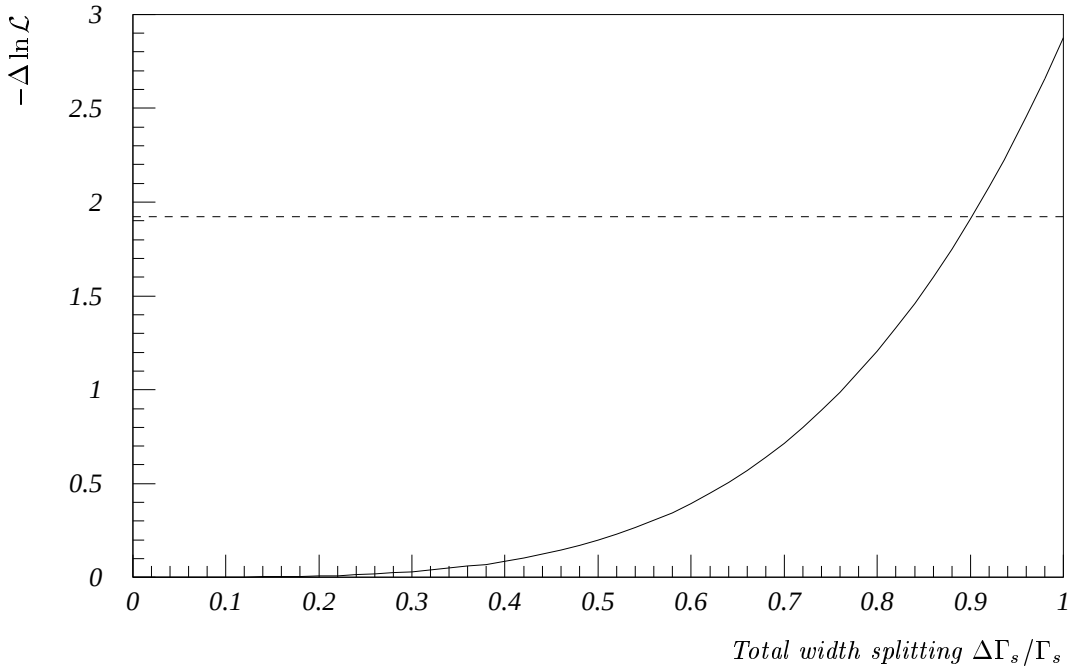


Figure 5.6: *Upper limit on the* $B_s^0$ *decay width difference* $\Delta\Gamma_s/\Gamma_s$. *The* 95% *C.L. upper limit is defined as the point where the log-likelihood function has decreased by* $\frac{1}{2}(1.96)^2 = 1.92$ *units from its maximum value.*

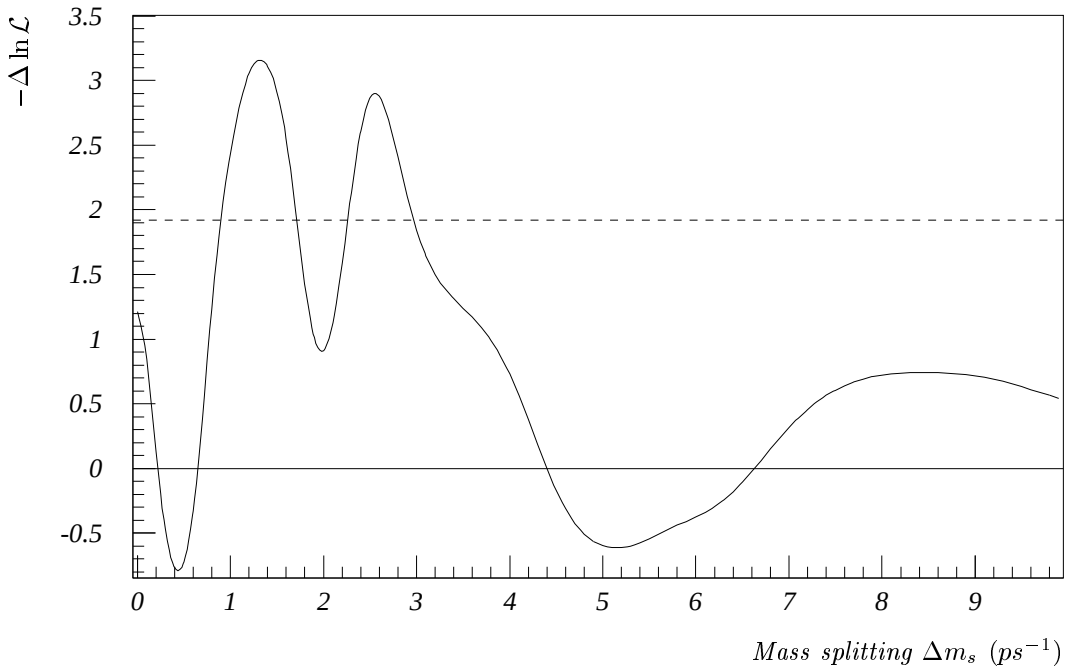


Figure 5.7: *The log-likelihood difference versus* $\Delta m_s$. *The* $\Delta \ln \mathcal{L}$ *that corresponds to* 95% *C.L. for Gaussian statistics is shown as a dashed curve. The likelihood reference point is* $\Delta m_s \to \infty$ *(A=0).*

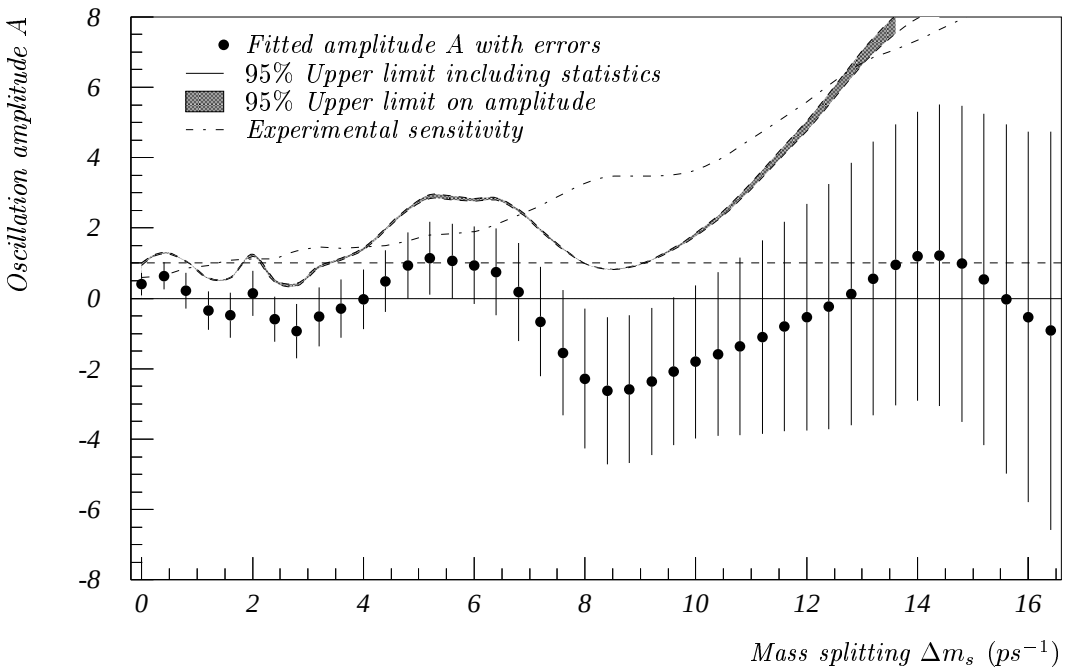


Figure 5.8: *Fitted amplitudes used to extract the lower limit on* $\Delta m_s$.

# Chapter 6

# Discussion and conclusions

## 6.1 Comparison to other analyses

The measurement of the $B_s^0$-lifetime and the limit on $\Delta m_s$ found in the present work can be compared to limits obtained from $B_s^0$-decays reconstructed in other decay channels, see figures 6.1 and 6.2. Comparing the errors of the lifetime fit, one can conclude that the effective number of events in the $\phi\ell^-$ analysis is about one half the effective number of events in the $D_s^-\ell^+$ analysis.

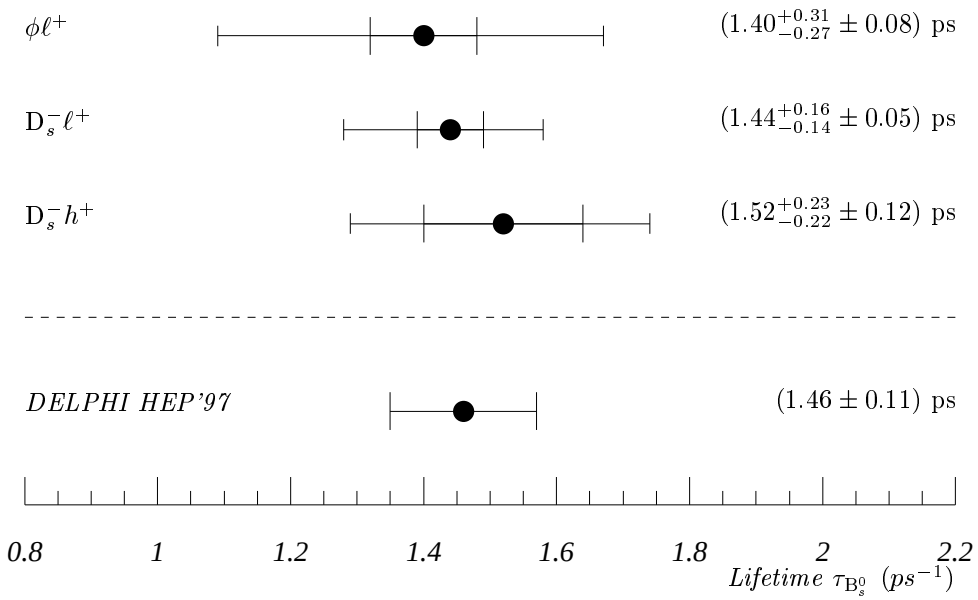


Figure 6.1: $B_s^0$ *lifetime measured by the DELPHI experiment. For each reconstruction channel, the systematic errors are quoted first, then systematic errors; for the combined measurement, only the total error is quoted. The combined result was prepared for the HEP'97-conference [64] (appendix A) and contains a previous version of the present analysis.*

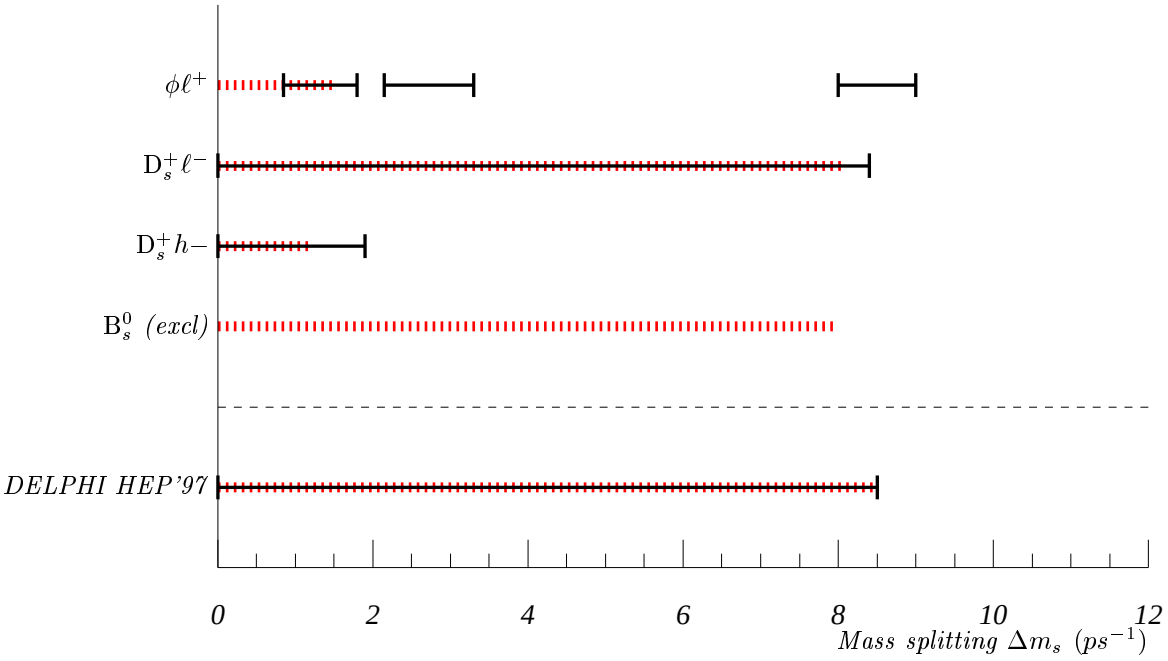


Figure 6.2: *Limits on $\Delta m_s$ from the DELPHI experiment. The sensitivity of each channel is shown as a dashed line and the excluded regions as solid bars. The analysis of exclusive final states ($B_s^0$ (excl)) [65] is not yet finalized, and only the expected sensitivity is shown. The combined result was prepared for the HEP'97-conference [64] (appendix A) and contains a previous version of the present analysis.*

## 6.2 Statistical correlation to $D_s^-\ell^+$ analysis

The task of combining $B_s^0$ mixing analyses of different subchannels and experiments is undertaken by the *LEP B Oscillations Working Group*[1] [74]. The combination procedure takes into account statistical as well as systematic correlations between analyses [75].

From the inclusive nature of the present analysis, it is clear that there are common events with DELPHI's $D_s^+\ell^-$ analysis [64], giving a statistical correlation between the analyses. A proper evaluation should use event lists to check the actual overlap between the selections of the two analyses. Unfortunately, such event lists are not available at this point. However, to get an indication of the correlation, the fractions of common signal events are considered. These are found in the $D_s^- \to \phi\pi^-$ and $D_s^- \to \phi\ell^-$ subchannels, which together account for $(46.3 \pm 3.6)\%$ of the signal events in the $D_s^-\ell^+$ analysis. The fraction of overlap signal events in the inclusive $\phi\ell^+$ analysis is found using a sample of simulated signal events. Of the events that pass the selection cuts, $(23.0 \pm 0.7)\%$ belong to the $D_s^- \to \phi\pi^-$ mode and $(23.4 \pm 0.8)\%$ belong to the $D_s^- \to \phi\ell^-$ mode. As the inclusive analysis does not utilize any lepton identification cuts, it is probably conservative to assume that all $D_s^- \to \phi\ell^-$ events are selected by the exclusive analysis. On the other hand, the non-linear discriminant discussed in 4.7 gives an enhanced statistical weight to events from the $D_s^- \to \phi\pi^-$ mode — these events are responsible for the peak at the nominal $D_s^-$ invariant mass seen in figure 4.24f. The enhanced weight might lead to a higher than expected

[1] The name of the group is misleading, as representatives from the CDF (Fermilab) and SLD (SLAC) participate as well.

correlation through the $D_s^- \to \phi\pi^-$ channel.

## 6.3 Impact of using a detailed likelihood function

The likelihood description used to extract the $\Delta m_s$ exclusion regions has a number of features not required for a bare minimum analysis:

- The multidimensional discriminant described in section 4.7 is supposed to increase the effective signal to background ratio, acting like a per-event signal fraction.
- The output of the Fisher discriminant constructed for the production charge tag is included as a continuous observable.
- The selected events are separated in two classes depending on the goodness of the vertex fit and the expected flight length resolution.

Stripping off these extra features and adding them back one by one, their importance for the sensitivity of the $\Delta m_s$ analysis is clearly exposed in figure 6.3: The point of 50% exclusion probability moves from 0.6 $ps^{-1}$ to 1.5 $ps^{-1}$, and would reach 2.4 $ps^{-1}$ with only 10% increased statistics. At $\Delta m_s = 2.4$ $ps^{-1}$ the amplitude error decreases by a factor 1/1.5, which is equivalent to a twofold increase in statistics. A particularly striking aspect is that the effect of splitting the sample in two resolution classes increases progressively towards higher oscillation frequencies, compared to the impact of the two other likelihood-enhancements. This reflects the fact that the analysis is resolution limited at large $\Delta m_s$.

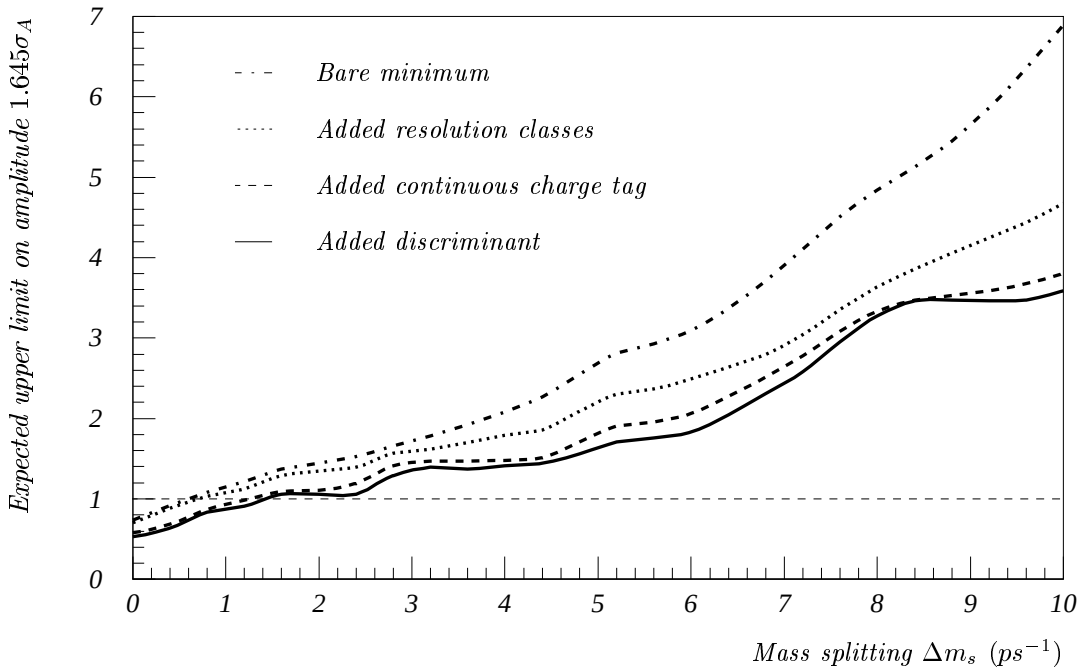


Figure 6.3: *Level of detail and $\Delta m_s$-sensitivity. Note that the fit error $\sigma_A$ is correlated to the fitted value of the amplitude A. This means that the sensitivity curves are prone to statistical uncertainty.*

## 6.4 Conclusions

Inclusive $B_s^0 \to \phi\ell^+$ decays in the DELPHI detector are reconstructed to study the $B_s^0$ lifetime, mixing and decay width splitting.

The $B_s^0$ lifetime is measured to be

$$\tau_{B_s^0} = \left(1.40^{+0.31}_{-0.27}\ (stat) \pm 0.08\ (syst)\right)\ \text{ps},$$

compatible with the current world average [5] $\tau_{B_s^0} = (1.54 \pm 0.07)$ ps, as well as to theoretical predictions that $\tau_{B_s^0}/\tau_{B_d^0} = 1.00 \pm 0.01$.

Using elaborate statistical methods, the sensitivity for time dependent $B_s^0\overline{B}_s^0$ mixing reaches $\Delta m_s < 1.5\ \text{ps}^{-1}$. No signal of time dependent mixing is seen, and the exclusion regions for $\Delta m_s$ listed in section 5.4 are compatible with the combined limit of $\Delta m_s > 10.2\ \text{ps}^{-1}$ 95% CL..

Analyzing data from the years 1991–1993 could theoretically double the available statistics, and should be included in the analysis. A twofold increase in statistics could push the mixing sensitivity up to $\Delta m_s < 3.0\ \text{ps}^{-1}$. However, as neither RICH particle identification nor the VD $z$ information are available in these datasets, the background as well as the resolution are expected to be worse than in the 1994–1995 data.

# Appendix A

# Contribution to the HEP 97 Conference

The article included in this appendix is written by the DELPHI B lifetimes/oscillation team. Section 5 and figures 10, 11 and 15 represent work done by the present author.

HEP'97 #457
Submitted to Pa 9
Pl 5

DELPHI 97-75 CONF 61
20 July, 1997

# Search for $B_s^0$ – $\overline{B_s^0}$ oscillations and measurement of the $B_s^0$ lifetime

Preliminary

DELPHI Collaboration

**P. Billoir[1], A. Borgland[2], G. Borisov[3], M. Canepa[4], P. Checchia[5], G. Eigen[2], O. Kouznetsov[6], X. Moreau[1], F. Parodi[4], E. Piotto[5], P. Privitera[7], O. Røhne[8], P. Roudeau[9] and A. Stocchi[9]**

[1] LPNHE, Paris, France
[2] Bergen University, Norway
[3] CEA, Saclay, France
[4] Genova University, Italy
[5] Padova University, Italy
[6] Dubna, JINR, Russian Federation
[7] Roma II University, Italy
[8] Oslo University, Norway
[9] LAL, Orsay, France

**Abstract**

This note presents a new limit on the mass difference of the physical $B_s^0$ states which partially updates the result given in a previous publication, which is appended in the Addendum. In the previous publication three analyses were performed using events containing a lepton emitted at large transverse momentum, $p_t$, relative to its jet axis $\ell - Q_{hem.}$, $\ell - \ell$ and $D_s\ell$. Among these three analyses only the $D_s\ell$ has been updated and two new analyses are presented : $D_s h$ and $\phi\ell$. Combining these measurements the limit turns out to be:

$$\Delta m_s > 8.5 \text{ ps}^{-1} \text{ at } 95\% \text{ C.L.}$$

The exclusion probability for this limit is 50%.
With the three samples, $D_s\ell$, $D_s h$ and $\phi\ell$, the lifetime of the $B_s^0$ meson has been measured :

$$\tau_{B_s^0} = (1.46 \pm 0.11)\ ps$$

Paper submitted to the HEP'97 Conference
Jerusalem, August 19-26

# 1 Introduction

In the Standard Model, $B_s^0 - \overline{B_s^0}$ mixing is a direct consequence of second order weak interactions. Starting with a $B_s^0$ meson produced at time $t$=0, the probability, $\mathcal{P}$, to observe a $B_s^0$ or a $\overline{B}_s^0$ decaying at the proper time $t$ can be written, neglecting effects from CP violation and a possible difference between the lifetimes of the heavy and light eigenstates :

$$\mathcal{P}_{B_s^0}^{unmix.(mix.)} = \mathcal{P}(B_s^0 \to B_s^0(\overline{B}_s^0)) = \frac{1}{2\tau_{B_s^0}} e^{-\frac{t}{\tau_{B_s^0}}} [1 \pm \cos(\Delta m_s t)] \quad (1)$$

Three important quantities have to be considered in this analysis:

- the determination of the sign of the $b$ quark at production time (production tag)
- the determination of the sign of the $b$ quark inside the B meson at decay time (decay tag)
- the measurement of the B decay proper time.

In the previous publication [1] three analyses were performed using events containing a lepton emitted at large $p_t$ relative to its jet axis [1]. In two of them ($\ell - Q_{hem.}$ and $\ell - \ell$) the proper time was measured using an inclusive vertex algorithm to reconstruct the decay distance and the energy of the candidate B hadron. In the $\ell - Q_{hem.}$ analysis the "production tag" was obtained from the value of the hemisphere charge measured in the hemisphere opposite to the lepton and the lepton charge was used to define the "decay tag". The $\ell - \ell$ analysis used events with two high $p_t$ leptons, identified in opposite hemispheres, and it was not possible to separate the notions of "decay" and "production" tags. In the third analysis, $(D_s\ell) - Q_{hem.}$, the identified lepton was accompanied, in the same hemisphere, by an exclusively reconstructed $D_s$. The "production tag" was obtained from the value of the hemisphere charge measured using both hemispheres.

This note presents an update of the $(D_s\ell)$ analysis using 94-95 data and corresponding to 2.2M hadronic $Z^0$ decays The improvements with respect to the previous analysis come from :

- the use of the data reprocessed with an improved tracking algorithm,
- the use of a new algorithm to define the "production tag".

The previous analysis of the data registered between 1991 and 1993 is included using the new tagging procedure.
Two other analyses are also presented : $D_s h$ and $\phi\ell$ analyses. The $D_s h$ analysis is similar to the $D_s\ell$ one but instead of the lepton it uses a charged hadron. It provides larger statistics but the $B_s^0$ purity of the sample is poorer. The $\phi\ell$ analysis is more inclusive and considers events where a high transverse momentum lepton is accompanied, in the same hemisphere, by a $\phi$ meson. These two new analyses use the data from 94-95 and benefit from the new algorithm to define the "production tag".
For the combined limit on $\Delta m_s$ the $(\ell - Q_{hem.})$ and $(\ell - \ell)$ analyses are also included ( these analyses have not been updated and their description is given in the addendum.)

[1] In addendum, this paper which is about to be published as CERN-PPE, is appended

## 2 The new algorithm for the tagging at the production time

The signature of the initial production of a $b(\overline{b})$ quark in the jet containing the $B_s^0$ or $\overline{B_s^0}$ candidate was determined using a combination of different variables sensitive to the initial quark state. For each individual variable $X_i$, the probability density functions $f_b(X_i)$ ($f_{\overline{b}}(X_i)$) for $b$ ($\overline{b}$) quarks were built and the ratio $R_i = f_{\overline{b}}(X_i)/f_b(X_i)$ was computed. The combined tagging variable was defined as:

$$x_{tag} = \frac{1-R}{1+R} \quad , \text{ where } \quad R = \prod R_i. \tag{2}$$

The variable $x_{tag}$ varies between -1 and 1. High values of $x_{tag}$ correspond to a high probability that a given hemisphere contained a $b$ quark in the initial state. If some of the variables $X_i$ are not defined for a given event, the corresponding ratios $R_i$ are set to 1, corresponding to equal probabilities for the initial state to be $b$ or $\overline{b}$. Definition (2) of $x_{tag}$ provides an optimal separation between $b$ and $\overline{b}$ initial states if the individual discriminating variables $X_i$ are independent.
A set of 9 discriminating variables was selected for this analysis. The event was divided into 2 hemispheres by the plane perpendicular to the direction of the $B_s^0$ candidate. One set (3 variables) was determined in the hemisphere which contains the $B_s^0$ meson, the other set (5 variables) in the hemisphere opposite to the $B_s^0$ meson, and 1 variable was common to both hemispheres.

All reconstructed charged tracks were used in the opposite hemisphere, and the following variables were defined:

- the mean jet charge which is defined as :

$$Q_{hem} = \frac{\sum_{i=1}^{n} q_i(\vec{p_i}\cdot\vec{e_s})^\kappa}{\sum_{i=1}^{n}(\vec{p_i}\cdot\vec{e_s})^\kappa} \tag{3}$$

  where $q_i$ and $\vec{p_i}$ are the charge and the momentum of particle $i$, $\vec{e_s}$ is the unit vector along the thrust axis, and $\kappa$=0.6;

- the weighted sum of charges of particles identified as kaon candidates: $Q_k = \sum q_i(\vec{p_i}\cdot\vec{e_s})^\kappa$;

- the sum of the charges of tracks having significant impact parameters with respect to the primary event vertex;

- the sum of the charges of the tracks compatible with the primary event vertex;

- the transverse momentum multiplied by the charge of the identified lepton candidate with highest momentum.

Only tracks not included in the $B_s^0$ candidate decay products were used to define the variables in the same hemisphere [2]. They were:

[2] In case of $D_s\ell$ analysis all the $B_s^0$ decay products are identified and removed, for more inclusive analyses this is possible only in a partial way

- the mean jet charge, computed using (3) with $\vec{e_s}$ directed along the reconstructed momentum of the $B_s^0$ candidate;
- the rapidity with respect to the direction of the thrust axis multiplied by the charge of the identified kaon candidate with the highest momentum having a trajectory compatible with the primary vertex,
- the momentum of any reconstructed $\Lambda^0$ candidate multiplied by the charge of the proton from its decay.

The last variable, common to both hemispheres, was the polar angle of the direction of the thrust axis.

## 2.1 Measurement of $\epsilon_{tag}$ in events with an exclusively reconstructed $D^*$

This section presents a measurement of the purity of the tagging at production time, $\epsilon_{tag}$, using the real data.
It has been obtained from the analysis of the $B_d^0 - \overline{B_d^0}$ mixing in events with an exclusively reconstructed $D^{*\pm}$. The $D^{*\pm}$ candidates were selected by reconstructing the decay chain $D^{*+} \to D^0\pi^+$ followed by $D^0 \to K^-\pi^+$ or $D^0 \to K^-\pi^+\pi^0$. The selection criteria rely mainly on the small mass difference between $D^{*+}$ and $D^0$ mesons [2]. The measurement of the $B_d^0 - \overline{B_d^0}$ mixing is performed by correlating *a)* the sign of the $D^{*\pm}$ charge, which tags the B flavour at the time of decay (since $D^{*-}$ in these events are mainly produced from $B_d^0$ and $D^{*+}$ from $\overline{B_d^0}$), with *b)* the global tagging variable, $xefo$, in the hemisphere opposite to the $D^{*\pm}$. If the $B^0$ meson decaying into a $D^{*\pm}$ has oscillated, the $D^{*\pm}$ charge and the variable $xefo$ of the hemisphere opposite to the $D^{*\pm}$ should be of unlike sign; if it has not oscillated, they should be of like sign. The mass difference $\Delta m_d$ between the two physical $\overline{B_d^0}$ states is obtained from the study of the $D^0$ decay distance distribution of unlike and like sign events. Details of the analysis can be found in [2]. The amplitude of the time dependent oscillation is sensitive to the probability of correctly tagging events as unmixed and mixed $B_d^0$. A fit was performed fixing the mass difference $\Delta m_d$ to the world average [3], and leaving $\epsilon_{tag}$ as a free parameter. The fit was repeated for different minimum values of the global tagging variable $xefo$. Results are reported in Table 1, together with the predictions from the Monte Carlo simulation. The fraction of events $f_{events}$ remaining after the cut on the tagging variable is also reported.

The tagging efficiency estimated with the $D^{*\pm}$ sample is consistent within its error with the expectations from the Monte Carlo simulation.

The selected sample of exclusively reconstructed $D^{*\pm}$ still contains a significant fraction of events originating from charm and light flavour quarks. In order to study the tagging variable distribution, the b-tag probability for all tracks of the event was required to be smaller than $10^{-3}$ [4]. The fraction of non-b events in the remaining sample is estimated to be 5%. The distribution of the product between the $D^{*\pm}$ charge and the tagging variable $xefo$ in the hemisphere opposite to the $D^{*\pm}$ is shown in Figure 1, together with the expectations from the Monte Carlo. Another check was performed by selecting events with an exclusively reconstructed $D^{*\pm}$ accompanied by a lepton of opposite charge. This sample is highly enriched in $B_d^0$, but has a limited statistics. However, it allows the study of the tagging variable $xefs$ in the same hemisphere as the $D^{*\pm}$-lepton candidate.

| | Data | | Simulation | |
|---|---|---|---|---|
| | $\epsilon_{tag}$ | $f_{events}$ | $\epsilon_{tag}$ | $f_{events}$ |
| $\|xefo\| > 0.$ | $0.68 \pm 0.02$ | 1. | 0.69 | 1. |
| $\|xefo\| > 0.1$ | $0.69 \pm 0.02$ | 0.88 | 0.71 | 0.89 |
| $\|xefo\| > 0.2$ | $0.71 \pm 0.02$ | 0.77 | 0.736 | 0.78 |

Table 1: *Values of $\epsilon_{tag}$ obtained from the analysis of exclusively reconstructed $D^{*\pm}$ for different cuts on the tagging variable $xefo$. Also reported is the fraction of events remaining after the cut. Expectations from the simulation are also given.*

The distribution of the product between the $D^{*\pm}$ charge and the tagging variable, $xefs$, is shown in Figure 2, together with the expectations from the Monte Carlo. A good agreement is observed.

# 3 The new ($D_s^{\pm}\ell^{\mp}$) analysis

$B_s^0$ meson candidates [3] were selected using an exclusively reconstructed $D_s$ meson correlated with a lepton of opposite charge emitted in the same hemisphere:

$$\overline{B_s^0} \longrightarrow D_s^+ \ell^- \overline{\nu} X.$$

$D_s$ mesons were identified in six non-leptonic and two semileptonic decay modes:

$$\begin{array}{ll}
D_s^+ \longrightarrow \phi\pi^+ & \phi \longrightarrow K^+K^-; \\
D_s^+ \longrightarrow \overline{K}^{*0}K^+ & \overline{K}^{*0} \longrightarrow K^-\pi^+; \\
D_s^+ \longrightarrow \overline{K_S^0}K^+ & \overline{K_S^0} \longrightarrow \pi^+\pi^-; \\
D_s^+ \longrightarrow \overline{K}^{*0}K^{*+} & \overline{K}^{*0} \longrightarrow K^-\pi^+, \quad K^{*+} \longrightarrow \overline{K}_S^0\pi^+; \\
D_s^+ \longrightarrow \phi\pi^+\pi^0 & \phi \longrightarrow K^+K^-; \\
D_s^+ \longrightarrow \phi\pi^+\pi^-\pi^+ & \phi \longrightarrow K^+K^-; \\
D_s^+ \longrightarrow \phi e^+\nu_e & \phi \longrightarrow K^+K^-; \\
D_s^+ \longrightarrow \phi\mu^+\nu_\mu & \phi \longrightarrow K^+K^-.
\end{array}$$

The selection criteria used for the first four hadronic modes and the two semileptonic modes were already presented in [1],[5]. In the following, the description of the selection criteria of the two new channels is given together with those of $\overline{K^{*0}}K^{*+}$ which was significately changed with respect to the one described in [1].

$D_s^+ \longrightarrow \overline{K}^{*0}K^{*+}$

$D_s^+$ candidates were selected by reconstructing $\overline{K}^{*0} \longrightarrow K^-\pi^+$ and $K^{*+} \longrightarrow K_s^0\pi^+$. $K_s^0$ candidates were reconstructed in the mode $K_s^0 \longrightarrow \pi^+\pi^-$ by combining all pairs of oppositely charged tracks and applying the "tight" selection criteria described in [6]. The

[3] Unless explicitely stated otherwise, charge conjugate states are always implied.

$K_s^0$ was then combined with two charged tracks of the same sign, and a third track of opposite charge. If more than one $D_s^+$ candidate could be reconstructed by the same four tracks (by swapping the two pion candidates for example) the $D_s^+$ candidate minimizing the squared mass difference $(M(K^-\pi^+) - M_{PDG}(\overline{K}^{*0}))^2 + (M(K_s^0\pi^+) - M_{PDG}(K^{*+}))^2$ was chosen. If more than one candidate passed the selection criteria only one was kept. Its selection was based on the lepton transverse momentum and the $D_s^+$ momentum.

The three charged tracks were then tested for geometrical compatibility with a single vertex by requiring that the $\chi^2$ of the $D_s$ vertex was smaller then 40. To improve on the vertex position resolution two of the three tracks were required to have at least one VD hit.

To reduce the combinatorial background the following kinematical selection criteria were also applied:

- $p(\pi^+_{\overline{K}^{*0}}) > 1.0\ GeV/c$
- $p(\pi^+_{K^{*+}}) > 1.75\ GeV/c$
- $p(K^-) > 1.5\ GeV/c$
- $p(K_s^0) > 1.0\ GeV/c$
- $p(\overline{K}^{*0}) > 3.5\ GeV/c$
- $p(K^{*+}) > 4.0\ GeV/c$
- $p(D_s^+) > 10\ GeV/c$
- $|M(K^-\pi^+) - M_{PDG}(\bar{K}^{*0})| < 70\ MeV/c^2$
- $|M(K_s^0\pi^+) - M_{PDG}(K^{*+})| < 95\ MeV/c^2$

$D_s \longrightarrow \phi\pi\pi\pi$

The $\phi$ was reconstructed in the decay mode $\phi \longrightarrow K^+K^-$ by taking all possible pairs of oppositely charged tracks requiring that the invariant mass was within 14 $MeV/c^2$ of $M_{PDG}(\phi)$. Neither kaon candidate should be tagged by the RICH as pions ("loose" selection) [6]. Three tracks, each compatible with the pion hypothesis as given by the combined RICH and dE/dX measurements, were then added to the $\phi$ candidate to make a $D_s^+$. Three of the five tracks were required to have at least one VD hit and two of the three pion candidates were required to have a momentum above 1.25 GeV/c. If more than one candidate passed the selection criteria only one was kept. In addition the following selection criteria were made to reduce the combinatorial background:

- $p(K) > 1.75\ GeV/c$
- $p(\pi) > 0.7\ GeV/c$
- $p(\phi) > 4.0\ GeV/c$
- $p(D_s^+) > 9\ GeV/c$

$D_s^+ \longrightarrow \phi\pi\pi^0$

The $\phi$ was reconstructed in the mode $\phi \longrightarrow K^+K^-$ by taking all possible pairs of oppositely charged tracks requiring that the invariant mass was within 11 $MeV/c^2$ of $M_{PDG}(\phi)$. Neither kaon candidate should be tagged by the RICH as pion ("loose" selection) [6]. A third track, compatible with the pion hypothesis as given by the RICH, and a reconstructed $\pi^0$ were then added to the $\phi$ candidate. If more than one candidate passed the selection criteria only one was kept. $\pi^0$ candidates were obtained in two different ways.

All $\gamma\gamma$ pairs in the same hemisphere as the $\phi$ candidate and whose invariant mass was within 20 $MeV/c^2$ of the nominal $\pi^0$ mass were refitted with the nominal $\pi^0$ mass as a constraint. The probability of the fit was required to be larger than 1%. Because of the small opening angle high momentum $\pi^0$s are seen as one cluster in the DELPHI electromagnetic calorimeter HPC. These overlapping photon pairs are tagged by the DELPHI electron and photon reconstruction program ELEPHANT [7], and were used as $\pi^0$s in this analysis. The three charged tracks were tested for geometrical compatibility with a single vertex, and $\chi^2(D_s^+) < 60$ was required. To improve the vertex resolution all three charged tracks were required to have at least one VD hit. In addition the following selection criteria were made to reduce the combinatorial background:

- $p(K) > 1.25\ GeV/c$
- $p(\pi^+) > 0.6\ GeV/c$
- $p(\pi^0) > 3.5\ GeV/c$
- $p(\phi) > 2.5\ GeV/c$
- $p(\mathrm{D_s}^+) > 7\ GeV/c$

In all non-leptonic modes, the measured position of the $\mathrm{D_s}^+$ decay vertex, the $\mathrm{D_s}^+$ momentum, and their measurement errors, were used to reconstruct a $\mathrm{D_s}^+$ particle. A candidate $\mathrm{B_s^0}$ decay vertex was obtained by intercepting this particle with a "loosely" identified lepton [6] (electron or muon) of opposite charge in the same hemisphere (as for hadron identification, leptons were also classified as "tight", "standard" and "loose"). The lepton was required to have a high momentum ($p > 3$ GeV/c) and high transverse momentum ($p_t^{out} > 1.2$ GeV/c) to suppress fake leptons and cascade decays ($b \to c \to \ell^+$) of non-strange B hadrons; the lepton track had also to be associated to at least one hit in the VD. In addition, the following kinematic cuts were applied:

- $3.0 < M(\mathrm{D_s}^{\pm}, \ell^{\mp}) < 5.5$ GeV/c$^2$,
- $p(\mathrm{D_s}^{\pm}, \ell^{\mp}) > 14$ GeV/c,
- $\chi^2(\mathrm{B_s^0}$ vertex$) < 20$.

In the $\mathrm{D_s}^+$ mass region, a clear excess of "right-sign" combinations ($\mathrm{D_s}^{\pm}\ell^{\mp}$) over "wrong-sign" combinations ($\mathrm{D_s}^{\pm}, \ell^{\pm}$) was observed in each channel. Figure 4 and 5 show the signals obtained with 94-95 data. The mass distribution for non-leptonic decays was fitted using two Gaussian distributions of equal widths to account for the $\mathrm{D_s}$ and $\mathrm{D}^+$ signals and an exponential for the combinatorial background. The $\mathrm{D}^+$ mass was fixed to the nominal value of 1.869 GeV/c$^2$ [8]. For $\mathrm{D_s}$ semileptonic decay modes, the $\mathrm{K^+K^-}$ invariant mass distribution for "right sign" events was fitted with a Breit–Wigner distribution to account for the signal, and a polynomial function to describe the combinatorial background.

Table 2 gives the measured number of events (background subtracted) in the $\mathrm{D_s}^+$ signal and the ratio between the numbers of combinatorial background events and the total.

| $D_s$ decay modes | Estimated signal 91-93 data | Estimated signal 94-95 data |
|---|---|---|
| $D_s \longrightarrow \phi\pi^+$ | $15 \pm 5$ $(40 \pm 8\%)$ | $45 \pm 9$ $(40 \pm 13\%)$ |
| $D_s \longrightarrow \overline{K}^{\star 0}K^+$ | $13 \pm 5$ $(36 \pm 9\%)$ | $35 \pm 9$ $(47 \pm 13\%)$ |
| $D_s \longrightarrow \overline{K^0_S}K^+$ | $18 \pm 6$ $(38 \pm 13\%)$ | $17 \pm 5$ $(23 \pm 5\%)$ |
| $D_s \longrightarrow \overline{K}^{\star 0}K^{\star +}$ | - | $19 \pm 5$ $(20 \pm 9\%)$ |
| $D_s^+ \longrightarrow \phi\pi^+\pi^0$ | - | $19 \pm 6$ $(43 \pm 11\%)$ |
| $D_s^+ \longrightarrow \phi\pi^+\pi^-\pi^+$ | - | $11 \pm 5$ $(37 \pm 15\%)$ |
| $D_s \longrightarrow \phi\ell^+\nu$ | $19 \pm 9$ $(38 \pm 12\%)$ | $42 \pm 11$ $(21 \pm 7\%)$ |

Table 2: *Numbers of* $D_s$ *signal events and ratios between the numbers of combinatorial background events and the total (in parentheses) for different* $D_s$ *decay channels. The level of the combinatorial background was evaluated using a mass interval of* $\pm 2\sigma$ $(\pm 1\Gamma)$ *centred on the measured* $D_s$ $(\phi)$ *mass.*

## 3.1 Results from the $(D_s^{\pm}\ell^{\mp})$ analysis

The time resolution for the fraction of events in which $D_s$ mesons are expected to come from $B_s$ semileptonic decays is summarized in Table 3 for the different $D_s$ decay modes corresponding to the 94-95 data analysis

| $D_s$ decay channel | First Gaussian fraction (%) | First Gaussian ($\sigma_1$) resolution (ps) | Second Gaussian ($\sigma_2$) resolution (ps) |
|---|---|---|---|
| $K^0_S K^+$ (94-95) | 68 | $\sqrt{(0.155^2 + 0.071^2\ t^2)}$ | $3.6\sigma_1$ |
| other non-leptonic (94-95) | 83 | $\sqrt{(0.108^2 + 0.074^2\ t^2)}$ | $2.9\sigma_1$ |
| $\phi\ell^+\nu$ (94-95) | 76 | $\sqrt{(0.128^2 + 0.069^2\ t^2)}$ | $3.6\sigma_1$ |

Table 3: *(*$D_s^{\pm}\ell^{\pm}$ *analysis) : time resolution for different* $D_s$ *decay modes parametrized using the sum of two Gaussian distributions. The width of the narrower Gaussian depends on the generated proper time. The width of the other distribution has been taken to be proportional to the width of the narrower one.*

The fitting procedures are explained in the previous publication [1].
The average tagging purity of the $x_{tag}$ variable, calculated on MC for true $\overline{B_s}^0 \rightarrow D_s^+\ell^-\nu X$ decays, was $74.2 \pm 0.3\%$. In the fitting procedure the information of the tagging purity is used on an event by event basis, using the distribution of the discriminating variable $x_{tag}$ (Figure 6). In this case the tagging purity of the $x_{tag}$ variable, calculated on MC for true $\overline{B^0_s} \rightarrow D_s^+\ell^-\nu X$ decays, is equivalent to $78.0 \pm 0.3\%$.
The $D_s\ell$ events selected in the 91-93 data ( see Figure 3 and Table 2) were also included. For this set of data the new procedure of tagging at the production time was also implemented. In this case the average tagging purity of the $x_{tag}$ variable, obtained from the simulation, was $71.8 \pm 1.1\%$, which correspond to a tagging purity on an event by event

basis of $75.7 \pm 1.1\%$
The plot of the variation of the oscillation amplitude $A$ as a function of $\Delta m_s$ combining the two sets of data ( 91-93 and 94-95) is shown in (Figure 7). Considering only statistical uncertainties, the lower limit on $\Delta m_s$ is :

$$\Delta m_s > 8.4 \text{ ps}^{-1} \quad \text{at } 95\% \text{ C.L.}. \tag{4}$$

The exclusion probability for this limit is 47%. The expected limit at 50% exclusion probability corresponds to $\Delta m_s > 8.1$ ps$^{-1}$.

## 3.2 Study of systematic uncertainties

Systematic uncertainties were evaluated by varying the parameters which were kept constant in the fit according to their measured or expected errors.

- Systematics from the tagging purity.
  The studies done in Sec. 2.1 show that using the tagging variables in the opposite hemisphere and asking $|xefo| > 0.$, the agreement between the real and simulated data on the tagging purity is $\epsilon_{tag}(DATA) - \epsilon_{tag}(MC)$ = -0.01 $\pm$ 0.02. It was also shown that there is a good agreement between the real data and the simulated distributions for the tagging purities on the opposite and the same hemisphere. This test makes confident on the use of the tagging purity on the event by event basis. The evaluation of systematics coming from the control of the tagging purity was done in the following way. The functions which give the tagging probabilities versus the discriminating variable for $b$ and $\overline{b}$ were modified to obtain a variation on the averaged value of the tagging purity of $\pm 3.0\%$. This variation was done by keeping the functions fixed for the values -1,0,1 of the discriminating variable.

- Systematics from the background level.
  $f_{bkg.}$ was varied according to the statistical uncertainty of the fitted combinatorial background fractions present in the different $D_s$ or $K^+K^-$ mass distributions.

- Systematics from the $B_s^0$ purity.
  $f_{B_s}$ is the fraction of events in which $D_s$ mesons are expected to come from $B_s$ semileptonic decays. $f_{D_sD}$ is the expected fraction of cascade decays $B \to \overline{D}^{(\star)} D_s^{(\star)+} X$ followed by the semileptonic decay $\overline{D}^{(\star)} \to \ell^- \overline{\nu} X$ which gives right-sign $D_s^{\pm} \ell^{\mp}$ pairs. After the selection cuts, the relative fractions are $f_{D_sD}/f_{B_s} = 0.106 \pm 0.020$ for non-leptonic $D_s$ decays and $f_{D_s\ D}/f_{B_s} = 0.102 \pm 0.022$ for semileptonic decays, where the errors take into account both the statistical error from simulation and the errors on measured branching ratios. The errors on these two fractions were taken to evaluate the systematics.

- Systematics from the resolution on the B decay proper time.
  The systematic error coming from uncertainties on the resolution functions was evaluated by varying by 10% the two parameters describing the quadratic time dependence of the narrower Gaussian (see Table 3). A variation of 10% of the resolution for background events was also considered.

Including these systematic uncertainties does not change the excluded $\Delta m_s$ regions significantly.

# 4 The new ($D_s^{\pm}h^{\mp}$) analysis

This approach is similar to the ($D_s^{\pm}\ell^{\mp}$) analysis but, instead of the lepton, it uses a charged hadron. It provides larger statistics but suffers from an ambiguity in the choice of the hadron and from a lower $B_s^0$ purity. This approach was already used to measure the $B_s^0$ lifetime and several details of the analysis are given in [5]. In the present measurement only data from 94-95 were used and the $D_s$ was selected in the following decay modes:

$$
\begin{array}{ll}
D_s^+ \longrightarrow \phi\pi^+ & \phi \longrightarrow K^+K^-; \\
D_s^+ \longrightarrow \overline{K}^{*0}K^+ & \overline{K}^{*0} \longrightarrow K^-\pi^+; \\
D_s^+ \longrightarrow \overline{K_S^0}K^+ & \overline{K_S^0} \longrightarrow \pi^+\pi^-; \\
D_s^+ \longrightarrow \overline{K}^{*0}K^{*+} & \overline{K}^{*0} \longrightarrow K^-\pi^+, \quad K^{*+} \longrightarrow \overline{K_S^0}\pi^+;
\end{array}
$$

The selection criteria for the first two channels was already given in [1],[5]. The selection of for the two new modes is explained in the following.
The common feature for the last two $D_s$ decay modes is the presence of a $\overline{K_S^0}$ which was searched for using the standard DELPHI procedure [6]. For the $D_s \to \overline{K_S^0}K^+$ decay mode the momentum of the $\overline{K_S^0}$ was requested to be above 4 $GeV/c$. The momentum of the $K^+$ should be greater than 3.5 $GeV/c$ and it should be identified at least as a "standard" kaon [6]. For the $D_s \to \overline{K}^{*0}K^{*+}$ decay mode no explicit cuts on the momenta of pseudoscalar mesons were imposed but the momenta of the vector mesons $\overline{K}^{*0}$ and $K^{*+}$ were requested to be larger than 5.5 $GeV/c$. The invariant mass of the vector meson candidates had to be within $\pm$ 40 $MeV/c^2$ of their nominal mass values. It was also requested at least "standard" indentification for the charged kaon from $\overline{K}^{*0}$ decay. In the $D_s$ decay channels, all particles which were not explicitly identified as protons, kaons or leptons, were classified as pions. The $D_s$ decay length had to be positive and the $\chi^2$–probability of the fitted $D_s$ vertex larger than $10^{-4}$.

The selection of the hadron accompanying the $D_s$ candidate is based on an impact parameter technique. A sample of tracks coming predominantely from B hadron decays is preselected using both the impact parameter with respect to the primary vertex and to the $D_s$ vertex. The hadron was then searched for amongst the preselected tracks in the event requiring that its charge is opposite to that of the $D_s$ and that it has the highest momentum among the candidates opposite in charge to the $D_s$. Details of the preselection as well as of the hadron selection are given in [5].

Figure 8 shows the $D_s^+$ signals after the accompanying hadron selection, for the four decay modes. The mass distribution for non-leptonic decays was fitted using two Gaussian distributions of equal widths to account for the $D_s$ and $D^+$ signals and an exponential for the combinatorial background. The $D^+$ mass was fixed to the nominal value of 1.869 GeV/c$^2$ [8]. Table 4 gives the measured number of events (background subtracted) in the $D_s^+$ signal and the ratio between the numbers of combinatorial background events and the total.

| $D_s$ decay modes | Estimated signal in 94-95 data |
|---|---|
| $D_s \longrightarrow \phi\pi^+$ | $341 \pm 30$ $(0.459 \pm 0.027\%)$ |
| $D_s \longrightarrow \overline{K}^{*0}K^+$ | $185 \pm 22$ $(0.508 \pm 0.034\%)$ |
| $D_s \longrightarrow \overline{K^0_S}K^+$ | $70 \pm 14$ $(0.489 \pm 0.06\%)$ |
| $D_s \longrightarrow \overline{K}^{*0}K^{*+}$ | $61 \pm 17$ $(0.592 \pm 0.063\%)$ |

Table 4: *($D_s^{\pm}h^{\mp}$) analysis : numbers of $D_s$ signal events and ratios between the numbers of combinatorial background events and the total (in parentheses) for different $D_s$ decay channels. The level of the combinatorial background was evaluated using a mass interval of $\pm 2\sigma$.*

## 4.1 Results from the ($D_s^{\pm}h^{\mp}$) analysis

The procedure to evaluate the sample composition is explained in details in [5]. Following this procedure the sample composition is:

- fraction of $B_s^0$ with a $D_s$ : $f_{B_s^0} = 40.1 \pm 3.6$ %
- fraction of $B_s^0$ with two $D_s$ : $f_{D_s,D_s} = 16.4 \pm 1.7$ %
- fraction of non-$B_s^0$ with one $D_s$ : $f_{D_s,D} = 7.4 \pm 0.8$ %
- fraction of non-$B_s^0$ with two D : $f_{B_s} = 36.2 \pm 3.9$ %

The time resolution for the fraction of $B_s^0$ decaying into one $D_s$ is summarized in Table 5 for different $D_s$ decay modes corresponding to the 94-95 data analysis.

| $D_s$ decay Channel | First Gaussian fraction (%) | First Gaussian ($\sigma_1$) resolution (ps) | Second Gaussian ($\sigma_2$) resolution (ps) |
|---|---|---|---|
| $K^0_S K^+$ (94-95) | 50 | $\sqrt{(0.177^2 + 0.168^2\ t^2)}$ | $3.2\sigma_1$ |
| other non-leptonic (94-95) | 60 | $\sqrt{(0.138^2 + 0.086^2\ t^2)}$ | $3.8\sigma_1$ |

Table 5: *($D_s^{\pm}h^{\mp}$) analysis : time resolution for different $D_s$ decay modes parametrized using the sum of two Gaussian distributions. The width of the narrower Gaussian depends on the generated proper time. The width of the other distribution has been taken to be proportional to the width of the narrower one.*

The average tagging purity of the $x_{tag}$ variable, calculated on MC for true $\overline{B_s}^0 \rightarrow D_s^+h^-X$ decays, was $71.4 \pm 0.4\%$. In the fitting procedure the information of the tagging purity is used on an event by event basis, using the distribution of the discriminating variable $x_{tag}$. In this case the tagging purity of the $x_{tag}$ variable is equivalent to $74.0 \pm 0.3\%$. The plot of the variation of the oscillation amplitude $A$ as a function of $\Delta m_s$ is shown in (Figure 9). Considering only statistical uncertainties, the lower limit on $\Delta m_s$ is :

$$\Delta m_s > 2.0 \text{ ps}^{-1} \quad \text{at } 95\% \text{ C.L.}. \tag{5}$$

The exclusion probability for this limit is 36%. The expected limit at 50% exclusion probability corresponds to $\Delta m_s > 1.2$ ps$^{-1}$.

### 4.2 Study of systematic uncertainties

Systematic uncertainties were evaluated by varying the parameters which were kept constant in the fit according to their measured or expected errors.

- Systematics from the tagging purity.
  The followed procedure is described in Sec. 3.2. The tagging purities for the other categories were used on average and fixed at the values found in the simulation. A conservative variation of $\pm$ 3 % on the expected tagging purity for the signal was considered to evaluated the systematics.

- Systematics from the background level.
  $f_{bkg.}$ was varied according to the statistical uncertainty of the fitted combinatorial background fractions present in the different $D_s$ mass distributions.

- Systematics from the $B_s^0$ purity.
  The $B_s^0$ purity was varied according to the value given in Sec 4.1 : $f_{B_s^0} = 40.1 \pm 3.6$ %. This error is mainly coming from the error on the $B_s^0$ production on jets and on the measured branching ratios for the different processes contributing to the final sample.

- Systematics from the expected resolution on the B decay proper time.
  The procedure is described in Sec. 3.2

Including the systematics uncertainties the lower limit on $\Delta m_s$ is:

$$\Delta m_s > 1.9 \text{ ps}^{-1} \quad \text{at } 95\% \text{ C.L.}.$$

## 5 The $(\phi\ell^{\mp})$ analysis

This analysis is more inclusive and concerns events where a high $p_t^{out}$ lepton is accompanied by a $\phi$ meson in the same jet. The selection of a high $p_t^{out}$ lepton enriches the sample in direct semileptonic decays and the presence of the $\phi$ increases its $B_s^0$ purity. This approach was already used to measure the $B_s^0$ lifetime and details of the analysis are described in [5]. In the present measurement only data from 94-95 were used. The invariant mass distribution for $K^+K^-$ is shown in Figure 10. A signal of 557 $\pm$ 49 $\phi$ mesons has been fitted using a Breit-Wigner distribution. The mass, $m_\phi = 1.0198 \pm 0.0003$ GeV/c$^2$, and the width, $\Gamma = 5.7 \pm 0.6$ MeV, are in agreement with the simulation predictions. Within $\pm 1\Gamma$ there are 406 $\pm$ 36 events and the combinatorial background amounts to $f_{comb} = 0.399 \pm 0.035$.
The reconstructed tracks with opposite charged of the identified lepton were considered as candidates for an extra $D_s$ decay hadron. The track with the highest momentum, projected along the $\phi$ direction was selected. A $D_s$ pseudotrack was then reconstructed using the kaon pair and the hadron candidate and used to form a common vertex (the candidate

$B_s^0$ vertex) with the identified lepton. The mass of the $K^+K^-h$ system combination was required to be lower than the $D_s$ mass and once the lepton had been attached, to not exceed the $B_s$ mass. Finally to improve the momentum and the direction resolutions for the B candidate, the electromagnetic energy was added to the $K^+K^-h$ system until the mass of the system exceeded 2.5 GeV/$c^2$. After this procedure the signal yielded to 429 $\pm$ 52 $\phi$ candidates fitted with a Breit-Wigner distribution. Within one gamma there are 319 $\pm$ 52 events and the fraction of the combinatorial background was $f_{comb} = 0.420 \pm 0.050$

## 5.1 Results from the $(\phi\ell^{\mp})$ analysis

The sample composition is calculated following the procedure explained in [5]. Inside the $\phi$ signal the proportion of semileptonic $B_s^0$ decays is found to be : $f_{B_s^0} = 0.42 \pm 0.07$. The time resolution for these events has been parametrized using a narrow and two wider Gaussians. The narrow Gaussian contains 45 % of the signal and its time dependent width is described by $\sqrt{(0.135^2 + 0.054^2\ t^2)}$

The plot of the variation of the oscillation amplitude $A$ as a function of $\Delta m_s$ is shown in (Figure 11). Considering only statistical uncertainties, the lower limit on $\Delta m_s$ is :

$$\Delta m_s > 2.0 \text{ ps}^{-1} \quad \text{at } 95\% \text{ C.L.}. \tag{6}$$

The exclusion probability for this limit is 37%. The expected limit at 50% exclusion probability corresponds to $\Delta m_s > 1.0$ ps$^{-1}$.

## 5.2 Study of systematic uncertainties

- Systematics from the tagging purity.
  The followed procedure is described in Sec. 3.2. The tagging purities for the other categories were used in average and fixed to the values found in the simulation. A conservative variation of $\pm$ 3 % on the expected tagging purity for the signal was considered to evaluate the systematics.

- Systematics from the background level.
  $f_{bkg.}$ was varied according to the statistical uncertainty of the fitted combinatorial background fractions present in the different $D_s$ mass distributions.

- Systematics from the $B_s^0$ purity.
  The $B_s^0$ purity was varied according to the value given in Sec 5.1 : $f_{B_s^0} = 0.42 \pm 0.07$. This error is mainly coming from the error on the $B_s^0$ production on jets and on the measured branching ratios for the different processes contributing to the final sample.

- Systematics from the resolution on the B decay proper time.
  The procedure is described in Sec. 3.2

Including the systematics uncertainties the lower limit on $\Delta m_s$ is:

$$\Delta m_s > 1.9 \text{ ps}^{-1} \quad \text{at } 95\% \text{ C.L.}.$$

# 6 Combined limit on $\Delta m_s$

Results presented in the previous sections have been combined and previous analysis from $\ell - Q_{hem.}$ and $\ell - \ell$ channels have been included [1], taking into account correlations between the events samples and between systematic uncertainties affecting the different amplitude measurements, according to the procedure defined by the LEP Working Group on the B Oscillations [3]. The resulting amplitude distribution is shown in Figure 12 and corresponds to the limit:

$$\Delta m_s > 8.5 \text{ ps}^{-1} \text{ at } 95\% \text{ C.L.} \quad (7)$$

The exclusion probability for this limit is 50%, so the combined limit corresponds to the combined sensitivity.

# 7 Updated measurement of the $B_s^0$ lifetime

The $B_s^0$ lifetime measurement has been updated, with respect to the previous analysis [5], using three selected events samples presented in this paper. In each analysis the $B_s^0$ lifetime and the proper time distribution for the background have been fitted simultaneously using events in the signal region ($\pm 2\sigma$ [$\pm 1\Gamma$] from the $D_s^+$ [$\phi$] mass peak) and events from combinatorial background samples ("side–bands" events in $D_s^{\mp} h^{\pm}$ and $\phi\ell^{\pm}$ analyses and "side–bands" and "wrong sign" events in ($D_s^{\mp}\ell^{\pm}$) analysis). Results are summarized in Table 6 and the different sources of systematics are given.

In each analysis the fitting method has been checked on pure $B_s^0$ simulated events generated with a lifetime of 1.6 *ps*:

$$\begin{aligned}
\tau_{B_s^0}(D_s^{\pm}\ell^{\mp})^{MC} &= 1.601 \pm 0.025 \quad \Delta^{shift} = +0.001 \pm 0.025 \ ps \\
\tau_{B_s^0}(D_s^{\pm}h^{\mp})^{MC} &= 1.630 \pm 0.050 \quad \Delta^{shift} = +0.030 \pm 0.050 \ ps \\
\tau_{B_s^0}(\phi\ell^{\mp})^{MC} &= 1.667 \pm 0.041 \quad \Delta^{shift} = +0.067 \pm 0.041 \ ps
\end{aligned}$$

The difference ($\Delta^{shift}$) between the generated and the fitted lifetimes was interpreted as a possible remaining bias due to limitations of the model used in the fit or to uniformity in acceptance induced by selection cuts. The measured lifetimes have been corrected for these differences and the statistical error on $\Delta^{shift}$ has been added to the systematic uncertainties. The corrected $B_s^0$ lifetime measurements for the three analyses turn out to be :

$$\begin{aligned}
\tau_{B_s^0}(D_s^{\pm}\ell^{\mp}) &= 1.44^{+0.16}_{-0.14}(stat.) \pm 0.05(syst.) \ ps \\
\tau_{B_s^0}(D_s^{\pm}h^{\mp}) &= 1.52^{+0.23}_{-0.22}(stat.) \pm 0.12(syst.) \ ps \\
\tau_{B_s^0}(\phi\ell^{\mp}) &= 1.33^{+0.25}_{-0.23}(stat.) \pm 0.10(syst.) \ ps
\end{aligned}$$

The results of the lifetime fits for the three samples are shown in Figure 13, Figure 14 and Figure 15.
In a previous publication [5] a measurement of the $B_s^0$ meson lifetime was obtained using an inclusive $D_s^+$ sample.To combine this measurement with those presented in this note statistical correlations have been taken into account as well as the common systematic

(from branching ratios and lifetimes). The full covariance matrix was calculated and the combined $B_s^0$ meson lifetime was found to be:

$$\tau_{B_s^0} = (1.46 \pm 0.11)\ ps$$

# References

[1] DELPHI Coll. "Search for $B_s^0 - \overline{B_s^0}$ oscillations" submitted to CERN-PPE ( given in the Addendum ).

[2] P. Aarnio, et. al. (DELPHI Collaboration), Measurement of $B_d^0 - \bar{B}_d^0$ oscillations, CERN/PPE 97–51 (12 May 1997), (Submitted to Zeit. f. Physik C)

[3] The LEP Oscillations Working Group " LEP Combined Results on $B^0$ Oscillations " LEPBOSC NOTE 97/1

[4] G.V. Borisov and C. Mariotti, Nucl. Instr. and Meth. **A372** (1996) 181.

[5] DELPHI Coll., P. Abreu et al., Z. Phys. **C71** (1996) 11.

[6] P. Abreu. et al., DELPHI Coll.,"Performance of the DELPHI Detector", CERN-PPE/95-159, submitted to Nucl. Instr. and Methods

[7] M. Feindt, C. Kreuter and O. Podobrin , ELEPHANT Reference Manual DELPHI-Note 96-82 PROG 217 ( 1996)

[8] Particle Data Group, R. M. Barnett et al., Phys. Rev. **D54** (1996) 1.

| $(D_s^{\pm}\ell^{\mp})$ **analysis** | |
|---|---|
| $\tau_{B_s^0}$ | $1.44^{+0.16}_{-0.14}(stat.) \pm 0.05(syst.)\ ps$ |
| $f_{comb}$, $f_{\mathrm{D}_s D}$, $f_{rif}$ | $+0.040 - 0.040$ |
| $\tau_{B^+}$, $\tau_{B_d^0}$, $\tau_{\Lambda_b}$ | $+0.005 - 0.005$ |
| $t$ resolution | $+0.010 - 0.010$ |
| Syst. tot | $+0.040 - 0.040$ |
| $D_s^{\pm}h^{\mp}$ **analysis** | |
| $\tau_{B_s^0}$ | $1.55^{+0.23}_{-0.22}(stat.) \pm 0.11(syst.)\ ps$ |
| $f_{comb}$, $f_{charm}$, $f_{B_{(not\ s)}}$ | $+0.098 - 0.099$ |
| $\tau_{B^+}$ | $+0.016 - 0.017$ |
| $\tau_{B_d}$ | $+0.015 - 0.015$ |
| $\tau_{\Lambda_b}$ | $+0.008 - 0.008$ |
| $t$ resolution | $+0.030 - 0.030$ |
| Syst. tot | $+0.011 - 0.011$ |
| $\phi\ell^{\mp}$ **analysis** | |
| $\tau_{B_s^0}$ | $1.40^{+0.26}_{-0.24}(stat.) \pm 0.09(syst.)\ ps$ |
| Sample composition | $+0.057 - 0.057$ |
| $\tau_{B^+}$ | $+0.016 - 0.016$ |
| $\tau_{B_d}$ | $+0.017 - 0.017$ |
| $\tau_{\Lambda_b}$ | $+0.004 - 0.004$ |
| $t$ resolution param. | $+0.046 - 0.046$ |
| Energy estimation | $+0.049 - 0.049$ |
| Syst. tot | $+0.09 - 0.09$ |

Table 6: *Results on $B_s$ lifetime*

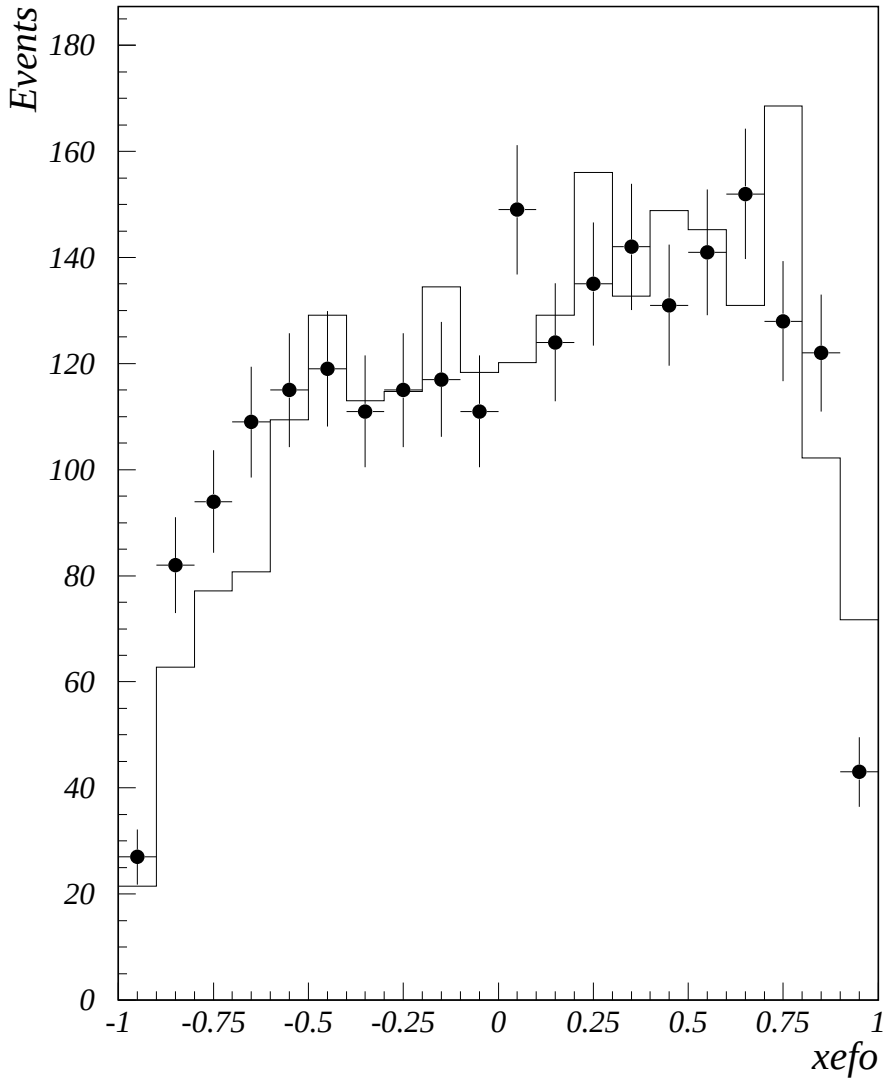


Figure 1: *Distribution of the global tagging variable xefo in the hemisphere opposite to the $D^{*\pm}$ candidate. The full dots with the error bars represent the data. The histogram is the Monte Carlo prediction.*

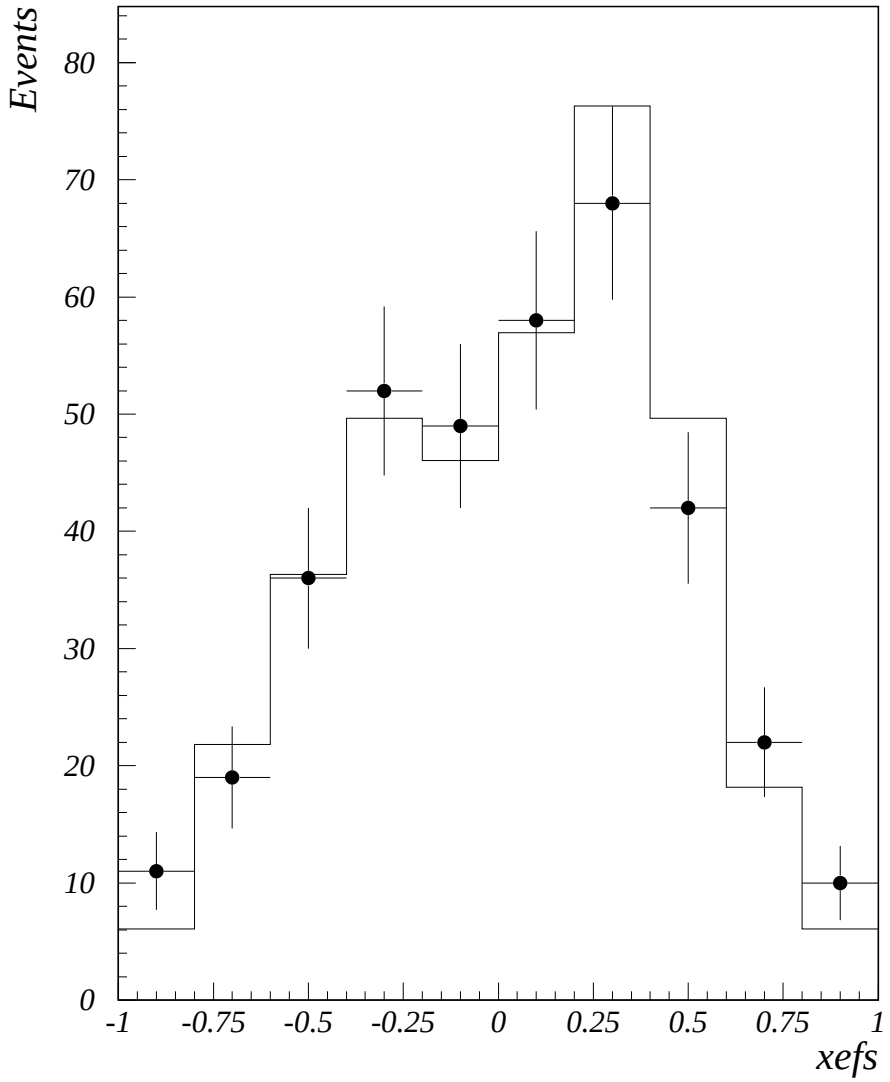


Figure 2: *Distribution of the global tagging variable $xefs$ in the same hemisphere as the $D^{*\pm}$-lepton candidate. The full dots with the error bars represent the data. The histogram is the Monte Carlo prediction.*

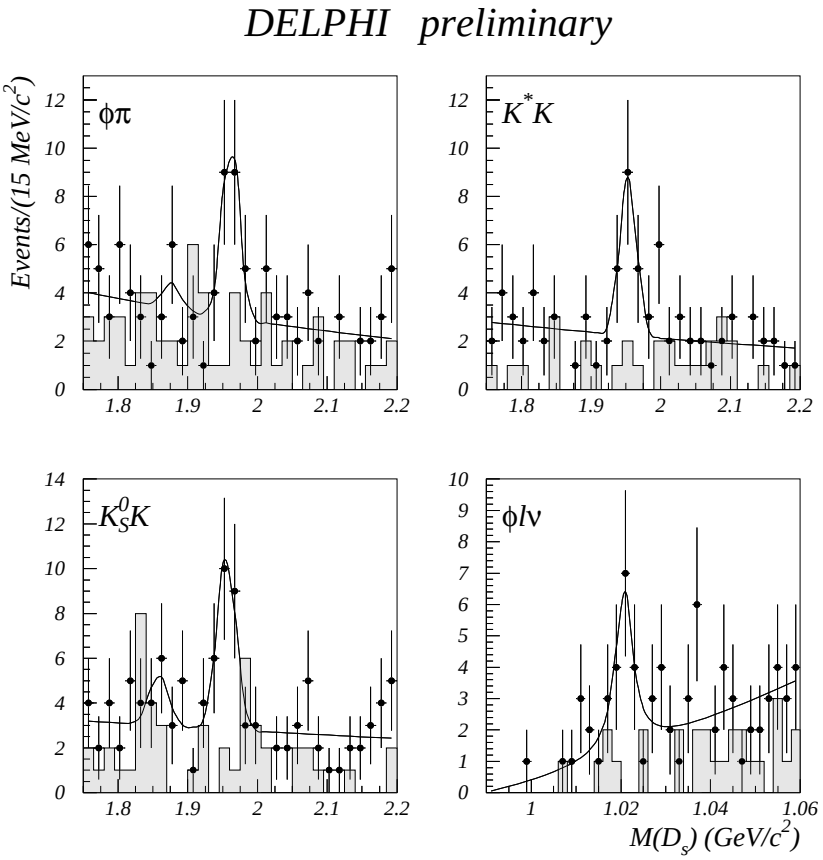


Figure 3: *($D_s^{\pm}\ell^{\mp}$) analysis on 91-93 data:*
*The first three plots show the invariant mass distributions for* $D_s$ *candidates in three non-leptonic decay modes. The fourth plot shows the* $K^+K^-$ *invariant mass distribution for* $D_s$ *candidates selected in the two semileptonic decay modes. The corresponding distribution for wrong-sign combinations are given by the* shaded histograms. *The curves show the fits described in the text.*

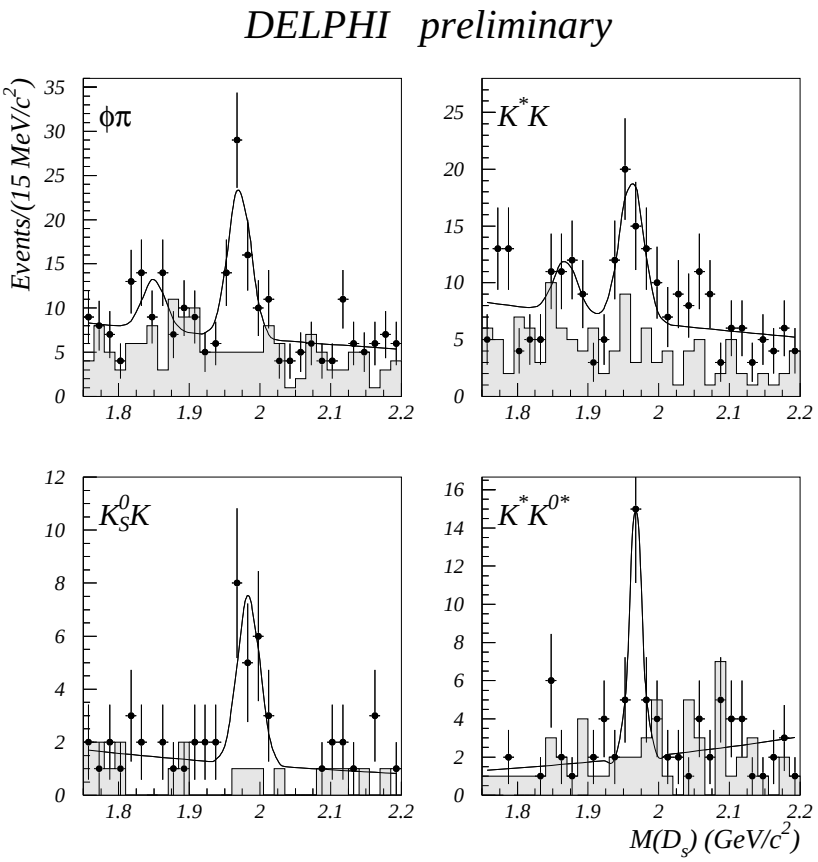


Figure 4: *($D_s^{\pm}\ell^{\mp}$) analysis on 94-95 data:*
*The four plots show the invariant mass distributions for* $D_s$ *candidates in four non-leptonic decay modes ($\phi\pi$, $K^{\star}K$, $K^0K$, $K^{\star 0}K^{\star +}$). The corresponding distribution for the wrong-sign combinations are given by the* shaded histograms. *The curves show the fits described in the text.*

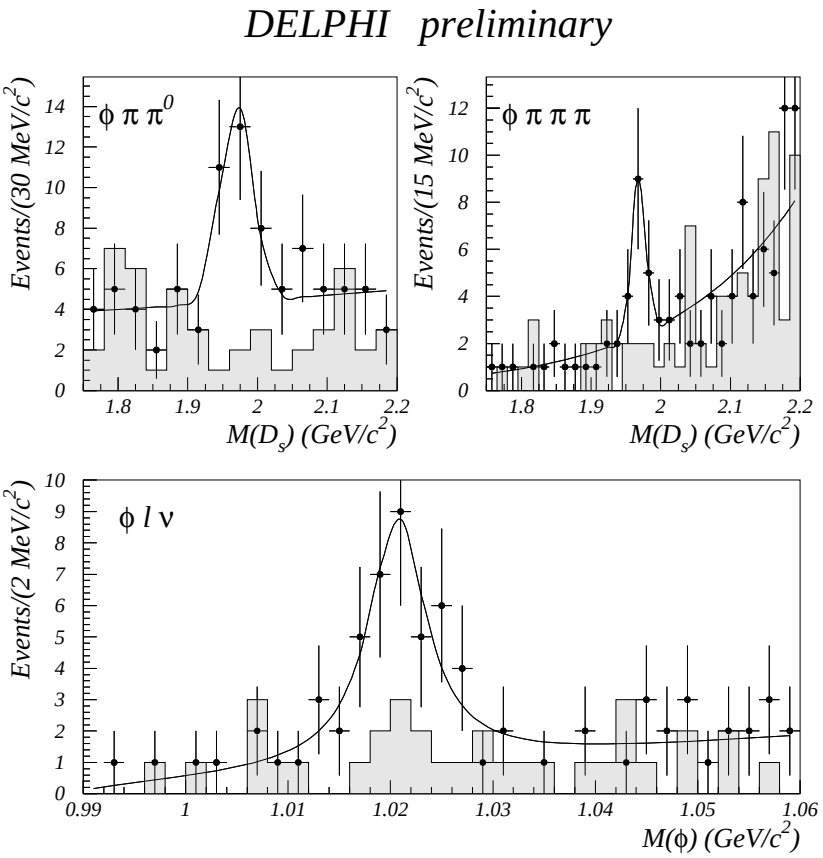


Figure 5: *($D_s^{\pm}\ell^{\mp}$) analysis on 94-95 data:*
*The first two plots show the invariant mass distributions for* $D_s$ *candidates in two non-leptonic decay modes ($\phi\pi\pi^0$, $\phi\pi\pi\pi$). The third plot shows the $K^+K^-$ invariant mass distribution for* $D_s$ *candidates selected in the two semileptonic decay modes. The corresponding distribution for wrong-sign combinations are given by the* shaded histograms. *The curves show the fits described in the text.*

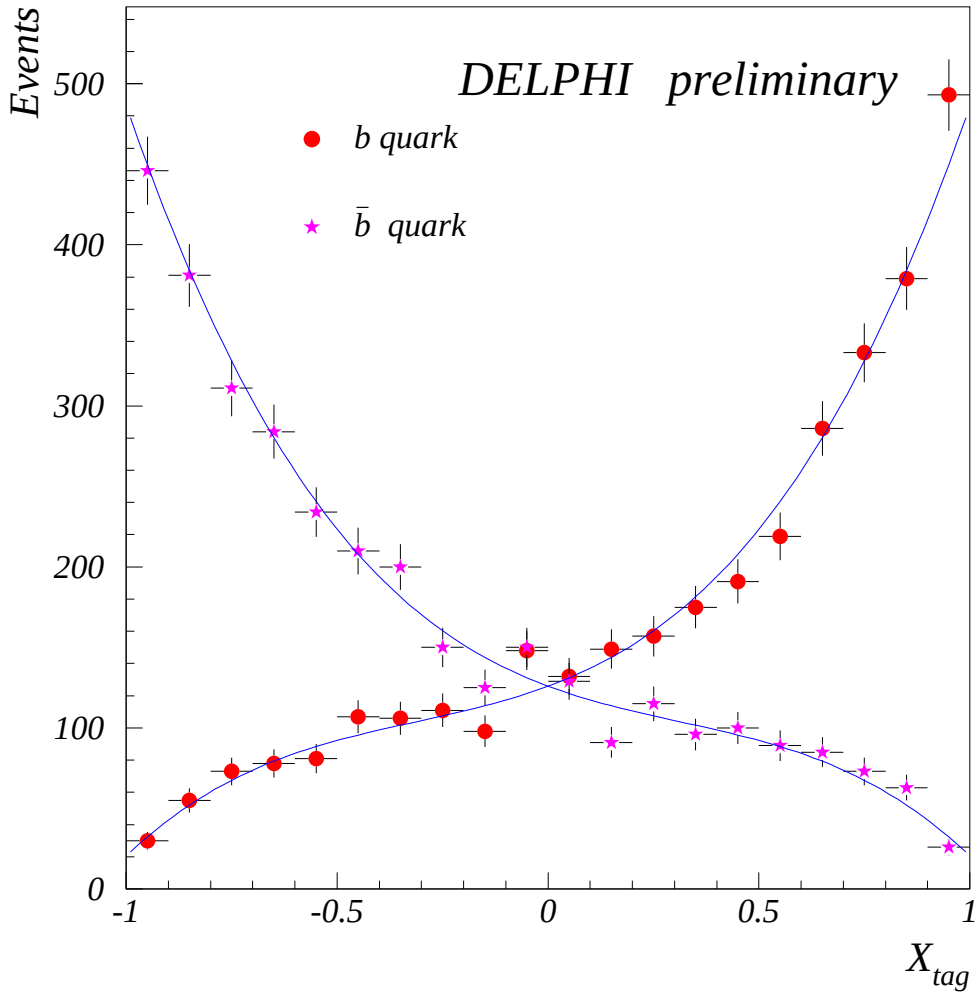


Figure 6: $(D_s^{\pm}\ell^{\mp})$ *analysis:*
*Distributions of the tagging variable* $x_{tag}$ *for* $b$ *and* $\bar{b}$ *initial states. They have been calculated using a dedicated simulated sample of events of the type* $\overline{B_s^0} \to D_s^+\ell^-\nu X$. *The parametrization used in the amplitude fit is superimposed.*

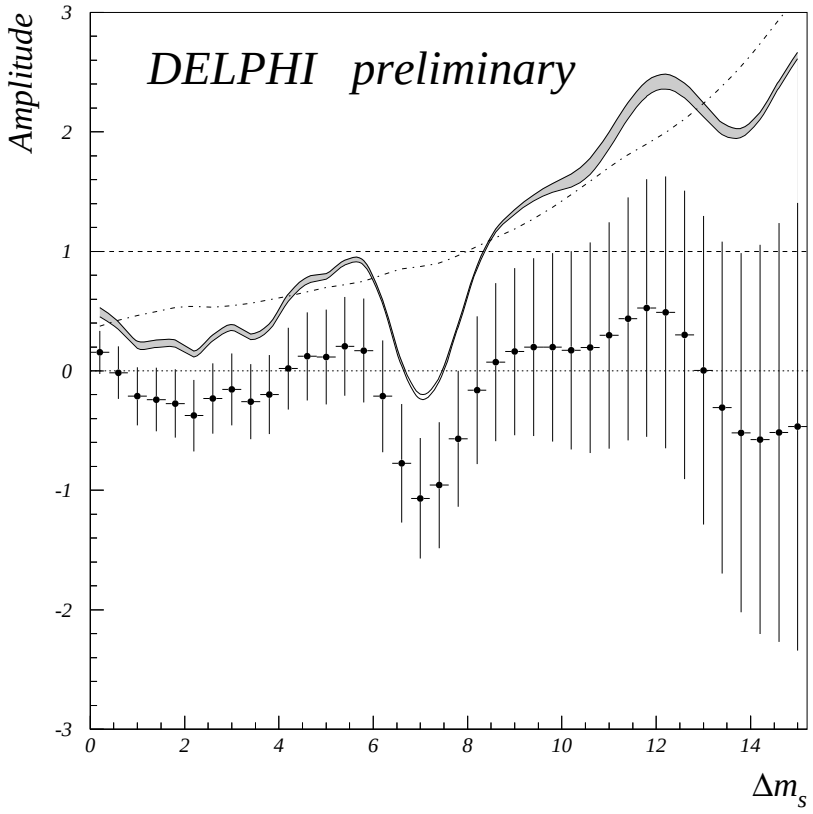


Figure 7: *($D_s^{\pm}\ell^{\mp}$) analysis on 91-95 data:*
*Variation of the oscillation amplitude A as a function of $\Delta m_s$. The lower continuous line corresponds to $A + 1.645\sigma_A$ where $\sigma_A$ includes statistical uncertainties only, while the shaded area shows the contribution from systematics. The dashed-dotted line corresponds to the sensitivity curve. The lines at A=0 and A=1 are also given*

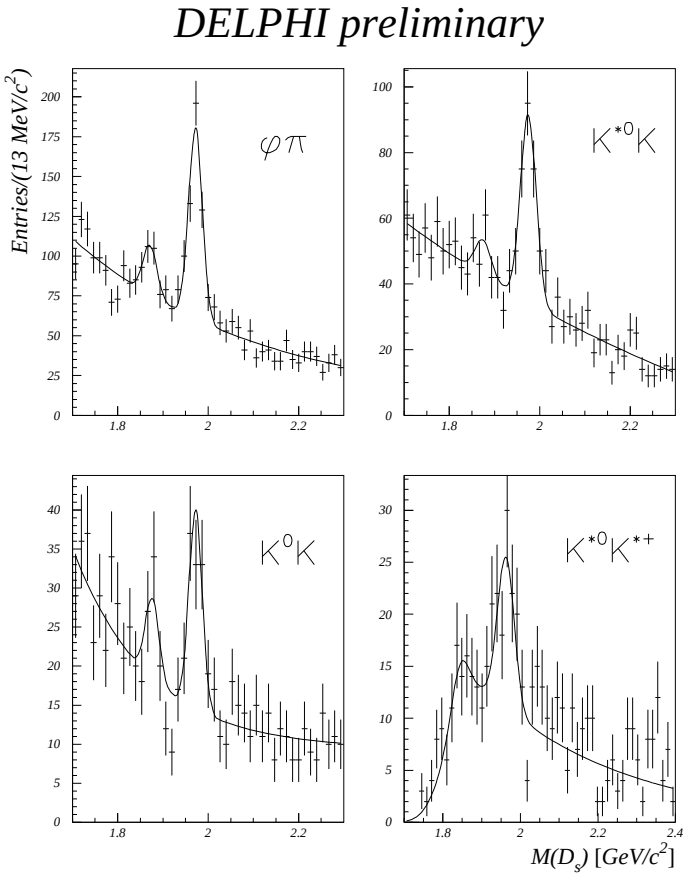


Figure 8: *($D_s^{\pm}h^{\mp}$) analysis on 94-95 data:*
*The plots show the invariant mass distributions for* $D_s$ *candidates in four non-leptonic decay modes as in Table. 4. The selected* $D_s$ *candidates are accompanied by a hadron of opposite electric charge, measured, in the same event hemisphere. The curves show the fits described in the text.*

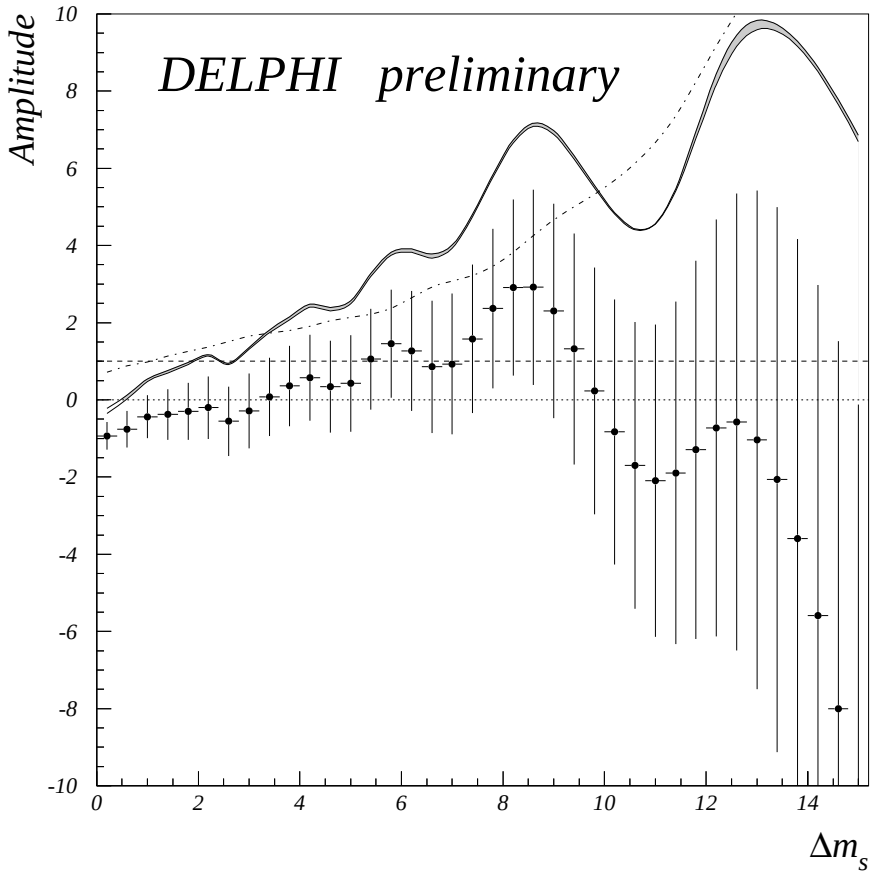


Figure 9: *($D_s^{\pm}h^{\mp}$) analysis on 94-95 data:*
*Variation of the oscillation amplitude A as a function of $\Delta m_s$. The lower continuous line corresponds to $A + 1.645\sigma_A$ where $\sigma_A$ includes statistical uncertainties only, while the shaded area corresponds to the contribution from systematics. The dashed-dotted line shows the sensitivity curve. The lines at A=0 and A=1 are also given.*

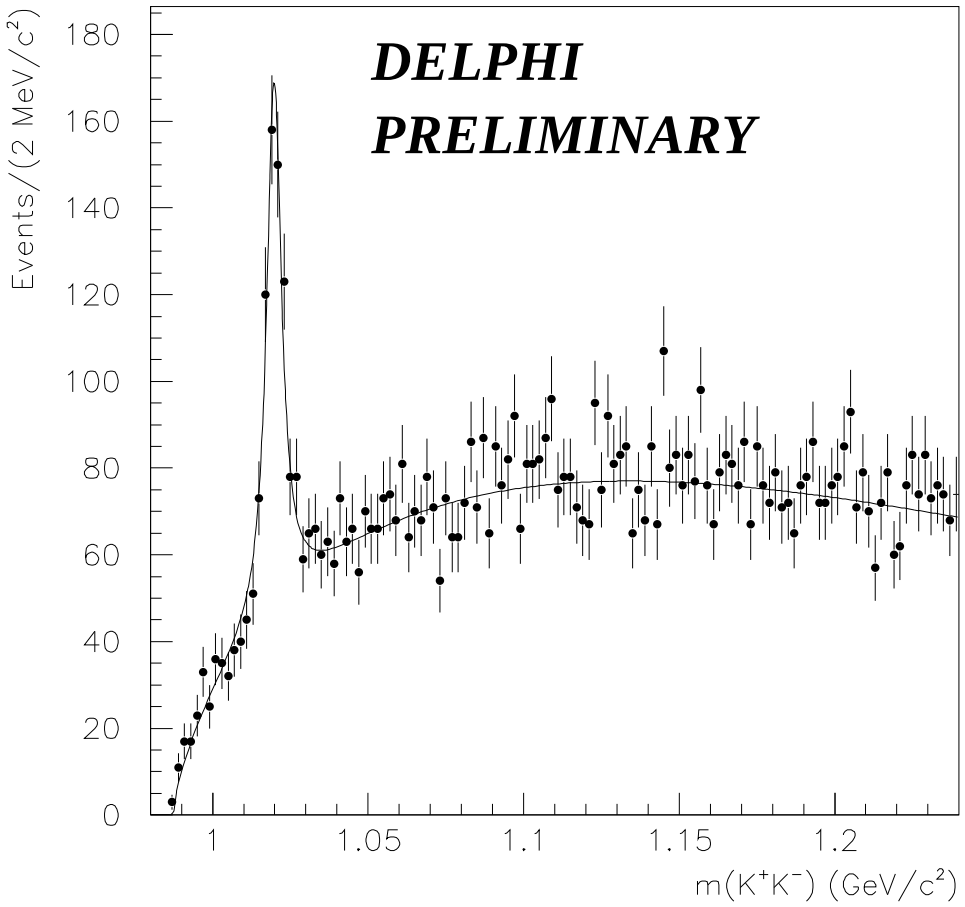


Figure 10: *($\phi\ell^{\mp}$) analysis on 94-95 data:*
*The plots show the invariant $K^+K^-$ mass distributions for $\phi$ candidates accompanied by a lepton of opposite electric charge, identified in the same hemisphere, and with $p_t^{out}$ above 1.0* GeV/c. *The curve shows the fit described in the text.*

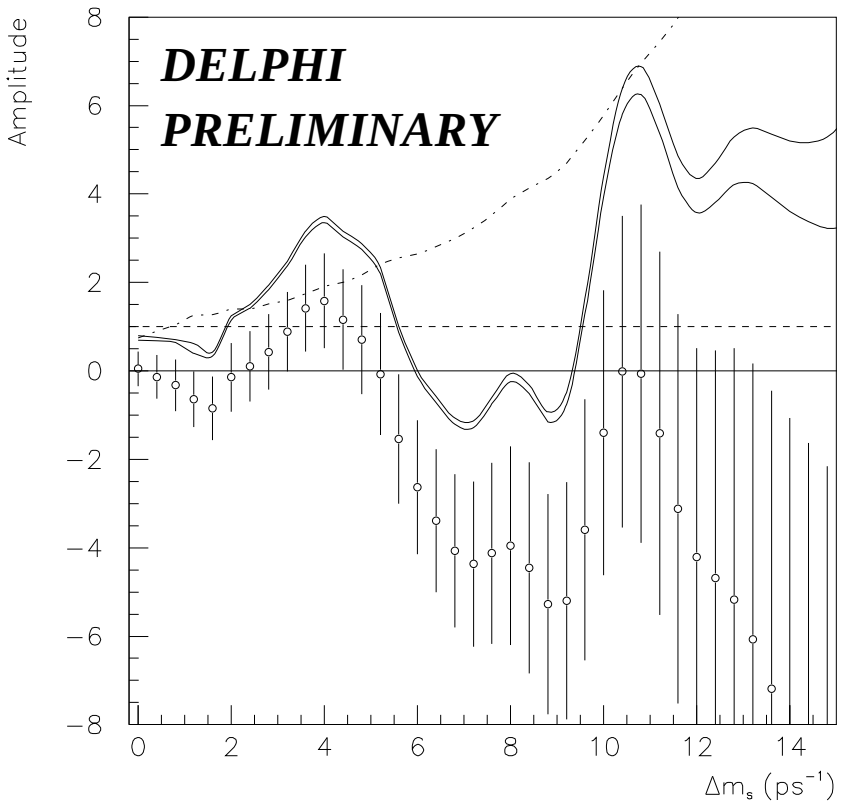


Figure 11: *($\phi\ell^{\mp}$) analysis on 94-95 data :*
*Variation of the oscillation amplitude A as a function of $\Delta m_s$. The lower continuous line corresponds to $A + 1.645\sigma_A$ where $\sigma_A$ includes statistical uncertainties only, while the shaded area shows the contribution from systematics. The dashed-dotted line corresponds to the sensitivity curve. The lines at A=0 and A=1 are also given*

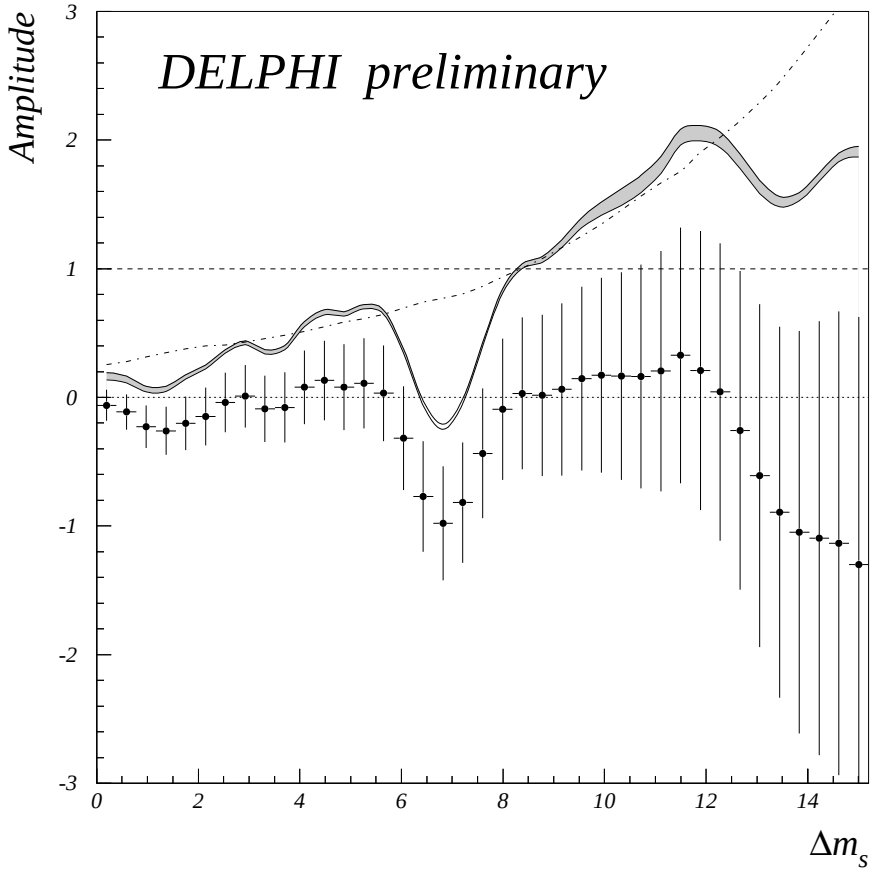


Figure 12: *Combination of the three analyses described in this paper and including previous inclusive analyses described in [1] :*
*Variation of the oscillation amplitude A as a function of* $\Delta m_s$. *The lower continous line corresponds to* $A + 1.645\sigma_A$ *where* $\sigma_A$ *statistical uncertainties only, while the shaded area corresponds to the contribution from systematics. The dashed-dotted line shows the sensitivity curve. The lines at A=0 and A=1 are also given.*

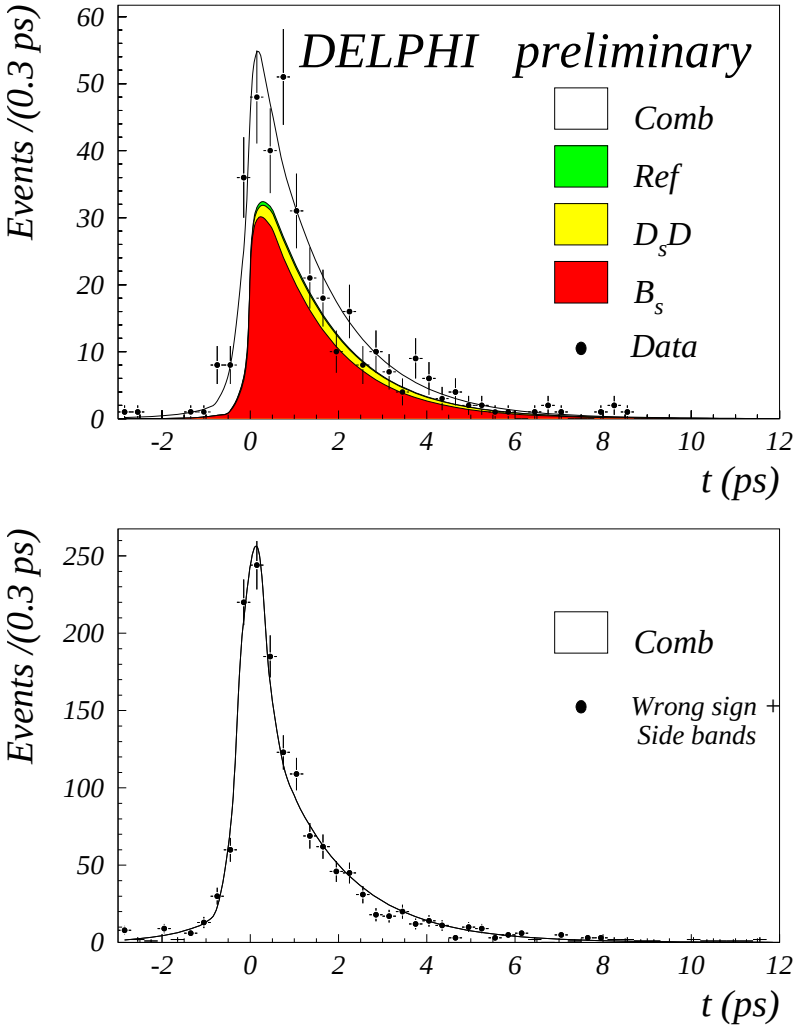


Figure 13: *($D_s^{\pm}\ell^{\mp}$) analysis on 91-95 data:*
*the upper plot shows the likelihood fit for events in the signal mass region. The points with error bars represent the data and the curves correspond to the different fitted contributions in the selected sample. The lower plot shows the corresponding distribution for the "wrong sign" events and events situated in side band regions with the fit superimposed.*

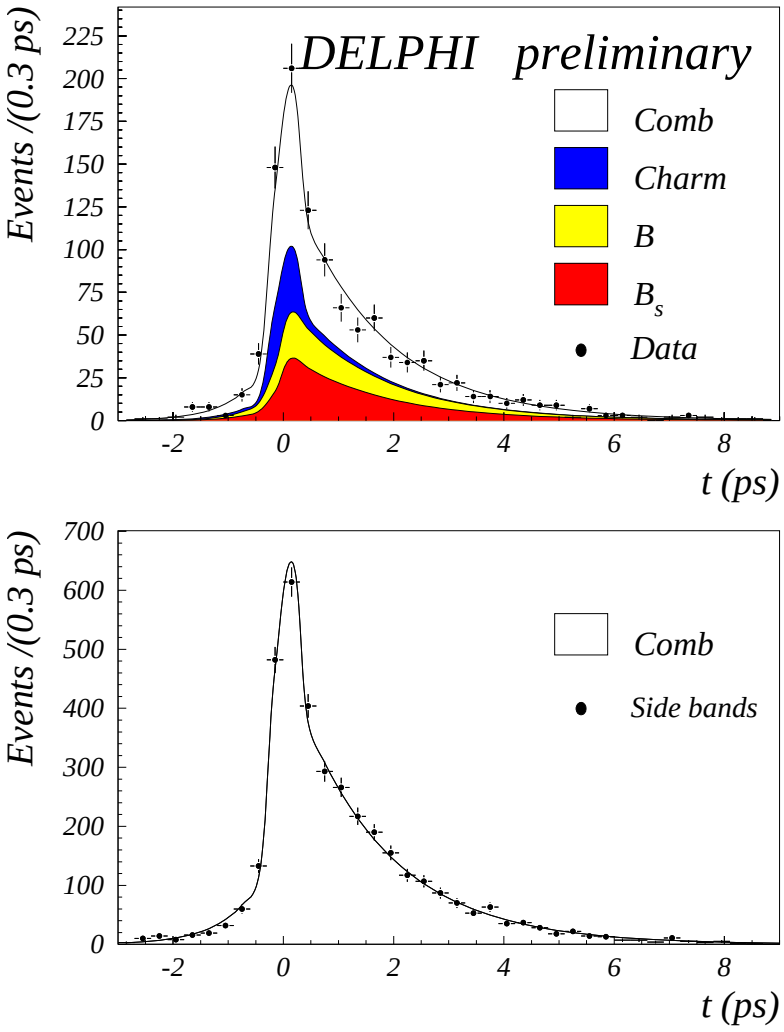


Figure 14: *($D_s^{\pm}h^{\mp}$) analysis on 94-95 data:*
*the upper plot shows the fitted proper time distribution for events in the signal mass region. The solid line shows the result of the maximum likelihood fit. The points with error bars represent the data. The fitted proper time distribution for events in the side–band (2.1-2.3* $GeV/c^2$*) mass region is shown in the lower plot.*

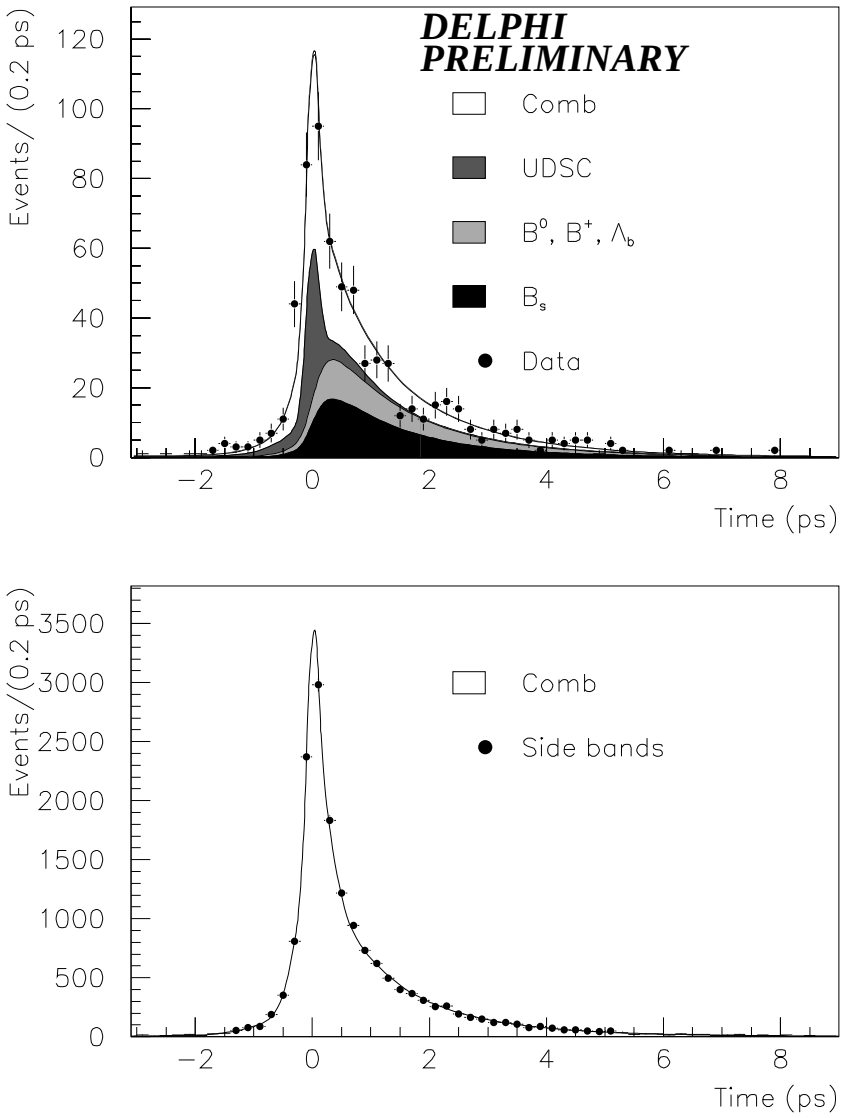


Figure 15: *($\phi\ell^{\mp}$) analysis on 94-95 data:*
*the upper plot shows the fitted proper time distribution for events in the signal mass region. The solid line shows the result of the maximum likelihood fit. The points with error bars represent the data.*

# Appendix B

# General remarks on cut optimization and parameter estimation

When trying to discover particles or resonances in the presence of an irreducible background, the kinematic cuts are usually chosen so they maximize some ratio of signal $S$ to background $B$, like one of

$$\begin{aligned} r_1 &= \frac{S}{B} \\ r_2 &= \frac{S}{\sqrt{B}} \\ r_3 &= \frac{S}{\sqrt{S+B}} \end{aligned}$$

All these three measures have the desirable property that their optimum is invariant when scaling $S$ and $B$ with a common factor. This means that the cuts will not have to be re-optimized when the luminosity of the experiment increase. Furthermore, it can be argued that $r_3 = \frac{S}{\sqrt{S+B}}$ is the ratio of the number of signal events and the expected fluctuation of the number of events, and it is often identified as the statistical significance of the sample. Roughly speaking, optimizing the cuts with respect to statistical significance is equivalent to selecting the set of cuts that maximizes the probability of making a discovery. A detailed analysis [76] of statistical significance involving the theory of hypothesis testing and likelihood methods is not included here.

Consider the measurement of some cross section $\sigma$ in the presence of a background. The expected number of signal events $s$ is given by the detector efficiency $\epsilon$ and the experiment luminosity $\mathcal{L}$, and the level of background $b$ is of course proportional to the luminosity:

$$\begin{aligned} s &= \epsilon\sigma\mathcal{L} \\ b &= \rho\mathcal{L} \end{aligned}$$

The probability of observing $n$ events is given by the Poisson distribution

$$\begin{aligned} p(n;\mu) &= \frac{e^{-\mu}\mu^n}{n!} \\ \mu &= s + b = (\epsilon\sigma + \rho)\cdot\mathcal{L}. \end{aligned}$$

The maximum likelihood estimator $\hat{\sigma}$ of the cross section is found by differentiating with respect to $\sigma$ and finding the root.

$$\hat{\sigma} = \frac{n - \rho\mathcal{L}}{\epsilon\mathcal{L}} \tag{B.1}$$

Using the well known mean $\langle n \rangle = \mu$ and variance $\langle (n-\mu)^2 \rangle = \mu$ of the Poisson distribution, $\hat{\sigma}$ is shown to be unbiased, and its variance can be found:

$$\langle \hat{\sigma} \rangle = \frac{\langle n \rangle - \rho\mathcal{L}}{\epsilon\mathcal{L}} = \frac{\mu - \rho\mathcal{L}}{\epsilon\mathcal{L}} = \sigma \tag{B.2}$$

$$\langle (\hat{\sigma} - \sigma)^2 \rangle = \left\langle \left( \frac{n-\mu}{\epsilon\mathcal{L}} \right)^2 \right\rangle = \frac{\mu}{(\epsilon\mathcal{L})^2} = \frac{\epsilon\sigma + \rho}{\epsilon^2\mathcal{L}} \tag{B.3}$$

Assuming that the signal efficiency $\epsilon = \epsilon\alpha$ and the background level $\rho = \rho\alpha$ are both functions of some cut parameter $\alpha$, this parameter should be chosen such that $\sigma_\sigma^2$ given by equation B.3 is minimal. This is equivalent to maximizing $\frac{s}{\sqrt{s+b}}$

Contrary to simple event counting experiments, it is not immediately clear how to optimize kinematic cuts in the general case when some measured distribution is used for *parameter estimation* involving the maximum likelihood method. The reason is that it is not clear which part of the background deteriorates the statistical significance of the signal the most. To illustrate this, consider the following series of Monte Carlo experiments: Signal and background decays are generated in known fractions. In each experiment, the signal lifetime is fitted using the maximum likelihood estimate, assuming the background lifetime to be known. Repeating the experiment, the uncertainty of the lifetime fit can be evaluated as the variance of the fitted values. For a background-free experiment with $n$ signal events, the measurement variance is known from elementary statistics to be

$$\sigma_\tau = \frac{\tau}{\sqrt{n}}. \tag{B.4}$$

With reference to equation B.4, the statistical significance of a sample with background can be conveniently described in terms of the *effective* number of events $n_{\text{eff}}$:

$$n_{\text{eff}} = \left( \frac{\tau}{\sigma_\tau} \right)^2 \tag{B.5}$$

In figure B.1 it is clearly seen that a long-lived background is much more harmful than the short-lived one.

Having shown that different kinds of background have very different impact on the uncertainty of a parameter estimation, it is still not evident that this will affect the procedure for optimizing the cuts. To elaborate, consider an experiment trying to measure the lifetime $\tau_{\text{S}}$ of a signal $S$ in the presence of two classes of background $B_1$ and $B_2$ of lifetimes $\tau_{\text{B1}} << \tau_{\text{B2}} \simeq \tau_{\text{S}}$. Furthermore, consider two cut variables $c_1$ and $c_2$ that mainly affects $B_1$ and $B_2$, respectively. It is not hard to imagine a distribution of signal and background that would allow a simultaneous change of the cuts in $c_1$ and $c_2$ leaving $S$ and $B_1 + B_2$ unchanged. Naively trying to optimize $\frac{S}{\sqrt{S+B}} = \frac{S}{\sqrt{S+B_1+B2}}$, one would not realize that cutting hard in the variable that affects the long-lived background while leaving the short-lived background untouched, would give the most accurate lifetime measurement.

The upshot of this reasoning is that cut optimization needs more consideration than just plotting $\frac{S}{\sqrt{S+B}}$ versus the cut variable and look for a maximum. If the background $B$ is composite, the different classes will affect the measurement differently. On the other hand, making a precise definition of statistical significance and subsequently optimizing with respect to this measure might be too intricate. Because of finite statistics of the Monte Carlo simulations, the distributions of the variables in question will only be approximately known. Hence, a mathematically correct but complicated procedure cannot be justified. A reasonable compromise between simplicity and accuracy might be to consider only the part $B'$ of the background that is affected by a given cut and optimize with respect to $\frac{S}{\sqrt{S+B'}}$.

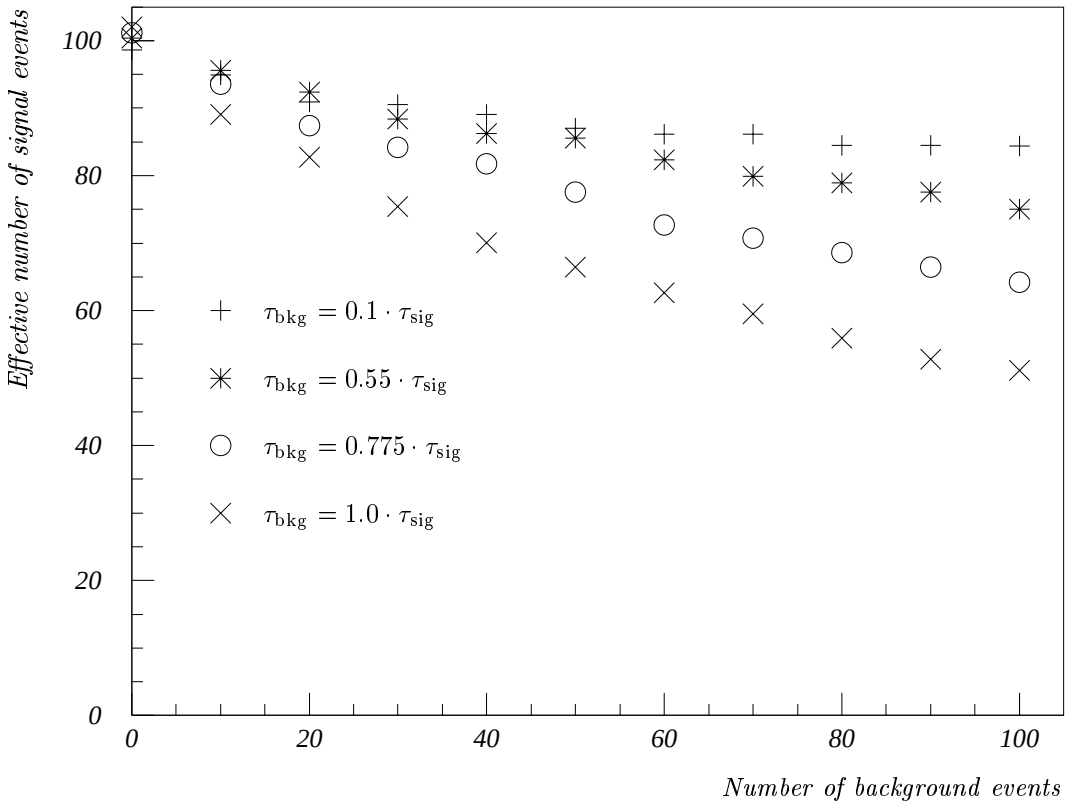


Figure B.1: *Lifetime fit in the presence of a known background. While the generated number of signal events is fixed, the effective number of signal events is computed using the variance of the fitted lifetime and defining* $\sigma_\tau = \frac{\tau}{\sqrt{n_{\rm eff}}}$.

# Appendix C

# Unbiased flight length estimation

After doing the primary and secondary vertex reconstruction, the full three dimensional flight path is available. It is expected that just taking the length of the measured path produces a bias towards longer path lengths: Take $\mathbf{x} = (x, y, z)$ as the true flight path and $\delta\mathbf{x} = (\delta x, \delta y, \delta z)$ as the measurement error. Assume that the error $\delta\mathbf{x} = (\delta x, \delta y, \delta z)$ is Gaussian with zero mean and a covariance matrix of $\langle \delta\mathbf{x}\delta\mathbf{x} \rangle = \mathbf{W}^{-1}$. Now, the mean measured flight length will be:

$$\langle l \rangle = \int \mathrm{d}^3\delta x |\mathbf{x} + \delta\mathbf{x}| \frac{|\mathbf{W}|^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}} e^{-\frac{1}{2}\delta\mathbf{x}^T\mathbf{W}\delta\mathbf{x}} = |\mathbf{x}| + \langle |\delta\mathbf{x}| \rangle + 2\langle \mathbf{x} \cdot \delta\mathbf{x} \rangle \tag{C.1}$$

The second term is positive while the third term is zero:

$$\langle l \rangle = |\mathbf{x}| + \langle |\delta\mathbf{x}| \rangle > |\mathbf{x}|. \tag{C.2}$$

Alternatively, consider the geometrical argument of figure C.1. Measurements falling outside the dashed arc give a result longer than the true length. One is easily convinced that measuring a longer length is more probable than measuring a shorter length.

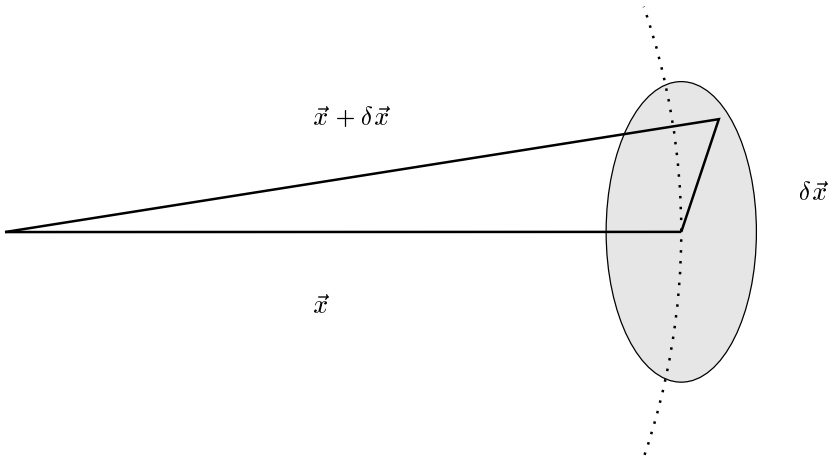


Figure C.1: *Geometrical bias of flight length estimation. The larger part of the error ellipse (shaded) falls outside the circular arc (dashed), indicating that the absolute value is a biased estimator of the flight length*

A more important reason for not using the length of the three dimensional flight path directly is that the measurement accuracy can be quite different in the $x$, $y$ and $z$ directions. With

$$l = \sqrt{x^2 + y^2 + z^2}, \tag{C.3}$$

simple error propagation gives

$$\sigma_l^2 = \left(\frac{x}{l}\sigma_x\right)^2 + \left(\frac{y}{l}\sigma_y\right)^2 + \left(\frac{z}{l}\sigma_z\right)^2 \tag{C.4}$$

Clearly, the error will be dominated by the coordinate that is measured with the poorest accuracy. To improve this, the flight path is projected onto the direction of the reconstructed momentum.

# Appendix D

# On the determination and interpretation of $1\sigma$ errors

The *classical statistics* or *frequentist* viewpoint and the *Bayesian* viewpoint offer two alternative descriptions of uncertainty in a measurement process. Frequently, concepts such as $1\sigma$ error and 95% confidence level are used without specifying which viewpoint they refer to. In two prominent examples, Gaussian and Poissonian statistics, the frequentist and the Bayesian viewpoint happen to give identical numbers. In the general case, this is not true.

### Frequentist $1\sigma$ errors

The concept of *sampling* is at the heart of the frequentist viewpoint. Any measurement process can be repeated, at least in principle. Each time a given experiment is performed, the outcome $\vec{x}$ of the experiment is *sampled* from a virtual population of possible outcomes. Repeating the experiment, the outcomes will be distributed according to some *probability density function* (pdf): $f(\vec{x};\vec{p})$ with parameters $\vec{p}$. A *statistic* is a function $s_i(\vec{x})$ of a single experimental outcome. An *estimator* $\hat{p}_i(x) = s_i(\vec{x})$ is a statistic that is constructed to estimate a parameter of the pdf. Being a function of the experimental outcome, the estimator itself is distributed according to some pdf. Any useful estimator should have a correct mean: $\langle \hat{p}_i \rangle = p_i$. The frequentist $1\sigma$ errors on an estimator $\hat{p}$ should be defined such that the interval $\langle \hat{p} - \sigma^- \ldots \hat{p} + \sigma^+ \rangle$ contains the true value in $\mathrm{erf}(\frac{1}{\sqrt{2}}) = 68\%$ of future experiments. In practice, the maximum likelihood method is often used to derive estimators. Furthermore, $1\sigma$ errors are calculated by moving the estimated parameter until there is a given required relative change in the likelihood function from the maximum likelihood. For the Gaussian case, this required likelihood ratio is:

$$\frac{\mathcal{L}(\hat{p} \pm \sigma)}{\mathcal{L}(\hat{p})} = e^{-\frac{1}{2}} \tag{D.1}$$

Using the more familiar $\chi^2(p) = -2 \log \mathcal{L}(p)$, this is expressed as

$$\Delta\chi^2 = \chi^2(\hat{p} \pm \sigma) - \chi^2(\hat{p}) = 1$$

.

### Bayesian $1\sigma$ errors

In Bayesian statistics, the parameters of the pdf are treated as statistical variables and the parameterized pdf $f(\vec{x};\vec{p})$ is promoted to a conditional pdf. $f(\vec{x}|\vec{p})$. This motivates the use of Bayes' theorem:

$$p(\vec{x},\vec{p}) = f(\vec{x}|\vec{p}) \cdot q(\vec{p}) = g(\vec{p}|\vec{x}) \cdot r(\vec{x}) \tag{D.2}$$

Doing the experiment means measuring $\vec{x}$. Having determined $\vec{x}$, equation D.2 can be solved to determine the conditional pdf for $\vec{p}$:

$$g(\vec{p}|\vec{x}) = \frac{f(\vec{x}|\vec{p})q(\vec{p})}{r(\vec{x})} \tag{D.3}$$

In this expression $r(\vec{x})$ is just a number that can be found using the normalization condition on $g(\vec{p}|\vec{x})$:

$$g(\vec{p}|\vec{x}) = \frac{f(\vec{x}|\vec{p})q(\vec{p})}{\int f(\vec{x}|\vec{p})q(\vec{p})\mathrm{d}p} \tag{D.4}$$

In this context, $q(\vec{p})$ represents any previous knowledge we might have about $\vec{p}$, being previous measurements or theoretical predictions. The Bayesian definition of $1\sigma$ errors is the shortest interval $\langle \hat{p} - \sigma^- \ldots \hat{p} + \sigma^+ \rangle$ such that the following equation is satisfied:

$$\int_{\hat{p}-\sigma^-}^{\hat{p}+\sigma^+} g(p|\vec{x})\mathrm{d}p = \mathrm{erf}(\frac{1}{\sqrt{2}}) = 68\% \tag{D.5}$$

Note that for a few but prominent cases, the frequentist requirement on the $1\sigma$ error can be written as an integral of a form identical to equation D.5 and the numerical values will be identical.

### Practical error determination

Having found the set of parameters that maximizes the log likelihood function, a crude estimate for the fit error on each parameter is evaluated using the second derivative of the log likelihood function at the maximum:

$$\sigma_p^2 = -\left( \left. \frac{\partial^2 \mathcal{L}}{\partial p^2} \right|_{\max} \right)^{-1} \tag{D.6}$$

This expression is exact in the Gaussian case and generally, by the *central limit theorem*, it converges to the correct value as the sample size increases. However, for finite statistics lifetime measurements, the log likelihood function is not even symmetric and equation D.6 is clearly not applicable. Conventionally, the statistical error on a fitted lifetime is quoted using what [77] calls the "Improved Error". To evaluate the "Improved Error", the maximum of the log likelihood function is studied as a function of $\tau_{\mathrm{B}_s^0}$. The asymmetric errors are defined by the equation

$$\log \mathcal{L}_{\max}(\hat{\tau}_{\mathrm{B}_s^0} - \sigma_\tau^-) = \log \mathcal{L}_{\max}(\hat{\tau}_{\mathrm{B}_s^0} + \sigma_\tau^+) = \log \mathcal{L}_{\max}(\hat{\tau}_{\mathrm{B}_s^0}) - \frac{1}{2} \tag{D.7}$$

or simply $\Delta \log \mathcal{L} = -\frac{1}{2}$. $\mathcal{L}_{\max}(\hat{\tau}_{\mathrm{B}_s^0})$ means that for each value of $\tau_{\mathrm{B}_s^0}$, likelihood function is maximized with respect to the remaining parameters. This way of defining the error interval is in general inconsistent with the Frequentist as well as with the Bayesian approach. The "Improved Error" is used because it is more economical to calculate than either of the two dogmatic methods. To get the Bayesian $1\sigma$ errors, one would have to integrate numerically $\mathcal{L}_{\max}(\hat{\tau}_{\mathrm{B}_s^0})$ whereas the Frequentist approach would require Monte Carlo simulation of a large number of experiments, evaluating the maximum likelihood estimate for the lifetime for each experiment.

# Bibliography


[1] S.L. Glashow. Parital–symmetries of weak interactions. *Nuclear Physics*, 22:579, 1961.

[2] S. Weinberg. A model of leptons. *Physical Review Letters*, 19:1264, 1967.

[3] A. Salam. Weak and electromagnetic interactions. In N. Svartholm, editor, *Elementary Particle Theory*. Almqvist and Wiksell, 1968.

[4] R. M. Barnett et al (Particle Data Group. Review of particle physics. *Physical Review*, D54:1–720, 1996.

[5] C. Casa et al (Particle Data Group). Review of particle physics. *The European Physical Journal*, C3:1, 1998. Available at `http://pdg.lbl.gov`.

[6] N. Cabibbo. Unitary symmetry and leptonic decays. *Physical Review Letters*, 10:531, 1963.

[7] M. Kobayashi and K. Maskawa. CP–violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49:652, 1973.

[8] S. Weinberg. –. *Transactions of the New York Academy of Sciences, Series II*, 38:185, 1977.

[9] H. Fritzsch. Calculating the Cabibbo angle. *Physics Letters*, 70B:436, 1977.

[10] L. Maiani. CP violation in purely lefthanded weak interactions. *Physics Letters*, 62B:183, 1976.

[11] L.-L. Chau and W.-Y. Keung. Comments on the parametrization of the Kobayashi–Maskawa matrix. *Physical Review Letters*, 53:1802, 1984.

[12] L. Wolfenstein. Parametrization of the Kobayashi–Maskawa matrix. *Physical Review Letters*, 51:1945, 1983.

[13] S. Erham. Beauty97 conference summary. *hep-ex/9805007*, Submitted to Nuclear Instruments and Methods in Physics research, 1998.

[14] T. Lohse et al. HERA-B: an experiment to study CP violation in the B system using an internal target at the HERA proton ring. *DESY Report*, PRC 94-02, 1994.

[15] Belle Collaboration. Letter of intent for a study of CP violation in B meson decays. *KEK Report*, 94-2, 1994.

[16] BaBar Collaboration. Letter of intent for the study of CP violation and heavy flavor physics at PEP II. *SLAC Report*, 443, 1994.

[17] A. Ali. *B* decays — introduction and overview. In S. Stone, editor, *B Decays*. World Scientific, 2nd edition, 1994. The measurement of $V_{ub}$ is discussed in section 3.8.

[18] M. Gell-Mann and A. Pais. Behavior of neutral particles under charge conjugation. *Physical Review*, 97:1387, 1955.

[19] T. D. Lee, R. Oehme, and C. N. Yang. Remarks on possible noninvariance under time reversal and charge conjugation. *Physical Review*, 106:340, 1957.

[20] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. –. *Physical Review Letters*, 13:138, 1964. Christenson and coworkers discovered CP–violation, observing the $K_L \to 2\pi$ decay. Later, CP–violation has also been observed in semileptonic decays of $K_L$.

[21] H. Schröder. $B\overline{B}$ mixing. In S. Stone, editor, *B Decays*. World Scientific, 2nd edition, 1994.

[22] Albrecht et al (ARGUS Collaboration). A study of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu u}$ and $B^0\overline{B}^0$ mixing using partial $D^{*+}$ reconstruction. *Physics Letters*, B324:249, 1994.

[23] Bartelt et al (CLEO collaboration). Evidence for penguin–diagram decays: First observation of $B \to K^*(892)\gamma$. *Physical Review Letters*, 71:674, 1993.

[24] Acciarri et al (L3 collaboration). —. *CERN Preprint*, EP/98-28, 1998. To be published in EPJ C.

[25] Abreu et al (DELPHI collaboration). Measurement of $B^0_d - \overline{B}^0_d$ oscillations. *Zeitschrift für Physik*, C76:579, 1997.

[26] Ackerstaff et al (OPAL collaboration). A study of B meson oscillations using hadronic $Z^0$ decays containing leptons. *Zeitschrift für Physik*, C76:401, 1997.

[27] Ackerstaff et al (OPAL) collaboration. An updated study of B meson oscillations using dilepton events. *Zeitschrift für Physik*, C76:417, 1997.

[28] D. Buskulic et al (ALEPH collaboration). Improved measurement of the $B^0_d - \overline{B}^0_d$ oscillation frequency. *Zeitschrift für Physik*, C75:397, 1997.

[29] Abe et al (CDF collaboration). Measurement of the $B^0 - \overline{B}^0$ oscillation frequency using $\pi -$ B meson charge–flavor correlations in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Physical Review Letters*, 80:2057, 1998.

[30] P. Paganini, F. Parodi, P. Roudeau, and A. Stocchi. Measurements of the $\rho$ and $\eta$ parameters of the $V_{CKM}$ matrix and perspectives. *hep-ph/9711261*, 1997.

[31] A. Abada et al. Meson spectroscopy and decay constants with Wilson fermions at $\beta = 6.4$. *Nuclear Physics*, B376:172, 1992.

[32] M. Beneke, G. Buchalla, and I. Dunietz. The width difference in the $B_s$–$\overline{B}_s$ system. *Physical Review*, D54:4419–4431, 1996.

[33] M. Neubert and C.T. Sachrajda. Spectator effects in inclusive decays of beauty hadrons. *Nuclear Physics*, B483:339–370, 1997.

[34] I. Bigi, B. Blok, M. Shifman, N. Uraltsev, and A. Vainshtein. Non-leptonic decays of beauty hadrons — from phenomenology to theory. In S. Stone, editor, *B Decays*. World Scientific, 2nd edition, 1994.

[35] P. Aarnio et al (DELPHI Collaboration). The DELPHI detector at LEP. *Nuclear Instruments and Methods in Physics Research*, A303:233, 1991.

[36] P. Abreu et al (DELPHI Collaboration). Performance of the DELPHI detector. *Nuclear Instruments and Methods in Physics Research*, A378:57–100, 1996.

[37] DELPHI Collaboration. DELPHI performance transparencies, 1996. Available at `<http://delphiwww.cern.ch/delfigs/export/pubdet.html>`.

[38] N. Bingefors et al. The DELPHI microvertex detector. *Nuclear Instruments and Methods in Physics Research*, A328:447, 1993.

[39] V. Chabaud et al. The DELPHI silicon strip microvertex detector with double sided readout. *Nuclear Instruments and Methods in Physics Research*, A368:314, 1996.

[40] V. Chabaud, A. Andreazza, P. Collins, and H. Dijkstra. Alignment of the DELPHI vertex detector. *DELPHI internal note*, 95-177 MVX 10, 1995.

[41] F. Hartjes et al. A drift chamber with variable drift velocity. *Nuclear Instruments and Methods in Physics Research*, A256:55, 1987.

[42] C. Brand et al. The DELPHI time projection chamber. *Nuclear Instruments and Methods in Physics Research*, A283:567–572, 1989.

[43] A. Amery et al. The DELPHI outer detector. *Nuclear Instruments and Methods in Physics Research*, A283:502–508, 1989.

[44] DELPHI. *DELANA Users's Guide*, 89-44 prog 137 edition, 1986.

[45] P. Billoir, R. Fruhwirth, and M. Regler. Track element merging strategy and vertex fitting in complex modular detectors. *Nuclear Instruments and Methods in Physics Research*, A241:115–131, 1985.

[46] W. Adam et al. Current achievements of the DELPHI ring imaging Cherenkov detector. *Nuclear Instruments and Methods in Physics Research*, A371:12, 1996.

[47] W. Adam et al. Particle identification algorithms for the DELPHI RICH detector. *Nuclear Instruments and Methods in Physics Research*, A371:240, 1996.

[48] H. G. Fischer et al. The DELPHI high density projection chamber. *Nuclear Instruments and Methods in Physics Research*, A265:218–222, 1988.

[49] M. Feindt, C. Kreuter, and O. Podobrin. ELEPHANT reference manual. *DELPHI internal note*, 96-82 PROG 217, 1996.

[50] C. Kreuter. Electron identification using a neural network. *DELPHI internal note*, 96-169 PHYS 658, 1996.

[51] F. Stichelbaut and G.R. Wilkinson. Performance of muon identification in DELPHI for the 93 and 94 data. *DELPHI internal note*, 95-140 PHYS 565, 1995.

[52] V. Bocci et al. (...). *Nuclear Instruments and Methods in Physics Research*, A362:361, 1995.

[53] V. Bocci et al. Basic concepts and architectural details of the DELPHI trigger system. *CERN preprint*, CERN/ECP 94-18, 1994. Contribution to the Nuclear Science Symposium and Medical Imaging Conference, Norfolk, USA, 1994, submitted to IEEE Trans. Nucl. Sci.

[54] R. Brun and J. Zoll. *ZEBRA User Guide.* CERN, Geneva, Switzerland, 1994.

[55] Yu. Belokopytov et al. The DELPHI database description. *DELPHI internal note*, 93-4 PROG 194, 1993.

[56] Yu. Belokopytov and V. Perevozchikov. CARGO database management package. *DELPHI internal note*, 93-5 PROG 195, 1993.

[57] Y. Sacquin. Description of the DELPHI DST content. *DELPHI internal note*, 94-161 PROG 210, 1994.

[58] Tz. Spassoff. DELPHI ShortDST content. *DELPHI internal note*, 97-146 PROG 221, 1997.

[59] T. Sjöstrand. High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4. *Computer Physics Communications*, 82:74–90, 1994.

[60] DELPHI. *DELSIM, DELPHI Event Generation and Detector Simulation User's guide*, 89-67 prog 142 edition, 1989.

[61] D. Buskulic (ALEPH collaboration). Study of the $B_s^0 - \overline{B}_s^0$ oscillation frequency using $D_s^- \ell^+$ combinations in $z$ decays. *Physics Letters*, B337:205–221, 1996.

[62] W. Adam et al (DELPHI collaboration). Search for $B_s^0 - \overline{B}_s^0$ oscillations. *Physics Letters*, B414:382–400, 1997.

[63] R. Barate et al (ALEPH collaboration). Study of $B_s^0$ oscillations and lifetime using fully reconstructed $D_s^-$ decays. *CERN Preprint*, PPE/97-157, 1997. Submitted to Zeitschrift für Physik C.

[64] P. Billoir et al (DELPHI $B_s^0$ subteam). Search for $B_s^0 - \overline{B}_s^0$ and measurement of the $B_s^0$ lifetime. *DELPHI internal note*, 97-75 CONF 61, 1997. Paper submitted to the HEP'97 Conference Jerusalem, August 1997.

[65] P. Billoir et al. The use of $B_s^0$ exclusively reconstructed events for $B_s^0 - \overline{B}_s^0$ oscillations measurement. *DELPHI internal note*, 98-31 CONF 126, 1998.

[66] O. Podobrin and M. Feindt. Inclusive measurement of the b fragmentation function. *DELPHI internal note*, 95-103 PHYS 538, 1995. Paper submitted to the "EPS-HEP 95" Conference, Brussels, 27th July – 2nd August 1995.

[67] P. Abreu et al. Measurement of the charged particle multiplicity of weakly decaying B hadrons. *Physics Letters*, B425:399–412, 1998.

[68] W. J. Murray. Measurement of the beam position in DELPHI. *DELPHI internal note*, 96-6 PHYS 590, 1996.

[69] R. A. Fisher. *Statistical Methods for Research Workers.* Oliver and Boyd, 10th edition, 1946.

[70] W. Ledermann (chief editor) E. Lloyd (volume editor), editor. *Handbook of Applicable Mathematics*, volume VI-B. Wiley, 1984.

[71] A. G. Frodesen, O. Skjeggestad, and H. Tøfte. *Probability and Statistics in Particle Physics.* Universitetsforlaget, Oslo, 1979.

[72] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Function.* Dover Publications, Inc., 9th edition, 1970.

[73] H.-G. Moser and A. Roussarie. Mathematical methods for $B^0\overline{B}^0$ oscillation analyses. *Nuclear Instruments and Methods in Physics Research*, A384:491–505, 1997.

[74] The LEP B Oscillations Working Group. Combined results on $B^0$ oscillations: Update for winter conferences 1998. *http://www.cern.ch/LEPBOSC/combined_results/winter_1998/winter_1998.ps*, LEPBOSC 98-1, 1998.

[75] L. Lyons and D. Gibaut. How to combine correlated estimates of a single physical quantity. *Nuclear Instruments and Methods in Physics Research*, A270:110–117, 1988.

[76] A. L. Read. Optimal statistical analysis of search results based on the likelihood ratio and its application to the search for the MSM Higgs boson at $\sqrt{s} = 161$ and 172 gev. *DELPHI internal note*, 97-158 PHYS 737, 1997.

[77] F. James. *Minuit: Function Minimization and Error Analysis.* CERN, Geneva, Switzerland, 94.1 edition. Available at `http://wwwinfo.cern.ch/asdoc/psdir/minuit.ps.gz`.