ALEPH 93 - 54
PHYSIC 93 - 45
C. Bauer et al.
12.3.93

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

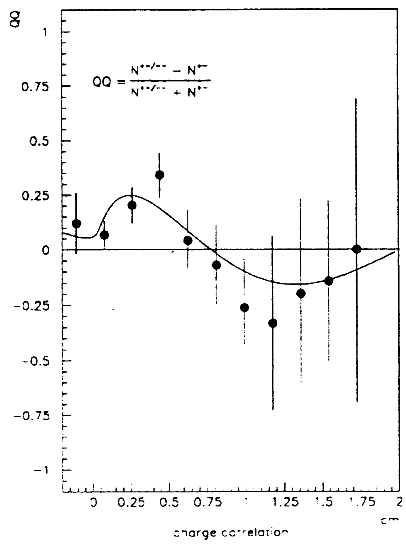
March 12, 1993

Observation of the time dependence of $B_d^0\bar{B}_d^0$ mixing

C. Bauer, H.-G. Moser, H. Seywerd

Abstract

The time dependence of $B_d^0\bar{B}_d^0$ oscillations has been observed using events with a D^* and a lepton in the opposite hemisphere. The time dependence of the oscillations can be derived from the vertex displacement of the D^0 from the D^* decay and the D^*-lepton charge correlation.



1 Introduction

Like in the K^0 system neutral B mesons (B_d^0 and B_s^0) can show the phenomena of mixing: time dependent oscillations from the particle state to the antiparticle state and vice versa. Hence a B^0 can transform into a $\bar{B}^0$ with the probability χ. The first evidence for $B^0 - \bar{B}^0$ mixing was found by UA1 [1]. However, this measurement could not distinguish between B_d^0 and B_s^0. They measured an average mixing parameter $\hat{\chi} = f_s\chi_s + f_d\chi_d$, where χ_i are the mixing parameters of the respective mesons and f_i their relative abundance. B_d^0 mixing was established by ARGUS [2] and confirmed by CLEO [3]. Other measurements of mixing followed later especially at LEP [4], with more precise results on $\hat{\chi}$. However, all these experiments could only measure the time integrated quantity χ. The characteristic time dependence has not been observed yet. This would be especially important for the B_s, which is expected to have a rather high oscillation frequency. In this case the time integrated measurement is, as will be shown later, insensitive to the actual oscillation frequency. Studies of the feasibility of time differential B_s analysis have been published [5][6][7]. Such measurements need very high statistics ($> 2\ 10^6$ *hadronic* Z^0) and a very good resolution of the B_s decay time. On the other hand the requirements to resolve B_d mixing, which should oscillate more slowly, are easier to meet. Using the ALEPH detector which is equipped with a high precision vertex detector a time differential analysis of $B_d^0\bar{B}_d^0$ oscillations using D^* lepton correlations has been made. Another analysis measuring the time dependence using dilepton events is described in [8].

This note is organized as follows: In section 2 a short overview on the phenomenology of $B\bar{B}$ mixing is given. The method to measure the time dependence of $B_d^0\bar{B}_d^0$ oscillations is discussed in section 3. In section 4 the decay length distributions of mixed and unmixed events including experimental effects are calculated. The events selection is explained in section 5. The treatment of the various background is shown in section 6. The decay length resolution is the topic of section 7. Section 8 treats the measurement of the oscillation frequency, the systematic errors of this measurement are discussed in section 9. Finally in section 10 the results are summarized.

2 Phenomenology of $B\bar{B}$ mixing

The flavour states B^0 and $\bar{B}^0$ are not eigenstates of the weak interaction, but a linear combination of the weak eigenstates B_1 and B_2 (assuming no CP violation):

$$B^0 = \frac{B_1 + B_2}{\sqrt{2}} \quad and \quad \bar{B}^0 = \frac{B_1 - B_2}{\sqrt{2}}.$$

These states have slightly different masses and decay widths m_1, Γ_1 and m_2, Γ_2 respectively and their time evolution in the rest frame is (using units with $\hbar = c = 1$):

$$a_i(t) \;=\; a_i(0)\ exp(-im_it - \Gamma_i\frac{t}{2}).$$

If at t=0 a pure B^0 state is produced then:

$$a_1(0) \; = \; a_2(0) \; = \frac{1}{\sqrt{2}}.$$

At $t \neq 0$ generally $a_1(t) \neq a_2(t)$ and the originally pure state will be a mixture of B^0 and $\bar{B}^0$. The probability that it decays at time t as a B^0 (*unmixed*) is:

$$p(t)^{unmixed} \; = \; \frac{\Gamma}{2} exp(-\Gamma t)[1 + cos(\Delta m t)], \tag{1}$$

and as a $\bar{B}^0$ (*mixed*):

$$p(t)^{mixed} \; = \; \frac{\Gamma}{2} exp(-\Gamma t)[1 - cos(\Delta m t)], \tag{2}$$

with $\Delta m \; = m_1 - m_2$ and $\Gamma \approx \Gamma_1 \approx \Gamma_2$ (assuming the differences in the decay width can be neglected). Integrating this from $t = 0$ to $t = \infty$ gives the total probability χ that a B^0 decays as a $\bar{B}^0$:

$$\chi \; = \frac{1}{2} \frac{(\frac{\Delta m}{\Gamma})^2}{1 + (\frac{\Delta m}{\Gamma})^2}$$

For large values of $\frac{\Delta m}{\Gamma}$ (> 3), χ saturates at its maximum value of 0.5. For B_d^0 χ has been measured by ARGUS and CLEO. The combined value of the two experiments is: $\chi_d \; = \; 0.159 \pm 0.025$ which corresponds to $\frac{\Delta m}{\Gamma} \; = \; 0.68 \pm 0.10$ [9].
A complete oscillation period takes about 9 lifetimes, the point when the B_d^0 has equal probability to decay as B_d^0 or $\bar{B}_d^0$ is at about 2.2 lifetimes. At LEP, where the B's have an average momentum of 31 GeV and a lifetime of about 1.5 ps this corresponds to a flight path of 5.9 mm.
The mass difference can be calculated from the box diagram [10][11]:

$$\Delta m \; = \; \left| \left(\frac{G_f^2}{6\pi^2} \right) B f_B^2 m_b m_W^2 F(m_t) \right|,$$

with the Fermi coupling constant G_f, the bag parameter B, the B-decay constant f_b, the b-mass m_b, the W-mass m_W. $F(m_t)$ is the dependence on the top mass (m_t):

$$F(m_t) \; \approx \; (V_{tb}^* V_{td})^2 \frac{m_t^2}{m_W^2} f(\frac{m_t^2}{m_W^2}),$$

with the Kobayashi-Maskawa Matrix elements V_{tb} and V_{td} and

$$f(x) \; = \; \frac{1}{4} + \frac{9}{4(1-x)} - \frac{3}{2} \frac{1}{(1-x)^2} - \frac{3}{2} \frac{x^2 ln x}{(1-x)}.$$

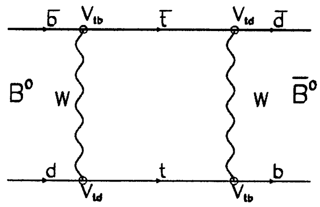


Figure 1: box diagram for $B_d^0\bar{B}_d^0$mixing

3 Measurements of the time dependent mixing at LEP

In order to measure the oscillation frequency Δm one has to tag the state of the B^0 at production time $t = 0$ and at the decay time $t = t_{decay}$. The standard way to tag the original state of the B^0 is via a semileptonic decay of the B in the opposite jet. The state at the decay time can be tagged by a D^*:

$$B_d^0 \rightarrow D^{*-} X \;\; ; \;\; \bar{B}_d^0 \rightarrow D^{*+} X.$$

Hence we can look at D^*-lepton charge correlations as a function of the B^0 decay time, which can be measured from the D^* decay vertex. A $D^{*-}l^-$ pair is called *like sign* and tags an *unmixed* event. A $D^{*-}l^+$ pair is called *unlike sign* and tags a *mixed* event. We can now define the charge correlation function QQ(t):

$$QQ(t) = \frac{N^{like} - N^{unlike}}{N^{like} + N^{unlike}} = \frac{N^{unmixed} - N^{mixed}}{N^{unmixed} + N^{mixed}}. \quad (3)$$

In case of a pure B^0 sample and perfect lepton tagging this quantity would oscillate between +1 and -1 with the frequency Δm. However, this is obscured by following processes:

- It is normally not possible to reconstruct the D^* decay vertex (which coincides with the B^0 vertex). The slow pion does not have enough p_t to allow precise vertexing. Hence the D^0 decay vertex must be used instead. This leads to a shift of the decay length distribution by approximately 0.9 mm and an additional smearing.

- The B on the lepton side may mix so causing the lepton to have the wrong sign correlation. This happens with the average mixing probability $\hat{\chi} = 0.128 \pm 0.10$

- The lepton can come from a $B \to D \to l$ cascade which gives also wrong sign correlations. This background can be controlled by cuts on the p and p_t of the lepton.
- The D^* can come from a $B^\pm$ decay. This gives unmixed events (if the lepton tag is correct). Unfortunately the branching ratios $B^0 \to D^*\ X$ and $B^- \to D^*\ X$ are not known. Nevertheless it can be assumed that most of the $D^{*\prime}s$ come from B^0. By the same token one can expect some background from decays like $B_s \to D^* K X$. This should be smaller than the $B^\pm$, and will be neglected.
- The D^* can come from $c\bar{c}$ events. This gives mixed sign correlations. However, because of the shorter lifetime of the D^0 compared to the $B^0 + D^0$ cascade this background dies out at long decay distances. Charm can also be reduced by a cut on the maximum D^* momentum and p_t cuts on the lepton.
- The lepton can be a decay lepton or any other mistag. These background leptons have no strong sign correlation with the D^* in the other hemisphere
- The D^* can be combinatorial background. Also in this case there should be no sign correlation with the lepton on the opposite side. The shape of this background can be determined from the sidebands of the D^* mass spectrum.
- Due to the limited vertex resolution the decay length spectrum will be smeared. As the resolution is in the order of 300 μm this effect is negligible compared to the length of an oscillation period (about 2.5 cm).
- The B^0 momentum is not measured. Hence the boost is not known and the decay time cannot be precisely reconstructed. This can be overcome by convoluting the decay time distribution with the B momentum spectrum to obtain the average decay length distribution. This results also in some smearing.

The background treatment is discussed in more detail in section 6.

4 Decay length distribution

Firstly the time distributions 1 and 2 are convoluted with the D^0 decay time. It is convenient to go already from time to decay length using:

$$f(\Gamma t)dt \to f(\frac{\Gamma}{\gamma\beta c} l)\ \frac{1}{\gamma\beta c} dl$$

where $\gamma\beta$ is the Lorentz boost of the decaying particle. Using $\Gamma_B = (\frac{\Gamma}{\gamma\beta c})_{B^0}$, $\Gamma_D = (\frac{\Gamma}{\gamma\beta c})_{D^0}$ and $M = (\frac{\Delta m}{\gamma\beta c})_{B^0}$ we get:

$$\frac{1}{n}\frac{dn}{dl}(l)^{unmixed}_{mixed} = \frac{\Gamma_B \Gamma_D}{2} \int_{l'=0}^{l'=l} exp(-\Gamma_B l')\ [1 \pm cos(Ml')] exp[-\Gamma_D(l-l')]\ dl' =$$

$$\frac{\Gamma_B \Gamma_D}{2}\left[\frac{1}{\Gamma_D-\Gamma_B}\left[exp(-\Gamma_B l)-exp(-\Gamma_D l)\right]\right.$$
$$\pm\frac{\Gamma_D-\Gamma_B}{(\Gamma_D-\Gamma_B)^2+M^2}\left[exp(-\Gamma_B l)\ cos(Ml)-exp(-\Gamma_D l)\right]$$
$$\pm\frac{M}{(\Gamma_D-\Gamma_B)^2+M^2} exp(-\Gamma_B l)\ sin(Ml), \quad (4)$$

or

$$=\frac{\Gamma_B \Gamma_D}{2}\left[\frac{1}{\Gamma_D-\Gamma_B}\left(exp(-\Gamma_B l)-exp(-\Gamma_D l)\right)\right.$$
$$\pm\sqrt{\frac{1}{(\Gamma_D-\Gamma_B)^2+M^2}}\ exp(-\Gamma_B l)\ cos(Ml+\delta)-\frac{\Gamma_D-\Gamma_B}{(\Gamma_D-\Gamma_B)^2+M^2}\ exp(-\Gamma_D l), \quad (5)$$

with $\delta = tan^{-1}(\frac{\Gamma_D-\Gamma_B}{M})-\frac{\pi}{2}$. This results in a rise of the signal with the D^0 decay constant, a damping of the oscillation amplitude by $\sqrt{\frac{(\Gamma_D-\Gamma_B)^2}{(\Gamma_D-\Gamma_B)^2+M^2}} \approx 0.94$ and a phase shift $\delta \approx -0.27$. In a next step the decay length function has to be convoluted with the vertex resolution. This can also be done analytically using the expressions:

$$\int_0^\infty \Gamma\ exp(-\Gamma l')\ exp\left(-\frac{(l-l')^2}{2\sigma^2}\right) dl' =$$
$$\Gamma exp\left(\frac{\Gamma^2\sigma^2}{2}-\Gamma\right)\left[1-ERF(\frac{\Gamma\sigma}{\sqrt{2}}-\frac{l}{\sqrt{2}\sigma})\right], \quad (6)$$

with the standard definition of the error function ERF. If there are oscillation terms the convolution becomes:

$$\int_0^\infty \Gamma\ exp(-\Gamma l')\ exp\left(-\frac{(l-l')^2}{2\sigma^2}\right) cos(Ml')\ dl' =$$
$$Re\left[\frac{\Gamma}{2}exp\left(\frac{\Gamma^2\sigma^2}{2}(1-iM\Gamma)-\Gamma(1-iM\Gamma)\right)\left(1-ERF(\frac{\Gamma\sigma}{\sqrt{2}}-\frac{l}{\sqrt{2}\sigma}-i\frac{M\sigma}{\sqrt{2}})\right)\right]. \quad (7)$$

The error function is now complex.
Finally the distribution has to be convoluted with the B^0 and D^0 momentum spectrum. This is done numerically. For the B^0 momentum spectrum we use a Peterson function:

$$p(z) = \frac{\alpha}{z\left(1-\frac{1}{z}-\frac{\epsilon}{1-z}\right)^2}, \quad (8)$$

with $z = \frac{p(B^0)}{p(beam)}$. ϵ and the range of z used for the convolution was determined by a fit to the momentum spectrum of MC events after selection cuts. For the D^0 momentum we used the average $p(D^0)$ as a function of $p(B^0)$, also obtained from MC events.
Using all this the evolution of the charge correlation function can be demonstrated :

- Figure 2 a) shows the optimal case, pure B^0 sample, no lepton mistag and perfect B^0 vertexing and momentum reconstruction.
- Figure 2 b) includes the effect of the D^0 decay length. The cosine is slightly shifted and its amplitude reduced.
- In figure 2 c) the momentum distribution of the B^0 and the vertex resolution is added. This leads also to a small shift and damping.
- The lepton mistag reduces the observed amplitude substantially as shown in figure 2 d). This is the largest effect.
- Adding the $c\bar{c}$ and $B^{\pm}$ background distorts the shape at short decay length, the distribution is also slightly shifted (figure 2 e)).
- Finally the addition of combinatorial background changes only the shape around $l = 0$ (figure 2 f)).

Despite all these effects the existence of a zero-crossing of QQ at $l \approx 1\ cm$ is not affected (however shifted). The alternative without time dependence has no zero-crossing and levels at $l > 0.5\ cm$ at a positive value.

5 Event selection

D^* events are identified in the decay $D^{*+} \longrightarrow D^0\pi^+$ where a selection on the mass difference $m(D^*) - m(D^0)$ allows a powerful rejection of combinatorial background. The D^0 candidates are reconstructed in three decay channels:

$$D^0 \rightarrow K^-\pi^+ \tag{9}$$
$$D^0 \rightarrow K^-\pi^+\pi^0 \tag{10}$$
$$D^0 \rightarrow K^-\pi^+\pi^-\pi^+ \ . \tag{11}$$

The events are selected from the 1991 and 1992 data sample using the hadronic event selection (CLAS 16) and VDET run selection.
Muons are selected using the identification flags 13 and 14. Electrons using the QEIDO flag and requiring 2σ dE/dx compatibility if dE/dx is available. Conversions are removed using the standard criteria. Jets are reconstructed with PCPA and the lepton p_t is calculated with the lepton removed from the jet.
For the $K\pi$ channel following selection cuts are used on the tracks:

- $\geq$ 1 $r - \phi$ VDET hit. No VDET requirement for the slow pion from the D^*
- $\geq$ 4 TPC hits
- χ^2/dof < 4 (track fit)

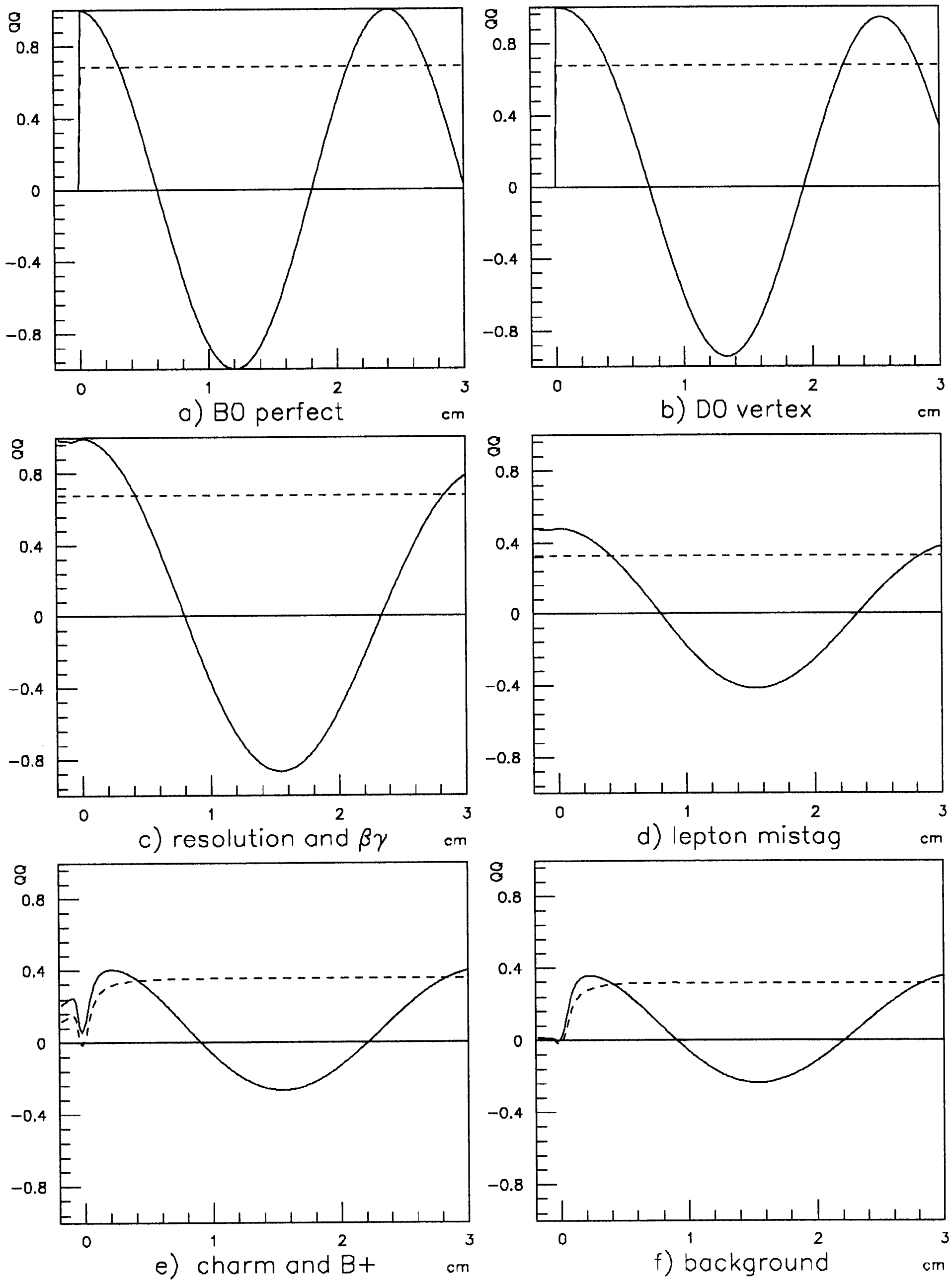


Figure 2: Charge correlation function: a) for a pure B^0 sample with perfect B^0 vertex reconstruction and perfect charge tagging, fixed B-momentum; b) using the D^0 vertex instead; c) convoluted with the vertex resolution and B^0 momentum distribution; d) including lepton mistag; e) adding charm and $B^{\pm}$ background; f) adding combinatorial background. The dashed line corresponds to a hypothetical time independent mixing. The values used correspond to the $K\pi$ sample.

- $d_0 < 2\ cm$ and $z_0 < 10\ cm$
- $\chi^2 < 20$ for the D^0 YTOP vertex

The physics cuts are

- standard V^0 rejection of any track used
- track momentum (K,π) $\geq$ 0.2 GeV/c
- 7 GeV/c $< p(D^*) <$ 25 GeV/c
- 1.845 $GeV/c^2 < m(D^0) <$ 1.885GeV/c^2
- 144 $MeV/c^2 < |m(D^*) - m(D^0)| <$ 147 MeV/c^2
- $|cos(\theta_{KD})| <$ 0.8 (decay angle of K towards flight direction of D^0 in its rest frame)

In the two other channels the combinatorial background is much higher. Hence more severe selection cuts have to be applied. For the $K\pi\pi\pi$ channel we use:

- p(K) > 2.5 GeV/c
- p(π) > 0.3 GeV/c
- K,π: 2 σ dE/dx requirement (if available)
- 4 tracks with $\geq$ 4 $r-\phi$ VDET hits, $\geq$ 2 z VDET hits in inner layer, $\geq$ 2 z VDET hits in outer layer,
- 1.849 $GeV/c^2 < m(D^0) <$ 1.879 GeV/c^2
- D^0 vertex probability > 0.01 (3 or 4 prong vertex)
- soft π: 0.5 $GeV/c < p <$ 4.2 GeV/c
- 0.144 $GeV/c^2 < m(D^*) - m(D^0) <$ 0.147 GeV/c^2
- 10.0 $GeV/c < p(D^*) <$ 25 GeV/c

Finally for $K\pi\pi^0$: The π^0 selection was:

- QPI0D0
- $\chi^2 < 25$
- $|m(\gamma\gamma) - m(\pi^0)| <$ 0.50 GeV/c^2

- $p(\pi^0) < 2.0$ GeV/c

The charged track, D^0, D^* selection criteria are:

- $p(\pi,K) > 1$ GeV/c
- $\geq 1\ r-\phi$ VDET hits
- K,π: 2 σ dE/dx requirement (if available)
- $\chi^2(YTOP) \ < \ 25.0$
- soft $\pi > 0.2$ GeV/c
- $|m(K\pi\pi^0) - 1.880\ GeV/c^2)| \ < \ 0.066\ GeV/c^2$ *
- $0.1434\ GeV/c^2 \ < \ m(D^*) - m(D^0) \ < \ 0.1486\ GeV/c^2$
- $11\ GeV/c^2 < p(D^*) < 25\ GeV/c$

The mass difference distributions for the various subsamples are shown in figure 3, the number of selected events in the various channels is listed in table 2.
The primary vertex is reconstructed using QFNDIP [12]. The decay length is calculated using:

$$l \ = \ \frac{t_i\sigma_{ij}^{-1}x_j}{t_i\sigma_{ij}^{-1}t_j},$$

t_i is the normalized momentum vector of the D^0, x_i the difference between D^0 and primary vertex, and σ_{ij} the total error matrix of primary and D^0 vertex.

6 Background

6.1 Lepton background

The lepton mistag is given by three sources:

- $B\bar{B}$ **mixing**: the fraction of leptons which tag the wrong charge due to mixing is given by the average mixing parameter $\hat{\chi} \ = \ 0.128 \ \pm \ 0.10$. This is an irreducible background.

*The D^0 mass is peak is not centered at the correct value, probably due to calibration of the electromagnetic calorimeter. Therefore the cut is made around the center of the observed distribution.

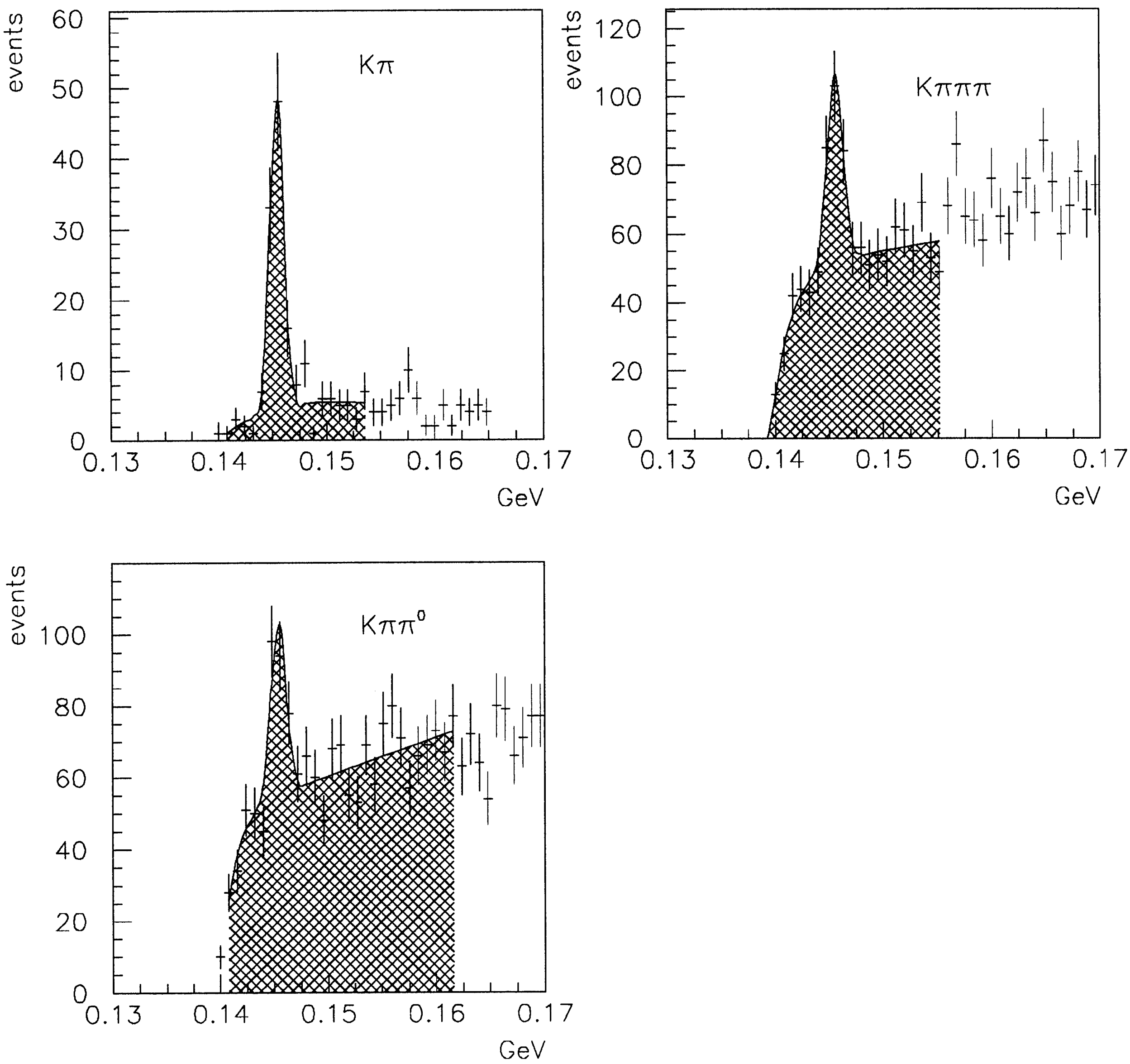


Figure 3: mass difference $m(D^*) - m(D^0)$ for the samples $K\pi$, $K\pi\pi\pi$ and $K\pi\pi^0$. The hatched area is used for the determination of the background

- **Second generation decays:** $B \rightarrow D \rightarrow l$ cascades give also wrong sign leptons. This background can be reduced by cutting on the p and p_t of the lepton. The values used in this analysis are taken from MC simulations. They are cross checked with the values obtained in the *high p_t lepton analysis* [13]. The fraction of second generation leptons

$$\eta = \frac{f(b \rightarrow c \rightarrow l)}{f(b \rightarrow l) + f(b \rightarrow c \rightarrow l)}$$

  can be taken from table 1.

- **Mis-identification:** This are decay muons, electrons from γ conversions, punch through and all other types of misidentified leptons. In first approximation there should be no charge correlation with the D^* in the other hemisphere, in fact there is a very small correlation due to some leading particle effect. The fraction

$$\delta = \frac{f(background)}{f(B) + f(background)}$$

  and the charge correlation is listed in table 1.

The total mistag probability per lepton in a b-event can then be expressed by:

$$A = (1 - \delta)\,(\hat{\chi} + \eta - 2\hat{\chi}\eta) + \delta d_{dec} \tag{12}$$

(q_{dec} is the fraction of misidentified leptons leading to 'mixed' events).

sample	ref [13]
beauty	
1. generation	0.805 ± 0.006
2. generation	0.139 ± 0.004
decays	0.060 ± 0.003
unlike fraction	0.46 ± 0.04
charm	
1. generation	0.61 ± 0.084
decays	0.252 ± 0.049
unlike fraction	0.615 ± 0.12

Table 1: Lepton purities. beauty 1. generation includes also $b \rightarrow \tau \rightarrow l$ and $b \rightarrow \bar{c} \rightarrow l$, which lead to correct charge assignments.

By the same token one can define a mistag probability 'C' for leptons in $c\bar{c}$ events. Here only misidentification can lead to mistagging.

6.2 D^* combinatorial background

The amount of combinatorial background is obtained from a fit to the $m(D^*)-m(D^0)$ mass spectrum. The signal is fitted with a gaussian, the background by a polynominal times an exponential damping term. The background is obtained by integrating the background fit between the limits of the D^* selection cuts. In case of the $K3\pi$ and $K\pi\pi^0$ samples special attention is paid for rejecting multiple combinations of the same event. The fitted background and the number of events are listed in table 2.
In order to parametrize the decay length distribution of the background events in the range

$$m(D^*) - m(D^0) \ < 143MeV/c^2 \ \ or \ \ 148MeV/c^2 < \ m(D^*) - m(D^0) \ < 165MeV/c^2$$

are selected and fitted with 2 gaussians (for $l < 0$) and 2 gaussians plus one exponential for $l > 0$. The exponential accounts for the combinatorial background from $b\bar{b}$ events which has substantial decay length. The events in the sidebands are also used to calculate the charge correlation for the combinatorial background. It is verified that the charge correlation does not depend on the decay length. From this the fraction of combinatorial background leading to 'mixed' events 'B_{bkg}' can be derived. The shape of the background distributions is shown in figure 4. The charge correlation of the background is plotted in figure 5.

sample	total events	background	l > 0 cm events	background
$K\pi$	106	18.0 ± 2.0	101	15.5± 1.9
$K\pi\pi^0$	367	222.0 ± 20.0	297	178.0± 18.0
$K\pi\pi\pi$	191	102.0 ± 10.0	169	91.0± 9.0

Table 2: background from a fit to real data sidebands of the D^* - D^0 mass difference. The signal range is from 144 to 147 MeV/c^2.

6.3 Charm background

The momentum spectrum of D^* from B and from $c\bar{c}$ is very different, $c\bar{c}$ events have on average higher momentum. Although D^0 from direct charm have on average a smaller decay length than the total flight path of D^0 from B, there is a small fraction with substantial decay length due to the large Lorentz boost of direct D^0. Hence a cut on the maximum momentum of the D^* removes a large fraction of the charm background and especially events at large decay length. This is demonstrated in table 3. We choose a cut of 25 GeV/c.
Events from $c\bar{c}$ are also very efficiently removed by a cut on the lepton p_t as discussed

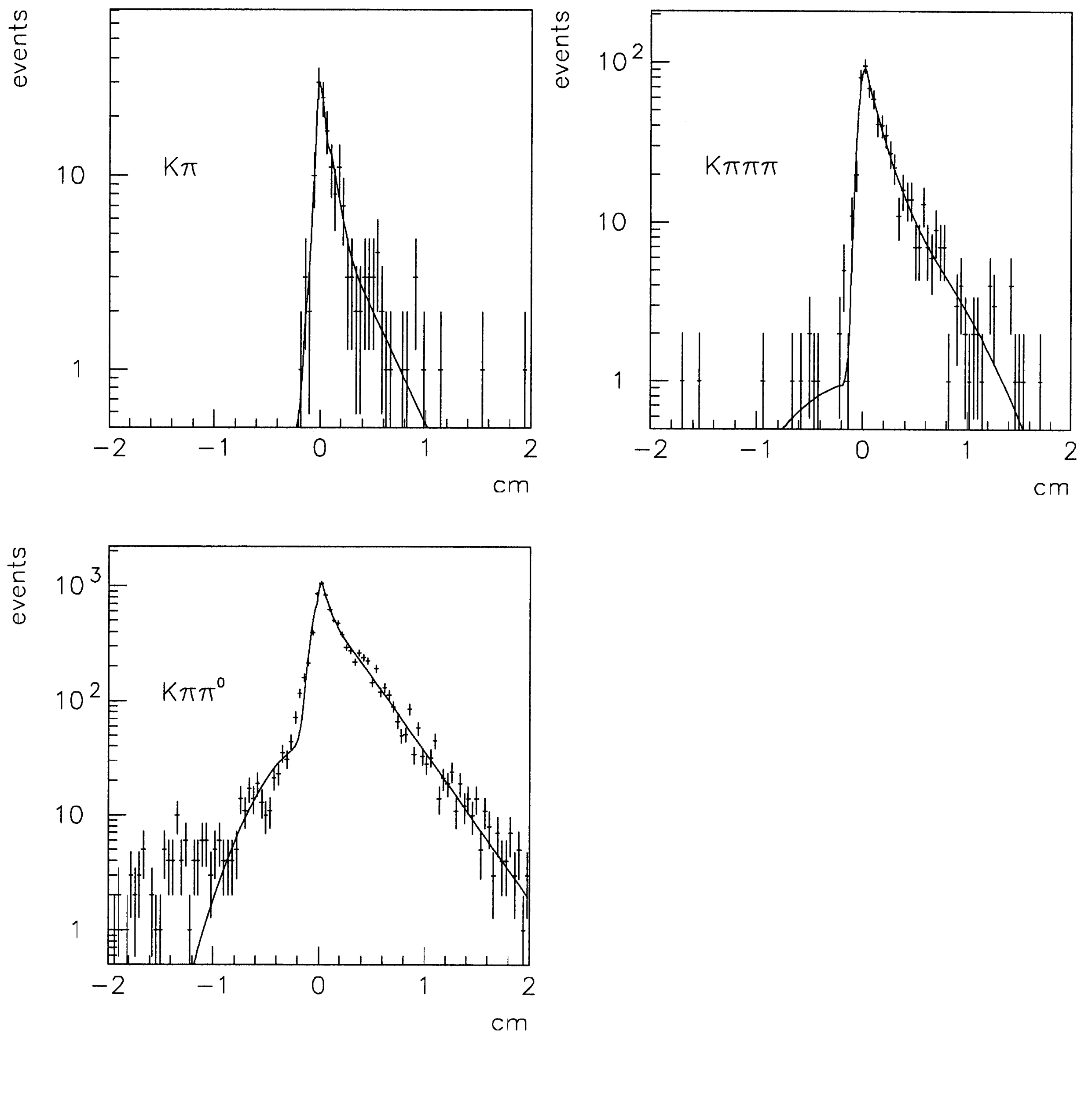


Figure 4: Decay length distributions of the combinatorial background

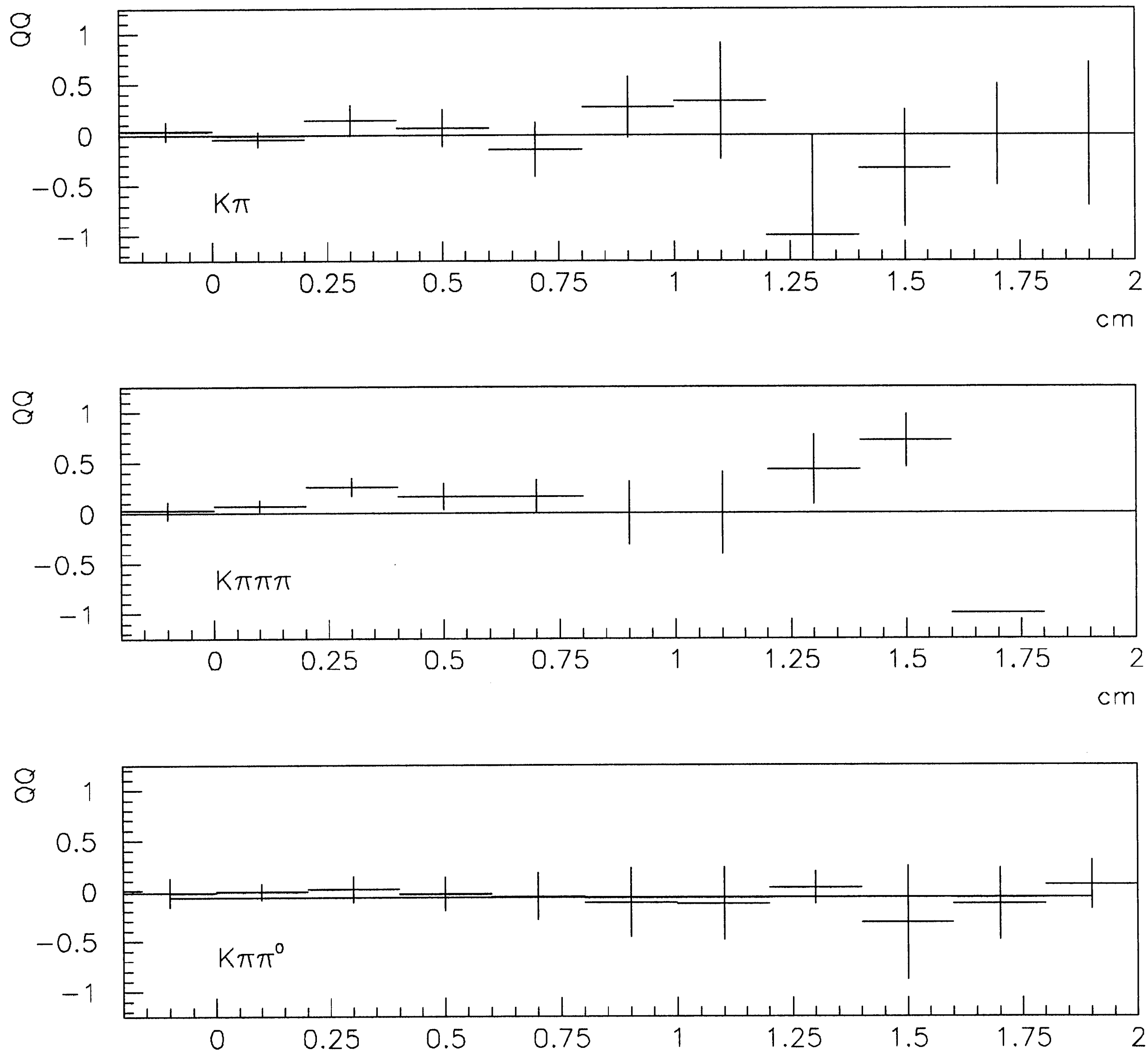


Figure 5: Charge correlation function of the combinatorial background. All the distributions are consistent with 0 in the total range of decay length

above. Altogether the charm background in the $K\pi$ sample is reduced to 9% according to the MC. Due to the different kinematical cuts the charm content is enhanced by a factor of 1.12 in the $K\pi\pi\pi$ and 1.18 in the $K\pi\pi^0$ samples.

source	45 GeV	30 GeV	25 GeV	20 GeV
beauty	1.00	0.98	0.91	0.77
charm	1.00	0.75	0.55	0.35
background	1.00	1.00	0.97	0.95
beauty	3.5 mm	3.4 mm	3.3 mm	3.2 mm
charm	1.7 mm	1.4 mm	1.2 mm	1.0 mm
background	1.0 mm	0.9 mm	0.9 mm	0.9 mm

Table 3: Efficiency of a cut on the maximal D^* momentum for D^* from B, $c\bar{c}$ and combinatorial background. Also listed is the average decay length of the D^0 after the cut

7 Resolution function

The resolution of the D^0 vertex is obtained using real data. For this pseudo D^0 are selected using track combinations around the D^0 mass satisfying the standard selection criteria (momentum, vertex χ^2). The distance of those tracks to the primary vertex is measured. If these tracks come from uds events, this distance should be 0. Any deviation from 0 measures the vertex resolution. Heavy flavour events are antitagged using a veto on leptons and requiring the uds probability from a lifetime b-tag to be greater than 90% in the hemisphere opposite to the selected tracks [14]. The distribution obtained is fitted with 3 gaussians, figure 6:

$$f(l-l_0) = \frac{A_1 exp\left(-\frac{(l-l_0)^2}{2\sigma_1^2}\right) + A_2 exp\left(-\frac{(l-l_0)^2}{2\sigma_2^2}\right) + A_3 exp\left(-\frac{(l-l_0)^2}{2\sigma_3^2}\right)}{\sqrt{2}(A_1\sigma_1 + A_2\sigma_2 + A_3\sigma_3)}. \quad (13)$$

The fit range was restricted to $-1\ cm < l < 0.2\ cm$ to reduce the contribution from remaining heavy flavour events. The values for the different subsamples are listed in table 4.

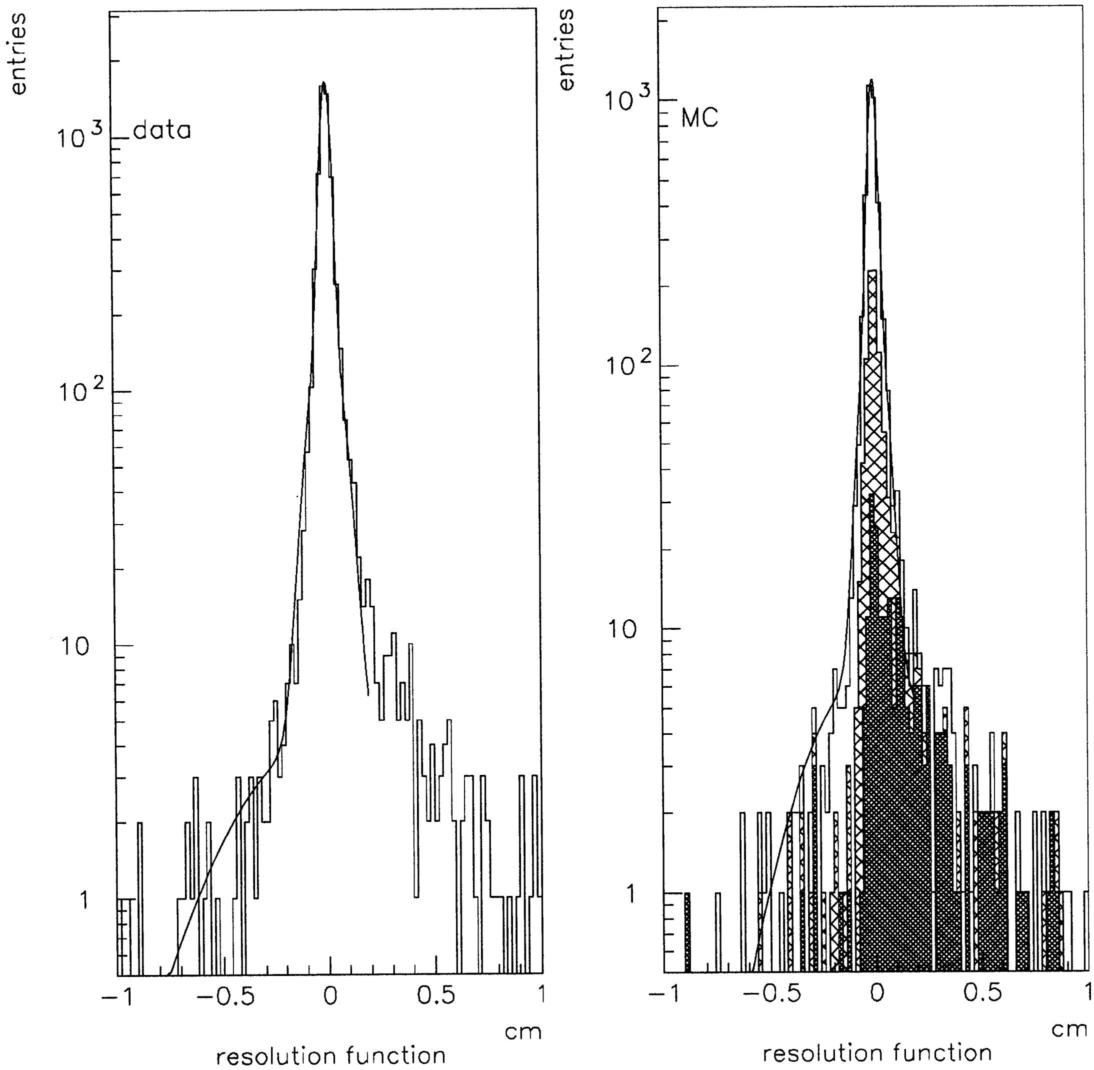


Figure 6: Resolution function. a) real data; b) MC events. The contribution of heavy flavour events is hatched (light = $c\bar{c}$, dark = $b\bar{b}$).

sample	$K\pi$	$K\pi\pi^0$	$K\pi\pi\pi$
A_1	1438	11905	2750
σ_1	0.022 cm	0.035 cm	0.030 cm
A_2	186	2073	401
σ_2	0.064 cm	0.1083 cm	0.083 cm
A_3	4.3	223	4.8
σ_3	0.36 cm	0.375 cm	0.43 cm

Table 4: parameters of the resolution function (13) for the different subsamples

8 Extraction of the oscillation frequency

8.1 Likelihood function

An unbinned maximum likelihood fit is performed to the like and unlike sign events. All three subsamples are fitted simultaneously. The likelihood function is:

$$P = exp\left(-\frac{(f_{B+} - f_{b+}^{exp})^2}{2\sigma(f_{b+}^{exp})^2}\right) exp\left(-\frac{(f_{c\bar{c}} - n_{c\bar{c}}^{exp})^2}{2\sigma(f_{c\bar{c}}^{exp})^2}\right) exp\left(-\frac{A - A_{exp})^2}{2\sigma(A_{exp})^2}\right)$$

$$\prod_{\nu=1}^{3}\left[\frac{e^{-\mu_{l\nu}}}{N_{l\nu}!}\mu_{l\nu}^{N_{l\nu}}\left(\prod_{i=1}^{N_{l\nu}}\frac{1}{n}\frac{dn}{dl}(l_i)_{l\nu}\right)\frac{e^{-\mu_{u\nu}}}{N_{u\nu}!}\mu_{u\nu}^{N_{u\nu}}\left(\prod_{i=1}^{N_{u\nu}}\frac{1}{n}\frac{dn}{dl}(l_i)_{u\nu}\right) exp\left(-\frac{(n_{bck,\nu} - n_{bck,\nu}^{exp})^2}{2\sigma(n_{bck,\nu}^{exp})^2}\right)\right], \tag{14}$$

with: $N_{l\nu}$, $N_{u\nu}$ the number of like and unlike sign events in the three subsamples ν, $\mu_{l\nu}$, $\mu_{u\nu}$ the expected (= fitted) number of events in the various subsamples, $n_{bck,\nu}$ the background, $n_{bck,\nu}^{exp}$ the background expected from the sidebands, $\sigma(n_{bck,\nu}^{exp})$ the error on the expected background; f_{B+} the fraction of D^* from charged B, f_{B+}^{exp}, $\sigma(f_{B+}^{exp})$ its expected value and error; $f_{c\bar{c}}$ the charm fraction, $f_{c\bar{c}}^{exp}$, $\sigma(f_{c\bar{c}}^{exp})$ its expectation and error. A is the lepton mistag fraction, A_{exp} its expected value and $\sigma(A_{exp})$ the error. The (normalized) shape of the decay length distributions $\frac{1}{n}\frac{dn}{dl}(l)$is given by (for each subsample ν, omitting this index in the following).
a) for unlike sign events:

$$\begin{aligned}\frac{1}{n}\frac{dn}{dl}(l)_u &= \frac{\mu_{B0}}{\mu_u}\left(\alpha\frac{1}{n}\frac{dn}{dl}(l)^{mixed} + (1-\alpha)\frac{1}{n}\frac{dn}{dl}(l)^b\right)\chi_d\,(1-A) \qquad (15)\\ &+ \frac{\mu_{B0}}{\mu_u}\left(\alpha\frac{1}{n}\frac{dn}{dl}(l)^{unmixed} + (1-\alpha)\frac{1}{n}\frac{dn}{dl}(l)^b\right)(1-\chi_d)\,A\\ &+ \frac{\mu_{B+}}{\mu_u}\frac{1}{n}\frac{dn}{dl}(l)^b\,A + \frac{\mu_{c\bar{c}}}{\mu_u}\frac{1}{n}\frac{dn}{dl}(l)^{c\bar{c}}\,(1-C)\end{aligned}$$

$$+ \frac{\mu_{bkg}}{\mu_u}\frac{1}{n}\frac{dn}{dl}(l)^{bck}\;(1-B_{bck}).$$

b) for like sign events:

$$\frac{1}{n}\frac{dn}{dl}(l)_l \;=\; \frac{\mu_{B0}}{\mu_u}\left(\alpha\frac{1}{n}\frac{dn}{dl}(l)^{unmixed} \;+\; (1-\alpha)\frac{1}{n}\frac{dn}{dl}(l)^b\right)(1-\chi_d)\,(1-A) \qquad (16)$$

$$+\; \frac{\mu_{B0}}{\mu_l}\left(\alpha\frac{1}{n}\frac{dn}{dl}(l)^{mixed} \;+\; (1-\alpha)\frac{1}{n}\frac{dn}{dl}(l)^b\right)\chi_d\,A$$

$$+\; \frac{\mu_{B+}}{\mu_l}\frac{1}{n}\frac{dn}{dl}(l)^b\,(1-A) \;+\; \frac{\mu_{c\bar{c}}}{\mu_l}\frac{1}{n}\frac{dn}{dl}(l)^{c\bar{c}}\,C$$

$$+\; \frac{\mu_{bkg}}{\mu_l}\frac{1}{n}\frac{dn}{dl}(l)^{bck}\;B_{bck},$$

μ_{B0},μ_{B+},$\mu_{c\bar{c}}$,μ_{bck} are the fitted values of the events from B^0, charged B, $c\bar{c}$ and combinatorial background, with $\sum_k \mu_k \;= \mu_l + \mu_u$. A is the lepton mistag fraction in B events, C in charm events and B_{bkg} in the combinatorial background. The fractions of charm and charge B are defined as:

$$f_{c\bar{c}} \;=\; \frac{\mu_{c\bar{c}}}{\mu_{B0}+\mu_{B+}+\mu_{c\bar{c}}},$$

$$f_{B+} \;=\; \frac{\mu_{B+}}{\mu_{B0}+\mu_{B+}}\;.$$

The overall sample composition can also be taken from table 5.

component	total	like	unlike
B^0	$(1-f_{bck})(1-f_c)(1-f_{B+})$		
B^0 mixed	$\chi(1-f_{bck})(1-f_c)(1-f_{B+})$	A	1-A
B^0 not mixed	$(1-\chi)(1-f_{bck})(1-f_c)(1-f_{B+})$	1-A	A
B^-	$(1-f_{bck})(1-f_c)f_{B+}$	1-A	A
charm	$(1-f_{bck})f_c$	C	1-C
background	f_{bck}	1-B_{bkg}	B_{bkg}

Table 5: composition of the sample in terms of like and unlike sign events

The shape distributions $\frac{dn}{dl}^{mixed,unmixed,b,c\bar{c},bck}$ for the contributions of mixed B, unmixed B, charged B, charm and combinatorial background are defined in the previous chapters. They depend explicitly on Δm (mixed, unmixed), τ_b (mixed, unmixed, b), τ_{D0} (mixed, unmixed, b, $c\bar{c}$). The mixing parameter χ_d is implicitly defined by the integrals of the mixed and unmixed distributions. The derivation is based on formulas 4 - 8.

The parameter α is normally set to 1, it can be varied between 0 and 1 to allow for a non time dependent component of mixing.

8.2 Fitting strategies

The fit is in principle able to fit for following free parameters (assuming $\alpha = 1$):

- The oscillation frequency Δm. The fit is mainly sensitive to the frequency Δm. Due to the number of like and unlike sign events there is a weak constraint on χ_d which measures $\frac{\Delta m}{\Gamma}$ †.
- The number of D^* from B^0, $B^\pm$ charm and background. The shapes of the respective decay length distributions are sufficiently different to allow their distinction by the fit, if $\Delta m \neq 0$.
- The mistag fraction A. It depends essentially on the observed oscillation amplitude which is reduced by (1-2A).

Of course some of these parameters (A, charm fraction, $B^\pm$) are known to some extend and this can be used to constrain the fit.
However, if $\alpha \neq 1$, that is there is a time independent component of mixing, a problem arises: The decay length distribution of mistagged $B^\pm$ is identical to a B^0 which has mixed. The two contributions become linearly dependent and the fit fails. This can be overcome by fitting only for $B^\pm$, leaving the mistag fraction A free. A then absorbs any mixing. Although the fit result is unphysical in this case, the difference in likelihood compared to the physical case with $\alpha = 1$ can be used to estimate the significance of the time dependent hypothesis.
The standard input parameters of the fit are listed in table 7.

8.3 Results

First the fit is performed for each subsample separately, with A, charm and $B^\pm$ fraction constrained to the expected values (see table 7):

$$K\pi \quad : \Delta m = 0.59 \pm 0.16\ ps^{-1} \tag{17}$$
$$K\pi\pi^0 \quad : \Delta m = 0.66 \pm 0.19\ ps^{-1} \tag{18}$$
$$K\pi\pi\pi \quad : \Delta m = 0.52 \pm 0.18\ ps^{-1}\ . \tag{19}$$

Then all subsamples are fitted simultaneously, still constraining the input parameters above:

$$\Delta m = 0.56^{+0.11}_{-0.10}\ ps^{-1}\ .$$

Finally A, the charm and $B^\pm$ fractions are free parameters:

$$\Delta m = 0.54 \pm 0.11\ ps^{-1} \tag{20}$$

†This is only true if the lepton mistag fraction is constrained. If the fit keeps the lepton mistag as a free parameter, no constraint on χ_d can arise from the event numbers.

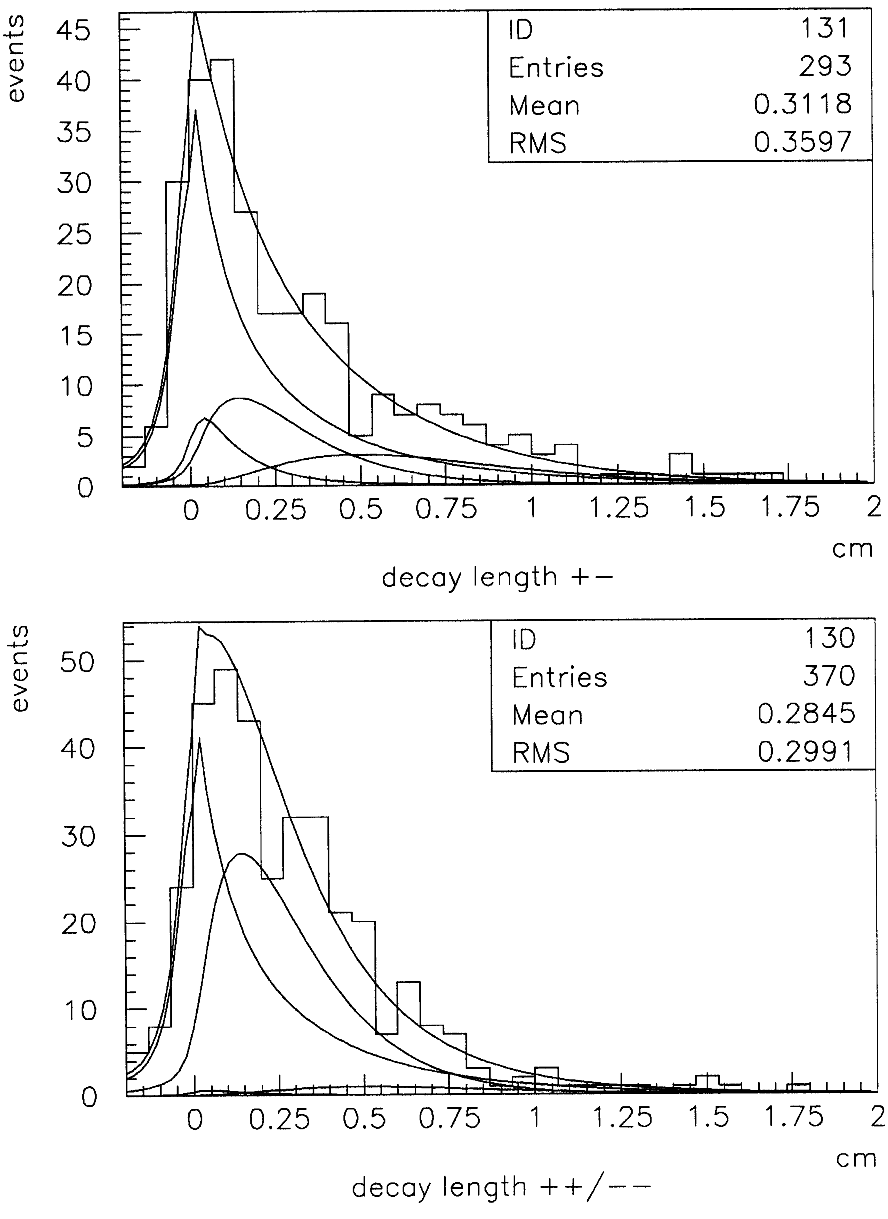


Figure 7: Result of the log likelihood fit to the like and unlike sign events. The histogram is the sum of all three subsamples. The contributions from background, charm, charged B, mixed and unmixed B^0 and the sum of all are indicated.

$$A = 0.238^{+0.081}_{-0.074} \quad (21)$$
$$f_{c\bar{c}} = 0.06 \pm 0.06 \quad (22)$$
$$f_{B+} = 0.00^{+0.22}_{-0.00} \,. \quad (23)$$

Despite fewer constraints the loss in precision is very small. The fit results of the lepton mistag A, the $B^\pm$ and charm fraction are very close to the expected values (for comparison, see table 7), providing a consistency check.
The result of this fit is shown in figure 7. The charge correlation function calculated from this fit is shown in figure 8.
If this fit is repeated using the time independent model, the difference in the log likelihood is 5.2, corresponding to 3.2 standard deviations or 5 10^{-4} probability.

9 Systematic errors

Following systematic errors were studied:

- **B-lifetime**: As the fit measures the oscillation frequency Δm the value of the B^0 lifetime gives only a small error. Converting into $\frac{\Delta m}{\Gamma}$ one has to consider the error on the B^0 lifetime. Despite the average B lifetime has been measured very precisely [15], the knowledge on the B^0 lifetime is less precise. Exclusive lifetimes were measured by ALEPH [16] and DELPHI [17]. Combining the results one obtains:
$$\tau_{B^0} = 1.40 \pm 0.18 \; ps,$$
$$\tau_{B^\pm} = 1.41 \pm 0.21 \; ps,$$
which are used for estimating the systematic error. Changing the input lifetime by its error of 0.18 ps changes the result on Δm by $\pm$ 0.006 ps^{-1}, whilst $\frac{\Delta m}{\Gamma}$ varies by $\pm$ 0.09.
The uncertainty of the D^0 lifetime (0.420 $\pm$ 0.008 ps) gives an error on Δm of ± 0.006 ps^{-1}.

- **$B^\pm$ fraction**: To date there exist no measurements of the inclusive branching ratios $B^0 \rightarrow D^{*-}X$ and $B^+ \rightarrow D^{*-}X$. In the simple spectator model one would expect that $B^0 \rightarrow D^{*-}X$ dominates over $B^+ \rightarrow D^{*-}X$, hence most of the D^* come from B^0. This is true in semileptonic decays as can be seen in the B^0-B^+ lifetime analysis [16] where the fraction of D^* from $B^\pm$ in D^*-lepton events was evaluated. Correcting for efficiencies the B^+ fraction is:
$$f_{B+} = 0.16^{+0.10}_{-0.16} \,.$$
This should be a reasonable assumption for the inclusive D^* sample. The presence of internal spectator decays of the $B^\pm$ (which are absent in semileptonic decays) should in first order only give D^{0*} events. As this parameter is allowed

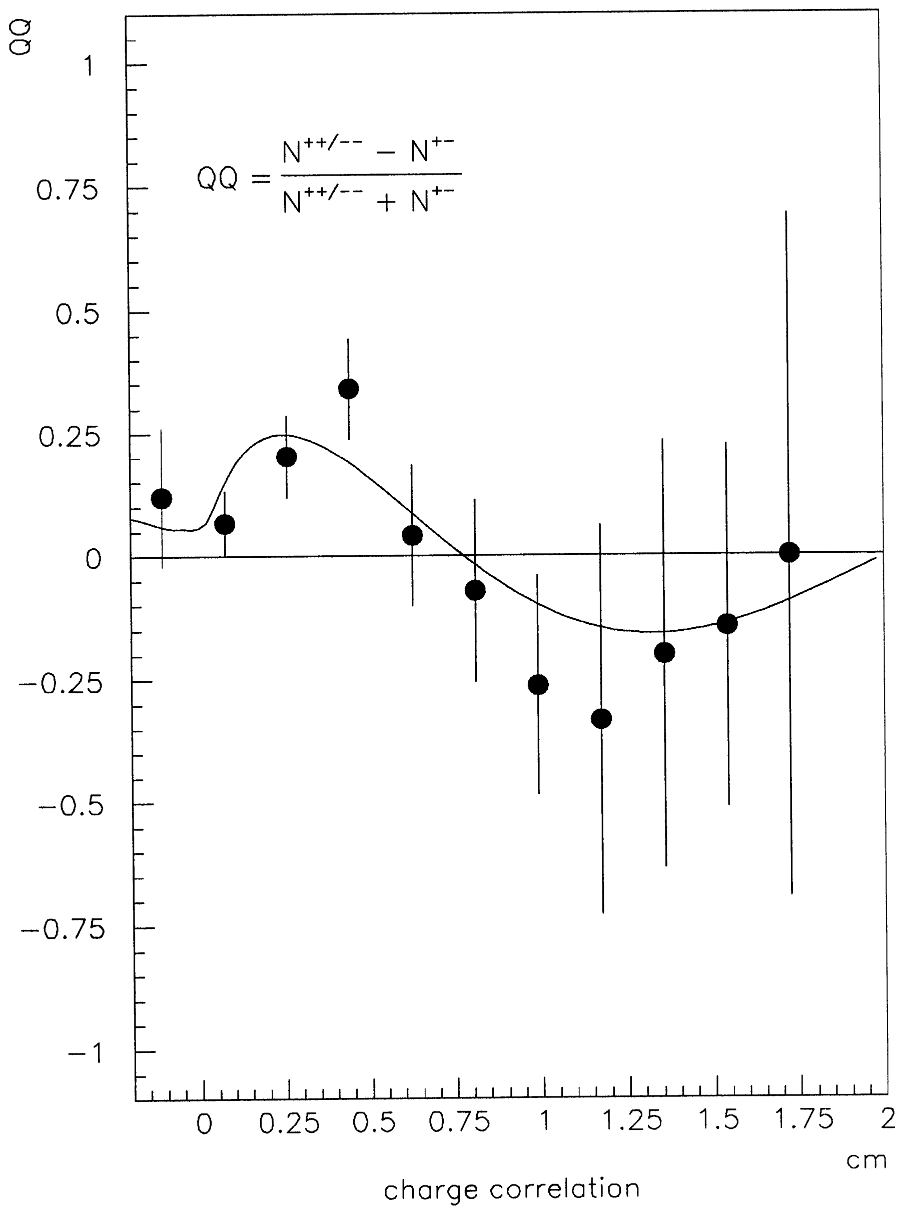


Figure 8: Charge correlation function QQ of the data (points). The result of the fit is indicated as solid line

to vary in the fit within its error, the systematic error is already included in the statistical error of the fit. In order to show the sensitivity of the result to the $B^{\pm}$ fraction, it was fixed to the expected value and varied by ± 0.16. This results in changes of Δm of $\pm$ 0.03 ps^{-1}. The unconstrained fit prefers a $B^{\pm}$ contribution of $0.0^{+0.22}_{-0.00}$, within the error consistent with the expectations.

- **Charm fraction**: Several sources of uncertainty can contribute to the error on the charm fraction. Firstly the error on the probability to get a D^* in a charm or beauty events [18]:

$$\frac{b}{c} \;=\; 0.87^{+0.15}_{-0.13}\;(stat)\;\pm 0.01\;(sys).$$

  Furthermore one has to consider the error on $BR(c \rightarrow l)$ ($\pm$ 5 %), charm fragmentation (for the D^* selection) ($\pm$ 9 %) (both from [19]) and the statistics of the MC ($\pm$ 10 %). Altogether the error on the charm fraction is $\pm$ 22 %. Also this systematic error is already include in the statistical error of the fit. Fixing it to the expected value and changing it by ± 0.22 introduces a change of $^{+0.007}_{-0.006}$ ps^{-1}. Leaving the charm fraction free results in 0.06 ± 0.06, in good agreement with the expectations.

- **Lepton mistag**: The mistag probability (with sign change) from MC is:

$$A \;=\; 0.245 \pm\; 0.010\;.$$

  This variation is already included in the statistical error of the fit. Fixing A to its expected value and changing by $\pm$ 1 s.d. results in a systematic error on Δm of $^{+0.006}_{-0.007}$ ps^{-1}. Leaving A as a free parameter the fit finds

$$A \;=\; 0.238^{+0.074}_{-0.081}\;,$$

  which is in excellent agreement with the MC prediction.

- **Fragmentation functions**: The error on the charm fragmentation affects mainly the selection efficiency and therefore the charm fraction, hence the error is treated there. Concerning the b fragmentation the B^0 momentum spectrum is affected, which enters directly in the measured value of Δm. We use $< x_b >= 0.712 \pm 0.012$ from reference [19] resulting in $\sigma(\Delta m) \;=\; 0.009\; ps^{-1}$.

- **Resolution function**: Various resolution function have been used to estimate the systematics, e.g. taking the real MC resolution etc, giving variations of ± 0.007 ps^{-1}.

- **Background subtraction**: Changing the expected background by ± 1 s.d. gives an error of ± 0.001 ps^{-1} on Δm. The uncertainty of the background charge correlation introduces an additional error of $^{+0.009}_{-0.022}$ ps^{-1}.

The systematic errors are summarized in table 6. The total systematic error on Δm is:

$$\sigma(\Delta m)_{sys} = {}^{+0.037}_{-0.042}\ ps^{-1}.$$

quantity	input error	error on Δm
$B^{\pm}$ fraction(*)	0.16 ± 0.16	$\pm\ 0.03\ ps^{-1}$
charm fraction(*)	0.09 ± 0.02	${}^{+0.006}_{-0.007}\ ps^{-1}$
lepton mistag(*)	0.245 ± 0.01	${}^{+0.006}_{-0.007}\ ps^{-1}$
B^0 lifetime	1.40 ± 0.18 ps	$\pm\ 0.006\ ps^{-1}$
D^0 lifetime	0.42 ± 0.008 ps	$\pm\ 0.004\ ps^{-1}$
fragmentation x_b	0.712 ± 0.012	$0.009\ ps^{-1}$
resolution	various	$\pm\ 0.007\ ps^{-1}$
background norm	1 s.d.	$\pm\ 0.001\ ps^{-1}$
backg. QQ	1 s.d.	${}^{+0.009}_{-0.022}\ ps^{-1}$
total		${}^{+0.016}_{-0.026}\ ps^{-1}$

Table 6: Systematic errors on Δm. The errors due to the parameters marked (*) are already included in the statistical error of the fit. They are listed here to show the sensitivity of the result to these parameters

10 Conclusions

Using D^*-lepton correlations the characteristic time dependence of $B^0_d \bar{B}^0_d$ oscillations has been observed. An alternative model without time dependence of the transition probability is disfavoured by 3.2 standard deviations. The measured mass difference Δm is:

$$\Delta m\ = [3.65\ {}^{+0.76}_{-0.69}\ (stat)\ {}^{+0.11}_{-0.17}(sys)\ ]\ 10^{-4}\ eV.$$

This is a direct measurement of the mass difference. Using $\tau_{B^0} = 1.40 \pm 0.18$ ps, this corresponds to a value of $\frac{\Delta m}{\Gamma}$ of:

$$\frac{\Delta m}{\Gamma}\ =\ 0.77\ {}^{+0.16}_{-0.15}\ (stat)\ \pm\ 0.10\ (sys).$$

The extra error of due to the B^0 lifetime is included in the systematic error.

References

[1] C. Albajar et al., Phys. Lett. B186 (1987) 247

parameter	value	error	fit result
$\hat{\chi}$	0.128	0.01	
τ_b (ps)	1.40	0.18	
f_c: cc/(cc+bb) $K\pi$	0.09	0.02	0.08 $\pm$ 0.02
f_c: cc/(cc+bb) $K\pi\pi\pi$	0.10	0.02	0.09 $\pm$ 0.02
f_c: cc/(cc+bb) $K\pi\pi^0$	0.11	0.02	0.09 $\pm$ 0.02
f_{B+}: $B^+/(B^0+B^+)$	0.16	0.16	0.10 $\pm$ 0.08
background (absolute) $K\pi$	15.5	1.9	15.6 $\pm$ 2.0
background (absolute) $K\pi\pi\pi$	91.	9.	81.0 $\pm$ 10.0
background (absolute) $K\pi\pi^0$	178.	18.	179 $\pm$ 16.
QQ(background) $K\pi$	0.00	0.06	
QQ(background) $K3\pi$	0.080	0.038	
QQ(background) $K\pi\pi^0$	0.007	0.030	
2. generation leptons	0.139	0.004	
decays (b-events)	0.060	0.003	
unlike fraction (b-decays)	0.46	0.04	
decays (charm events)	0.252	0.049	
unlike fraction (c-decays)	0.615	0.12	
mistag b: A (computed)	0.245	0.010	0.244 $\pm$ 0.010
mistag c: C (computed)	0.088	-	
mistag bck: B_{bkg}	0.503	-	

Table 7: input parameters for likelihood fit. Some parameters were allowed to vary in the fit within the errors of the expectations. The results of the fit are listed in this case

[2] H. Albrecht et al., Phys. Lett. B192 (1987) 245

[3] M. Artuso et al., Phys. Rev. Lett. 62 (1989) 2233

[4] D. Buskulic et al. (ALEPH), Phys. Lett B284 (1992) 177
D. Decamp et al. (ALEPH), Phys. Lett. B258 (1991) 236
B. Adeva et al. (L3), Phys. Lett. B288 (1992) 395
P.D. Acton et al. (OPAL), Phys. Lett. B276 (1992) 379
P. Abreu et al. CERN-PPE 92-203
C. Albajar et al. (UA1), Phys. Lett. B262 (1991) 171

[5] H.-G. Moser: *The possibility to measure the time dependence of $B_s\bar{B}_s$ oscillations using Fourier analysis*, Nucl. Inst. & Meth. A295 (1990) 435

[6] P. Roudeau: *Measurement of $B_s\bar{B}_s$ oscillations using semileptonic decays*, Report of the working group on high luminosity at LEP, CERN 91-02 (1991) 261

[7] A. Ali and D. London: *Prospects for Measuring the $B_s\bar{B}_s$ mixing ratio χ_s*; DESY 92-075, 1992

[8] R. Forty, ALEPH note in preparation

[9] H. Schroder in *B Decays*; edt. by S. Stone, World Scientific, Singapore (1992), page 292

[10] A.J. Buras, W. Slominski and H. Steger, Nucl. Phys. B245 (1984) 369

[11] A. Ali in *B Decays*; edt. by S. Stone, World Scientific, Singapore (1992), page 1

[12] D. Brown: *QFNDIP, a primary vertex finder*; ALEPH 92-47, PHYSIC 92-42

[13] A. Abbaneo et al. *Measurement of the $Z \rightarrow b\bar{b}$ forward backward asymmetry and the χ mixing parameter from high p_t leptons*; ALEPH 93-24, PHYSIC 93-16 (1993)

[14] D. Brown et al.: *Tagging b hadrons using track impact parameters*; ALEPH 92-135 PHYSIC 92-124 (1992)

[15] D. Buskulic et al.: *Updated measurement of the average b hadron lifetime*; Phys. Lett. B295 (1992) 174

[16] D. Buskulic et al.: *Measurement of the $\bar{B}^0$ and B^- meson lifetimes*; to be published in Phys. Lett. B

[17] P. Abreu et. al.: *A Measurement of the B meson Production and Lifetime using Dl^- Events in Z^0 decays*; CERN-PPE/92-174 (1992)

[18] A. Bencheikh et al.: *Charm Physics Results with 91 Data*; ALEPH 92-39, PHYSIC 92-36

[19] A. Falvard et al.: *Production and Semileptonic decays of Heavy Quarks at the Z*; ALEPH 93-30, PHYSIC 93-21