ALEPH 95-079
PHYSIC 95-73
M.-C. Lemaire and A. Roussarie
5.7. 95

Time dependent B_s mixing from lepton-kaon correlations

M.-C. Lemaire and A. Roussarie
DAPNIA/SPP, CE Saclay
91191 Gif-sur-Yvette Cedex, France

July 6, 1995

Abstract

A lower limit is set on the B_s^0 meson mixing parameter Δm_s using data collected from 1991 to 1994 by the ALEPH detector. A new method is presented where the time dependent charge asymmetry is measured in lepton-kaon correlations. Events containing a high p, p_T lepton and a fragmentation kaon are selected. The kaon charge, associated with the jet charge in the opposite hemisphere, tags the b quark charge at production while the lepton tags the B_s at decay. Topological vertexing is used to define the B vertex from which the decay time is measured. 4436 lepton-kaon correlations are selected, in which the B_s mistag fraction is as low as 19%. The selection of the kaon enriches the B_s sample by a factor 1.35. Two methods are used to search for the B_s^0 mixing: the well known log-likelihood difference method and a new method, the fitted amplitude method, proposed to allow an easy combination of various results. They give identical limits. Assuming the B_s fraction to be $(12\pm3)\%$, a lower limit of $\Delta m_s > 4.0 ps^{-1}$ (95% C.L.) is set by these methods.

1 Introduction

Like neutral kaons, neutral B_0 oscillate. Neglecting CP violation, the mass eigenstates B_1 and B_2 are expressed in terms of flavor eigenstates B_0 and $\overline{B_0}$ as following:

$$B_1 = \frac{B_0 + \overline{B_0}}{\sqrt{2}} \quad B_2 = \frac{B_0 - \overline{B_0}}{\sqrt{2}}$$

If a B_0 is produced at time t=0, the probability to have a B_0 at time t is:

$$P_{B_0 \to B_0} = \frac{1 + cos\Delta mt}{2} \cdot e^{-\frac{t}{\tau}}$$

where τ is the B_0 lifetime and Δm is the mass difference between the two mass eigenstates B_0 and $\overline{B_0}$. In the standard model mixing is calculated via box diagrams, involving top quark and W exchange, Δm_d and Δm_s depend on the Cabibbo Kobayashi Maskawa(CKM) matrix elements V_{td} and V_{ts} [1] but also on the top mass and poorly known strong interaction parameters: the bag parameter, B decay constant and QCD correction factors. Most of these uncertainties cancel in the ratio:

$$\frac{\Delta m_s}{\Delta m_d} = (1.16 \pm 0.10)^2 \left[\frac{V_{ts}}{V_{td}}\right]^2$$

where the estimate of the coefficients has been taken from [1]. This measurement is particularly important in connexion to the unitarity triangle since the ratio $|V_{td}/\lambda V_{ts}|$ (where λ is the sine of the Cabibbo angle) is a measure of one side.

Various 95% confidence level lower limits have been set on Δm_s, first with dilepton analyses by ALEPH at 3.9 ps^{-1} [2] and OPAL at 2.2 ps^{-1} [3]. The ALEPH lepton-jet charge method [4] has raised this limit to $\Delta m_s > 6.1 ps^{-1}$.

The present paper reports a search for the time dependent B_s mixing, using a high p_T lepton for tagging the b quark charge at decay time an a combination of the fragmentation kaon charge with the jet charge in the opposite hemisphere for tagging the charge at production time. Excluding the lepton, topological vertexing is used to reconstruct a 'charm' track which vertexed with the lepton gives the B vertex from which the decay length of the B hadron is measured. Combined with an estimation of the boost of the B hadron, it gives the proper decay time. The charge asymmetry of the lepton-kaon system is observed as a function of the reconstructed B_s proper time. To set a limit on Δm_s, a new method has been used. For any given Δm_s, a B_s oscillation amplitude $\mathcal{A}$ is fitted, through a maximum likelihood method, by comparing the observed charge asymmetry time dependence to the expected one in which the B_s oscillation would be rescaled by a factor $\mathcal{A}$. The fitted value of $\mathcal{A}$ is therefore compared to either 0 (no mixing is observed) or 1 (the mixing is observed for this value of Δm_s). This method is more pedagogical to display the mixing visibility and will allow an easy combination of mixing results originating from various analyses and experiments.

2 Event selection, proper time determination and initial flavor tagging

2.1 Event selection

The ALEPH detector has been described in detail elsewhere [5] with its performance [6]. The present analysis is based on 3 million hadronic Z^0decays recorded during the

1991-1994 data runs.

In order to work with well contained events in the detector, each hadronic decay is required to have $\mid \cos(\Theta_{thrust}) \mid \leq 0.85$, where Θ_{thrust} is the polar angle of the thrust axis. Events are selected as having at least one high momentum lepton $p > 3$ GeV/c. In each event, jets are reconstructed using charged and neutral particles (determined with an energy flow algorithm and clustered using the scaled-invariant mass technique with a clustering parameter $Y_{cut} = (6\ GeV/E_{cm})^2$. The event is required to have at least two jets. The lepton momentum must be less than 90% of the energy of the jet to which it belongs and the charged multiplicity in the lepton jet should be greater or equal to 3. The transverse momentum p_T of the lepton from its associated jet is calculated by excluding the lepton from the jet. At least one lepton must have a $p_T > 1.25\ GeV/c$. If in an event several leptons belonging to the same hemisphere satisfy all the criteria, only the highest transverse momentum one is selected. The events are then classified in double tagged and single tagged events corresponding respectively to events where each hemisphere contain a high p and high p_T lepton and events where only one hemisphere has such a lepton.

Large simulated sample of hadronic events (3.7 millions of $q\bar{q}$ and 0.57 million of $b\bar{b}$) have been analyzed. The Monte Carlo generator is based on JETSET 7.3 with updated branching ratios and using the Korner-Schuler model for semileptonic b decays. For single tagged events, the composition of the lepton sample is shown in table 1.

Table 1: Sample composition of the single tagged lepton events

event type	fraction %	
	e	μ
$b \rightarrow l$	82.7	73.5
$b \rightarrow \tau \rightarrow l$	1.2	1.1
$b \rightarrow c \rightarrow l$	7.7	7.9
$c \rightarrow l$	6.0	6.3
$K, \pi \rightarrow \mu$	-	6.3
photon conversions	1.0	-
Misid. hadron	1.1	4.0

2.2 Proper time determination

The proper time of a b hadron decay is given by:

$$t = gd = \frac{m_B}{p_B c} d$$

where g is the boost factor, m_B and p_B are respectively the B hadron mass and momentum, and d is the decay length which is determined after reconstruction of the interaction point and of the B decay vertex. The method used to reconstruct the event primary vertex is made insensitive to the lifetime information of the tracks by projecting them onto the plane perpendicular to the jet to which they belong and combines this with the envelope of the beam spot. The centre of the beam spot is periodically determined from hadronic events reconstructed and analyzed over 75 successive events. Using this algorithm on simulated $b\bar{b}$ events the average resolution on the position of the primary vertex projected along the b flight direction is 90μ.

The secondary vertex is determined as follows: excluding the lepton, all the charged tracks of the lepton hemisphere are used to search for a decay vertex. This is achieved by determining for each track the difference in χ^2 between assigning this track to the primary vertex or allowing it to originate from a second vertex. The candidate decay vertices are placed on a grid and the point which maximizes the sum of the difference in χ^2, made over the tracks of the lepton hemisphere except the lepton, is called the 'charm' vertex. Tracks within 3σ of the charm vertex are combined to form a charm track, requiring that there is at least two such tracks. The charm track is finally vertexed with the lepton track to give the B decay point. To improve the vertex resolution, the lepton track is required to have at least one vertex detector hit in both $r\phi$ and rz projections. The reconstructed decay length d is taken as the distance between the primary vertex and the B decay vertex projected on the B track direction.

In the Monte Carlo simulation one can compare the measured distance d_{reco} to the generated one d_{true} as a function of the true time t_{true} ($t_{true} = g_{true} d_{true}$ with $g_{true} = m_b/p_B$). It is found that the decay length resolution depends on the true time. It displays negative tails corresponding to fragmentation tracks wrongly assigned to the secondary vertex. To minimize this bias, it is required that the distance of the charm vertex to the primary vertex exceeds 300μ. The distributions of $g_{true}(d_{reco} - d_{true})$ obtained after this cut are displayed on fig.1. The negative tail, increasing with the B true time, is still visible but has been reduced by a factor 2 by the charm cut. The result of a three gaussian fit to the distributions of fig.1 are given in table 2, together with the cut efficiency.

Table 2: Average charm cut efficiency and result of the 3 gaussian fit to the distribution of $g_{true}.(d_{reco} - d_{true})$, for four slices of the B true proper time after the charm distance cut. F_i,M_i and σ_i are the fraction, mean and rms of gaussian number i. In the fit, M_2=M_1 is assumed.

Time slice (ps)	ϵ	F_1	M_1	σ_1	F_2	σ_2	F_3	M_3	σ_3
< 0.2	0.607	0.52	0.05	0.14	0.31	0.38	0.17	0.20	0.97
0.2-0.6	0.771	0.51	0.03	0.12	0.35	0.39	0.14	0.01	1.23
0.6-1.5	0.854	0.46	0.02	0.13	0.37	0.36	0.16	-0.20	1.18
> 1.5	0.976	0.48	0.01	0.13	0.34	0.38	0.18	-0.51	1.63

Table 3: Result of a three gaussian fit to the distribution $(g_{true} - g_{reco})/g_{true}$ with the same definitions as in table 2.

F_1	M_1	σ_1	F_2	M_2	σ_2	F_3	M_3	σ_3
0.21	0.05	0.05	0.44	0.02	0.11	0.34	-0.03	0.25

The knowledge of the B hadron momentum is needed to compute it's proper time. It is derived, as described in the ALEPH dilepton analysis [2], from the sum of three contributions obtained from charged particles, missing neutrino momentum and part of the neutral energy of the lepton jet. For the charged particles, the momentum of the lepton and of all the tracks of the charm vertex are summed. The missing neutrino momentum is the difference between the beam energy and the visible energy in the hemisphere containing the lepton [7]. Finally, 68 % of the neutral energy in the lepton jet is assigned

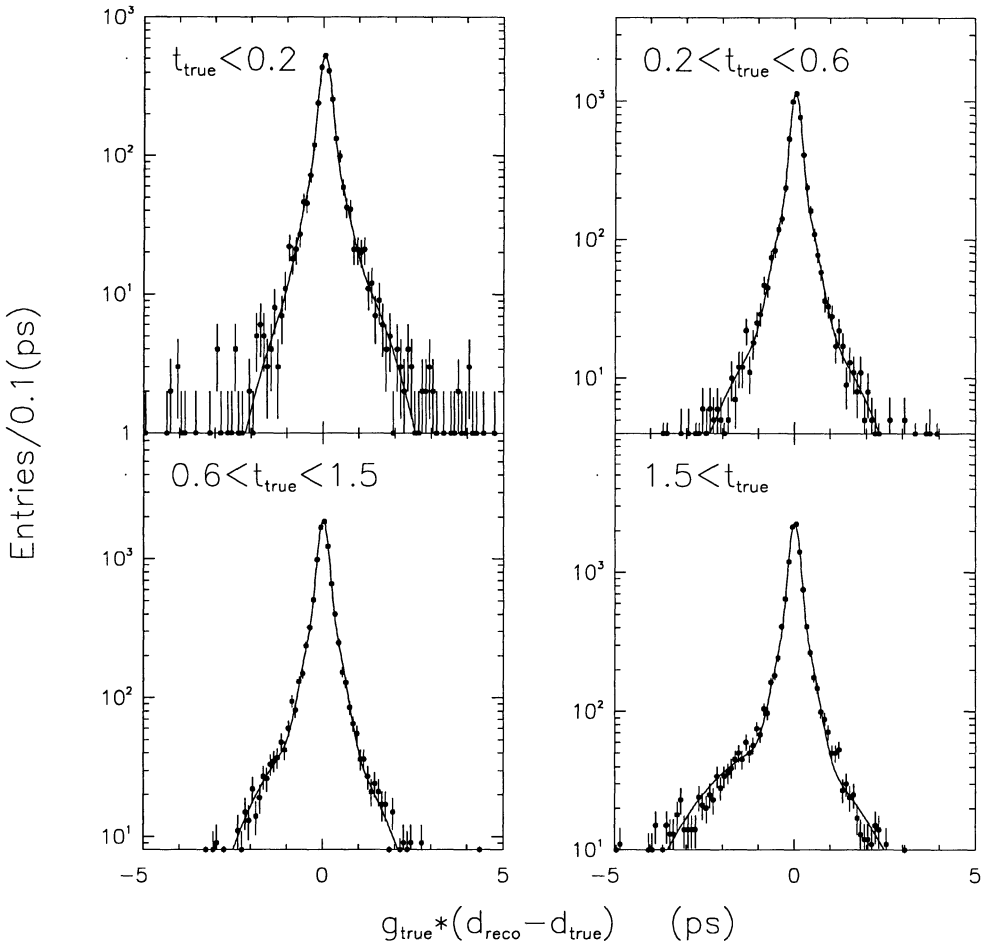

Figure 1: time resolution for different true time slices

to the b momentum. This fraction has been estimated from the Monte Carlo simulation. The resolution in the boost parameter g is plotted in fig.2 and the parameters of a three gaussian fit are listed in table 3. The cut on the charm distance which effectively reduces the contribution of the negative tails has also the following consequences: it creates a time dependence of the selection efficiency with t_{true} which is parametrized from the simulation and used later in the fit and it makes the boost resolution independent of the true time.

The determination of the time resolution (decay length and boost) and of the selection efficiency has been done independently for cascade and primary b leptons.

2.3 b charge tagging at production time

From strangeness conservation, the B_s meson hadronizes in association with the production of a fast strange hadron. While the lepton tags the b charge at decay time, the fragmentation kaon allows to tag the b charge at production time and enriches the sample in B_s. Charged tracks of momentum greater than 1.5 GeV are identified as kaons if

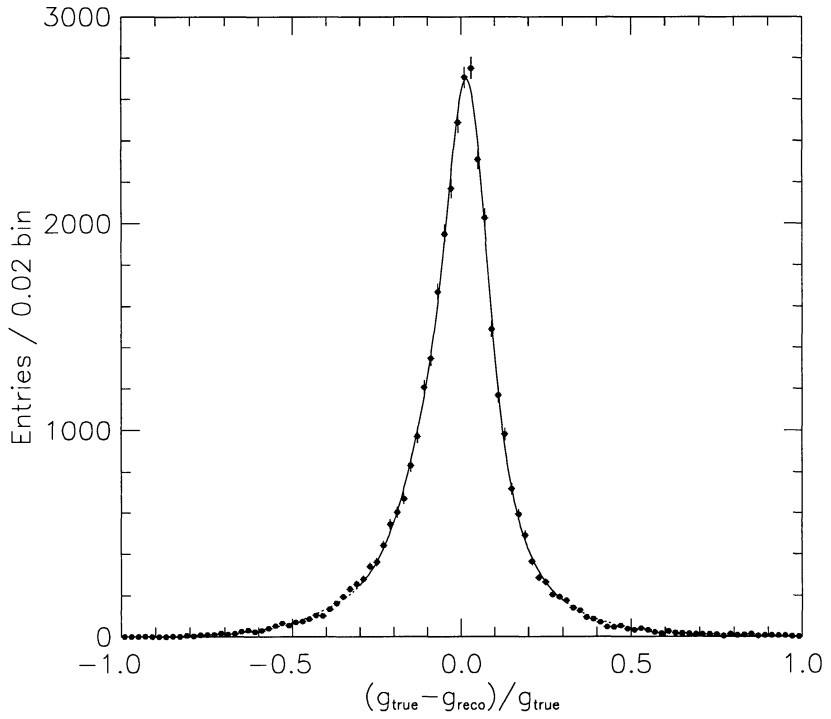

Figure 2: boost resolution

the sum of the two values of the dE/dx estimator, assuming first the particle to be a kaon and then to be a pion, is negative. The fragmentation kaon is taken as the most energetic kaon at the primary vertex. In these conditions, 22% of the lepton hemispheres contain such a kaon.

To enrich the secondary vertex with D_s, it is required that this vertex has either 0 or 2 kaons or if it has only one kaon, it's charge must be opposite to the lepton charge.

To reduce the mistag rate the product of the kaon charge and the b charge in the opposite hemisphere is required to be negative. This opposite hemisphere charge is the lepton charge for dileptons events or the jet charge for single lepton events (calculated with momentum weighting to a power $\kappa = 0.5$).

From the Monte Carlo simulations, the fraction of B_s is determined to be 0.165 ± 0.005 while the input production fraction is 0.122. The kaon selection enriches the B_s sample by a factor 1.35 ± 0.04 while the B_d fraction is only changed by a factor 1.01 ± 0.02. As it will be shown in more detail in section 4, the mistag rate achieved in B_s events is as low

as $19.0 \pm 1.3\%$. With this selection 4436 lepton-kaon correlations have been measured in the 91-94 data. The sample composition is given in table 4.

Table 4: Sample composition: f_b is the fraction of B events in the total sample, f_{bc} the fraction of cascade in the B sample, f_s and f_d the fraction of B_s ad B_d in the primary B lepton sample

source	fraction
f_b	0.921
f_{bc}	0.098
f_s	0.165
f_d	0.385

3 Likelihood fit

3.1 Formalism

A l^-K^- or l^+K^+ correlation is called good sign (G) and tags an unmixed event. A l^-K^+ or l^+K^- pair is called wrong sign (W) and tags a mixed event. The log-likelihood is the sum:

$$\ln L = \ln L_G + \ln Lw$$

Summing the contributions of the measured proper time bins, one can write:

$$\ln L_{G,W} = -\sum N_i^{G,W} \ln D_i^{G,W}$$

where $N_i^{G,W}$ are the measured rates and $D_i^{G,W}$ are the expected time distribution probabilities:

$$D_i^{G,W} = f_b[(1-f_{bc})P_{b\to l}^{G,W}(t_m^i) + f_{bc}C_{b\to c\to l}^{G,W}] + (1-f_b)P_{bkg}(t_m^i)$$

which explicits three components, primary and cascade b decay leptons and non b (called background).

The $(b \to l)$ probabilities are:

$$P_{b\to l}^{G,W}(t_m^i) = f_d P_{B_d}^{G,W}(t_m^i, \Delta m_d) + f_s P_{B_s}^{G,W}(t_m^i, \Delta m_s) + (1-f_d-f_s)P_{B_u}^{G,W}(t_m^i),$$

where the label B_u means the sum of the B_u meson and b baryon contributions. A similar expression can be written for the cascade probabilities $C_{b\to c\to l}^{G,W}(t_m^i)$

For $b \to l$ we have:

$$P_f^{G,W}(t_m) = \int P_f^{G,W}(t)\, R_b(t_m, t)\, \epsilon_b(t)\, e^{-\frac{t}{\tau_f}}\, dt$$

where $R_b(t_m, t)$, the time resolution function, is the probability to measure t_m when the true value is t, ϵ_b is the selection efficiency, function of the true time. The index f stands for d and s.

For $b \to c \to l$ we have a similar expression:

$$C_f^{G,W}(t_m) = \int C_f^{G,W}(t)\, R_{bc}(t_m, t)\, \epsilon_{bc}(t)\, e^{-\frac{t}{\tau_f}}\, dt$$

The time dependence of the charge correlation is given by the expressions:

$$P_f^G(t) = \frac{1 + A_f + B_f cos(\Delta m_f t)}{2}$$

$$P_f^W(t) = \frac{1 - A_f - B_f cos(\Delta m_f t)}{2}$$

$$P_u^G(t) = \frac{1 + A_u}{2}$$

$$P_u^W(t) = \frac{1 - A_u}{2}$$

For all flavours, f = d, s and u, it is assumed that the effect of the cascade $b \rightarrow c \rightarrow l$ is only to change the sign of the lepton. Then:

$$C_f^G(t) = P_f^W(t)$$

$$C_f^W(t) = P_f^G(t)$$

3.2 Monte Carlo values of the parameters

The coefficients A and B correspond respectively to the time independent and time dependent charge correlations asymmetries. The B_s and A_s parameters are determined from the Monte-Carlo for pure $B_s \rightarrow l$ events, similarly B_d and A_d are determined from a pure Monte Carlo $B_d \rightarrow l$ sample while a B_u mesons and b baryons sample is used to determine A_u.

The corresponding numbers are displayed on table 5 with the statistical errors of the Monte Carlo. Lepton-kaon charge correlations are induced by several ways. If the kaon

Table 5: A_f and B_f (where f stands for u,d,s) asymmetry parameters

Flavour	B	A
B_u	0.	0.285 ± 0.014
B_d	0.356 ± 0.043	0.053 ± 0.033
B_s	0.600 ± 0.055	0.022 ± 0.025

and the lepton, originate from the same B decay, the correlation is time-independent. This is a small effect, as it can be seen from the small values of A_s and A_d. If the kaon charge is correlated to the b quark charge at production time two cases occur:

- The correlation is time independent for B which do not mix (B_u meson and B baryons). This explains the large value of A_u.
- The correlation is time dependent for B which mix. This explains the large value of the B_d and B_s parameters.
  For B_d mesons the production time tagging is provided only by the opposite hemisphere charge requirement.
  For B_s mesons the tagging sensitivity is enlarged (larger B_s parameter) by the leading fragmentation kaon produced to compensate the strangess of the B_s meson (we

call it the B_s-Kpartner). From MC history, it has been seen in B_s events that $(47 \pm 4)\%$ of the selected kaons are B_s-Kpartner. For the remaining B_s event kaons, the charge asymmetry, created by the opposite jet-charge selection, is found to be $(26 \pm 4)\%$, consistent with the value of the A_u and B_d parameters.

The background (non b) is parametrized as follows:

$$P_{bkg}^{G,W} = f_{c\bar{c}} \frac{1 \pm A_{c\bar{c}}}{2} F_{c\bar{c}}(t_m^i) + f_{no.pr.} \frac{1 \pm A_{no.pr.}}{2} F_{no.pr.}(t_m^i) + f_{mis.id.} \frac{1 \pm A_{mis.id.}}{2} F_{mis.id.}(t_m^i)$$

where $f_{c\bar{c}}, f_{no.pr.}, f_{mis.id.}$ stand for the $c\bar{c}$, the non prompt and the misidentified hadron fractions. They satisfy the relation ship:

$$f_{c\bar{c}} + f_{no.pr.} + f_{mis.id.} = 1$$

The A coefficients are the asymmetry parameters, while the $F(t_m^i)$ functions correspond to the measured time distributions. Both are taken from Monte Carlo. The f and A values are given in table 6 with the Monte Carlo statistical error.

Table 6: Fractions and charge asymmetry parameters for the backgound

$f_{c\bar{c}}$	$f_{no.pr.}$	$A_{c\bar{c}}$	$A_{no.pr.}$	$A_{mis.id.}$
0.36 ± 0.02	0.42 ± 0.02	0.35 ± 0.07	-0.06 ± 0.06	$0.10 \pm .09$

3.3 Data distributions and chosen parameters

Fig.3 shows the proper time distribution of the 4436 data events. To check the sample composition and time resolution functions of the lepton-kaon-jet analysis, a fit to the time distribution is performed. The result is displayed on figure 3. The averaged B lifetime fitted on data is 1.55 ± 0.03 (statistical error only), which has to be compared to the world average value: 1.538 ± 0.022 [8]

For the data, the charge asymmetry:

$$Q_{asy}^i = \frac{N_i^G - N_i^W}{N_i^G + N_i^W}$$

is computed in each time bin i. Figure 4 displays Q_{asy} versus time. The time averaged charge asymmetry can be fitted on the data. As it is not possible to distinguish between A_u, A_d and A_s, A_d and A_s have been set to zero and only A_u is allowed to vary. Assuming $\Delta m_s = \infty$, and fixing the B_d parameter to the value of table 5 and $\Delta m_d = 0.49$ [11], the fit gives: $A_u = 0.22 \pm 0.05$ close to the Monte Carlo value of table 5. The χ^2 of the fit is 27.2 for 23 degrees of freedom.

The value of the B_d parameter can be checked on the data. It should not depend on the kaon requirement. Relaxing the kaon selection and defining the good and wrong sign classes of events from the opposite hemisphere charge alone, 50635 lepton-jet events are selected in the data. The previous analysis is repeated with the same fitting function. The sample composition, charge asymmetries of charm and background are taken from the ALEPH lepton-jet analysis [9]. The fit, where B_d is imposed to be equal to A_u, gives

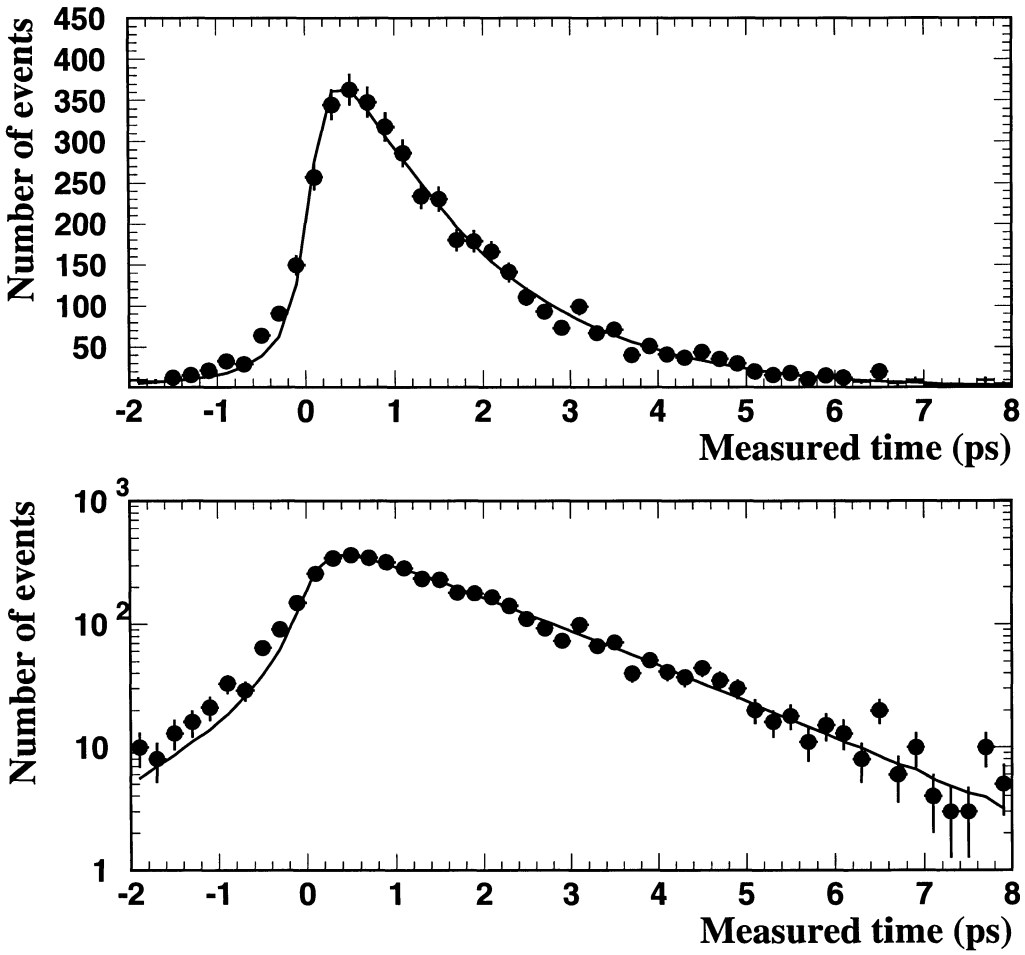

Figure 3: Lifetime fit

$B_d = 0.29 \pm 0.01$ and $\Delta m_d = 0.428 \pm 0.038(stat.)$ in fair agreement with the Monte Carlo value of table 5 and with [9].

The parameter values used in the present analysis are listed in table 7 together with the assumed systematic errors. The B lifetimes and Δm_d values and their errors are Winter conference averages [10, 11]. The value of the B_d and B_s fractions at the hadronization level (f_d and f_s) are measured values [12]. It should be noticed that they are the fractions of B_d and B_s after occurence of strong decays of the type $B_s^{**} \rightarrow B_d K$, which takes place before the weak mixing. These values are rescaled by the enrichment factor obtained by Monte Carlo. All other parameters are taken from Monte Carlo (see previous section) except the value of A_u which comes from the data. The parameters B_d and A_u are given the MC values with large systematics errors. In particular the consistency of the value taken for B_s (0.60 ± 0.12) with the value of $B_d = A_u = 0.29 \pm 0.01$, charge asymmetry due to the jet-charge tagging alone as measured on the data (see above), corresponds to a fraction of B_s-Kpartner contained in the selected kaon sample equal to $44 \pm 17\%$. This 40% relative uncertainty shows that the assumed systematic error on B_s is conservative.

All background parameters are also given large errors.

Table 7: Best estimated value and assumed systematic uncertainty of each parameter. $A_{background} = A_{no.pr.} = A_{mis.id.}$. The B_s and B_d fractions, f_s and f_d, are the product of the fractions at production by the selection enrichment factor. The values of these two quantities are given separately in two consecutive lines. The last column gives, for a value of $\Delta m_s = 4ps^{-1}$, the ratio systematic/statistical of the errors on either likelihood or amplitude (see section 4).

parameter	value	uncertainty	Syst / Stat %
$\tau_b(ps)$	1.54	± 0.02	0.4
τ_{Bd}/τ_b	1.013	± 0.047	0.8
$\tau_{Bs}(ps)$	1.56	± 0.12	5.6
Δm_d	0.49	± 0.03	0.4
f_d	$0.382 \times$	± 0.026	0.8
	1.01	± 0.05	
f_s	$0.122 \times$	± 0.032	35.0
	1.35	± 0.09	
f_b	0.921	± 0.042	13.1
f_{bc}	0.098	± 0.015	3.1
$f_{c\bar{c}}$	0.36	± 0.07	3.3
$f_{no.pr.}$	0.42	± 0.08	2.8
$A_{c\bar{c}}$	0.35	± 0.10	5.3
$A_{background}$	0.0	± 0.15	11.1
B_d	0.35	± 0.10	2.2
B_s	0.60	± 0.12	28.5
A_u	0.22	± 0.05	4.0
Total			49.2

4 Setting a Δm_s limit

Two methods have been used to search for the presence of the B_s^0 mixing: A likelihood difference method and a new amplitude fit method.

4.1 Likelihood difference method

4.1.1 Formalism

This is performed by calculating the difference in the log-likelihood values calculated for any Δm_s and for $\Delta m_s = \infty$. It is given by:

$$\Delta\mathcal{L}(\Delta m_s) = -\sum_{i=1} N_i \ln[\frac{D_i(\Delta m_s, \alpha^0)}{D_i(\infty, \alpha^0)}]$$

where N_i is the measured rate in time bin i and where D_i is the expected time distribution probability (given in section 3.1). In this expression and the following one's, the summation made on good and wrong sign distributions is omitted for simplicity. α^0 denotes

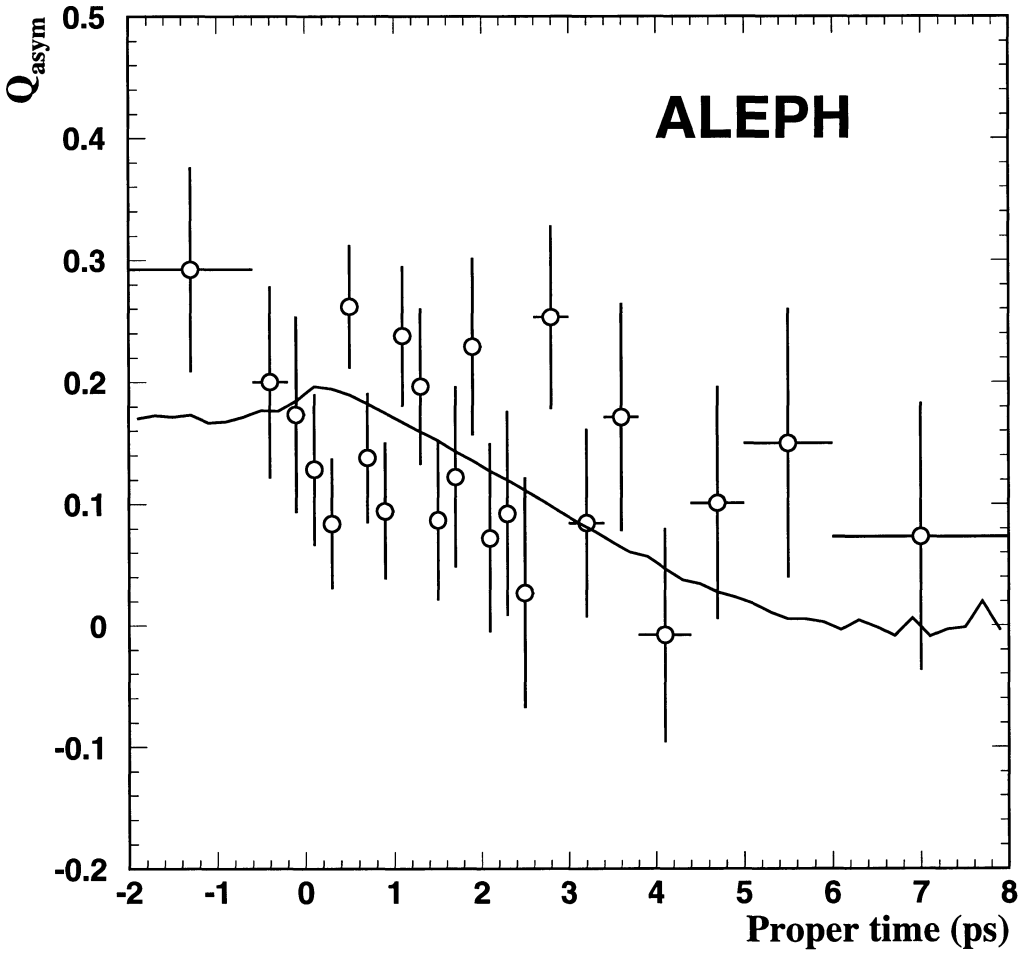

Figure 4: The lepton-kaon charge asymmetry as a function of reconstructed proper time, with the result of the fit shown superimposed, assuming $\Delta m_s = \infty$.

the set of all parameters (fractions f, charge correlation coefficients A and B, Δm_d and lifetimes) on which depend the charge correlation. The best estimate of these parameters is used, the same for data and simulation, to compute $\Delta\mathcal{L}(m_s)$.

A fast toy Monte Carlo makes use of the parametrized expected distributions D_i to generate simulated distributions (with a number of events equal to the present data statistics) for any value of Δm_s and of the parameters α^0. They are used to determine what is the expected distribution of $\Delta\mathcal{L}$ as a function of Δm_s. The originality of the present work is the use of an analytical determination of this distribution:

The average log-likelihood difference is given by:

$$\langle \Delta\mathcal{L}(\Delta m_s) \rangle = -N \sum_{i=1} D_i(\Delta m_s, \alpha^0) \ln[\frac{D_i(\Delta m_s, \alpha^0)}{D_i(\infty, \alpha^0)}]$$

where N_i has been replaced by its average $N \cdot D_i$ to get the expected average time distribution. Similarly, for samples generated with $\Delta m_s = \infty$, the average log-likelihood

difference is given by:

$$\Delta\mathcal{L}^{\infty}(\Delta m_s) = -N\sum_{i=1} D_i(\infty,\alpha^0)\ln[\frac{D_i(\Delta m_s,\alpha^0)}{D_i(\infty,\alpha^0)}]$$

The following relation can be demonstated:

$$\Delta\mathcal{L}^{\infty}(\Delta m_s) = -\langle\Delta\mathcal{L}(\Delta m_s)\rangle$$

The statistical rms of $\Delta\mathcal{L}(\Delta m_s)$ is given by:

$$\sigma^{stat}[\Delta\mathcal{L}(\Delta m_s)] = \sqrt{N\sum_{i=1} D_i(\Delta m_s,\alpha^0)[\ln\frac{D_i(\Delta m_s,\alpha^0)}{D_i(\infty,\alpha^0)}]^2}$$

The following relation can be demonstated:

$$\sigma^{stat}[\Delta\mathcal{L}(\Delta m_s)] = \sqrt{2\cdot\Delta\mathcal{L}^{\infty}(\Delta m_s)}$$

The systematic rms of $\Delta\mathcal{L}(m_s)$ originates from the fact that the best estimate of the parameters α^0 which is used, may not correspond to the true value of the data. For instance the true value of each parameter, say α_l, may differ from the used estimate α_l^0, by a systematic spread, assumed gaussian of rms σ_{α_l}. The systematic rms of $\Delta\mathcal{L}(\Delta m_s)$ which results from this parameter uncertainty is:

$$\sigma_l^{syst}[\Delta\mathcal{L}(\Delta m_s)] = N\sigma_{\alpha_l}\sum_{i=1}\frac{\partial D_i(\Delta m_s,\alpha^0)}{\partial\alpha_l}\ln[\frac{D_i(\Delta m_s,\alpha^0)}{D_i(\infty,\alpha^0)}]$$

The total systematic error is given by adding quadratically the effects of all parameters α_l:

$$\sigma^{syst}[\Delta\mathcal{L}(\Delta m_s)] = \sqrt{\sum_l[\sigma_l^{syst}\Delta\mathcal{L}(\Delta m_s)]^2}$$

The total error is:

$$\sigma^{tot} = \sqrt{\sigma_{stat}^2+\sigma_{syst}^2}$$

and, the $\Delta\mathcal{L}$ distribution being gaussian, a fact which has been demonstrated with the toy Monte Carlo, the 95 % confidence level is obtained as:

$$\Delta\mathcal{L}^{95} = \langle\Delta\mathcal{L}(\Delta m_s)\rangle + 1.645\times\sigma^{tot}[\Delta\mathcal{L}(\Delta m_s)]$$

It has been checked extensively that these analytical computations give the same results than the toy Monte Carlo. They allow a faster study of the errors, more Δm_s points, and the separation of the systematic contributions of all parameters. They have been used to obtain the results given below.

4.1.2 Results

On figure 5 is displayed, as a function of Δm_s, the log-likelihood difference for the data (full line) and the 95 % confidence level limit obtained with the statistical errors only (open circles). The limit is $\Delta m_s = 4.35ps^{-1}$ with statistical errors only.

The systematic errors of the parameters are given in section 3.3, table 7. From them the systematic rms of $\Delta\mathcal{L}$, $\sigma^{syst}[\Delta\mathcal{L}]$, can be computed. The ratio of systematic rms to the statistical one, $\sigma^{stat}[\Delta\mathcal{L}]$, is shown in figure 6 versus Δm_s and individual values are

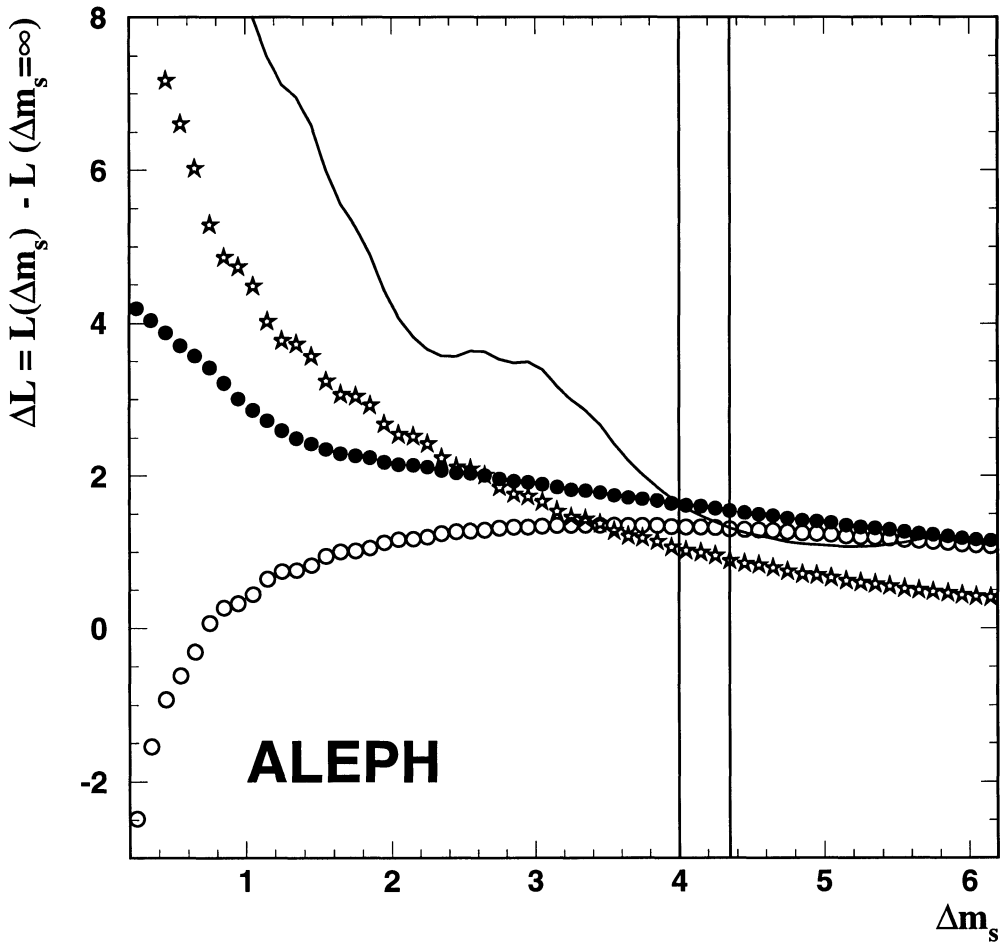

Figure 5: Log likelihood differences versus Δm_s, full curve: data, open circles: $\Delta\mathcal{L}$ 95% C.L. limit (stat. only), dots: $\Delta\mathcal{L}$ 95 % C.L. limit (stat. + syst.) , stars: $\Delta\mathcal{L}$ average on $\Delta m_s = \infty$ samples

given in table 7 for $\Delta m_s = 4ps^{-1}$. At low values of Δm_s, the main systematic is due to the correlated parameters A_u and B_d (charge correlation for non B_s) but above Δm_s =2 ps^{-1} the contribution of f_s and B_s (oscillation amplitude for B_s mesons) dominate.

On figure 5 is also displayed the 95 % confidence level limit obtained when adding statistical and systematic errors (black dots). When the systematics are included the limit become $\Delta m_s = 4.0ps^{-1}$.

The curve giving the variation of the log-likelihood difference $\Delta\mathcal{L}^{\infty}(\Delta m_s)$, averaged for samples generated with $\Delta m_s = \infty$, is displayed on figure 5 (stars). It crosses the 95%C.L. curves at values of Δm_s which represent the averaged potential limits allowed by this analysis. They are $\Delta m_s = 3.45ps^{-1}$ for statistics only and $2.65ps^{-1}$ when including the systematics. The data result is therefore somewhat lucky. To see how lucky it is, 400 toy Monte Carlo samples have been generated with $\Delta m_s = \infty$. In 32% of these samples, the likelihood curve crosses the 95% confidence level limit (with syst.) at a Δm_s value

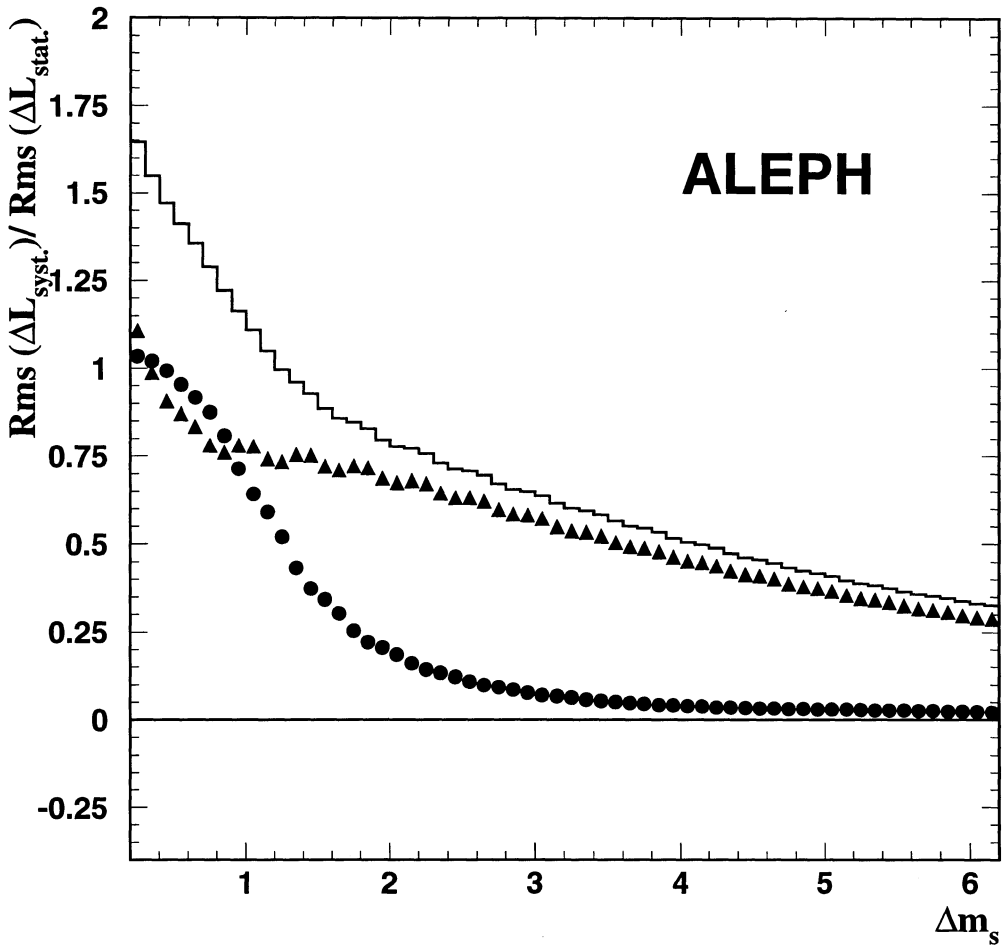

Figure 6: Ratio of the systematic rms of $\Delta\mathcal{L}$ to the statical rms versus Δm_s. Histogram: total systematics, dots: contribution of B_d,A_u, triangles: contributions of f_s and B_s.

greater than the data limit, $4.0ps^{-1}$

4.2 Amplitude fit method

To check the validity of the previous limit and to make possible the combination of limits derived by various analyses or experiments, a new method, the amplitude fit, has been performed.

4.2.1 Formalism and results

The log-likelihood formula given at the beginning of section 4.1 is modified as follows: the B_s^0 oscillation amplitude (parameter B_s) is rescaled by $\mathcal{A}$, the "fitted amplitude". For any given Δm_s, the value of $\mathcal{A}$ which minimizes the log-likelihood, is determined. This is done first with the optimal values of the parameters given in table 7. Let $\mathcal{A}_0$ and $\sigma[\mathcal{A}_0]$

be the fit result and the statistical error. Figure 7 displays the variation of $\mathcal{A}_0$ with Δm_s (points with statistical errors). The dependence of the log-likelihood with $\mathcal{A}$ is parabolic: the errors made on the amplitude are gaussian. This allows to compute the 95% CL limit on the amplitude to be, $\mathcal{A}_0 + 1.645 \cdot \sigma[\mathcal{A}_0]$, drawn on the figure. The fit result and the limit have to be compared to either 0 (no mixing is observed) or 1 (mixing is observed for this value of Δm_s). This method is more pedagogical to display the mixing visibility and will allow an easy combination of mixing results originating from various analyses and experiments.

In order to determine what systematic errors on the amplitude are induced by the parameter uncertainties, toy Monte Carlo samples including the expected B_s^0 mixing at the value of Δm_s under study are generated. It has been verified that the averaged value of the amplitude $\mathcal{A}$, fitted at this value of Δm_s, is equal to 1. These parameters are changed, one at a time, by a quantity equal to the assumed uncertainty (see table 7). We then fit the new value of the amplitude $\mathcal{A}$ with the error $\sigma[\mathcal{A}]$. A technical difficulty arises, because the error does not stay constant when varying some of the parameters (f_s,B_s,..). The systematic error on the amplitude cannot be simply derived from the difference $\mathcal{A} - \mathcal{A}_0$. To overcome this difficulty, the amplitude has been redefined. The total B_s^0 oscillating amplitude is the product of $\mathcal{A}$ by the following parameter's product:

$$\mathcal{F} = f_b \cdot (1 - 2 \cdot f_{bc}) \cdot f_s \cdot B_s$$

The value of $\mathcal{A} \cdot \mathcal{F}$ is imposed by the charge-asymmetry distribution and depends very little on $\mathcal{F}$. Let $\mathcal{F}_0$ be the value of $\mathcal{F}$ for the central value of the parameters. The amplitude is redefined as:

$$\mathcal{B} = \frac{\mathcal{F}}{\mathcal{F}_0} \cdot (\mathcal{A} - 1) + 1$$

With this new definition the central value $\mathcal{B}_0$ and the statistical error $\sigma[\mathcal{B}]$ of the amplitude are left unchanged. This error depends very little on the parameters. This allows to compute the systematic error on the amplitude due to each parameter as $\mathcal{B} - \mathcal{B}_0$, the difference of the fitted amplitudes when this parameter is changed by a quantity equal to the assumed uncertainty. This is done for all the parameters of table 7. The systematic errors are summed quadratically to get the total systematic error $\sigma^{syst}[\mathcal{A}]$ which, at the end, is added to the statistical error:

$$\sigma^{tot} = \sqrt{\sigma^2_{stat} + \sigma^2_{syst}}$$

The 95% CL limit on the amplitude, including all systematic errors, is $\mathcal{B}_0 + 1.645 \cdot \sigma^{tot}$. It is drawn on figure 7. The 95% CL limits derived on Δm_s are obtained from the crossing of the 95% curves with unity. They are almost identical to the limits obtained using the log-likelihood difference method of the previous section.

4.2.2 Averaging results

The amplitude fit method offers a unique advantage of making straightforward to combine Δm_s limits originating from various analyses and/or experiments, if they have presented their results according to it. It transforms each search for B_s^0 mixing in the measurement of the same physical quantity, the observed mixing amplitude, which should be the same for all experiments, equal to 0 or 1. Combining results means averaging these measurements in the usual way.

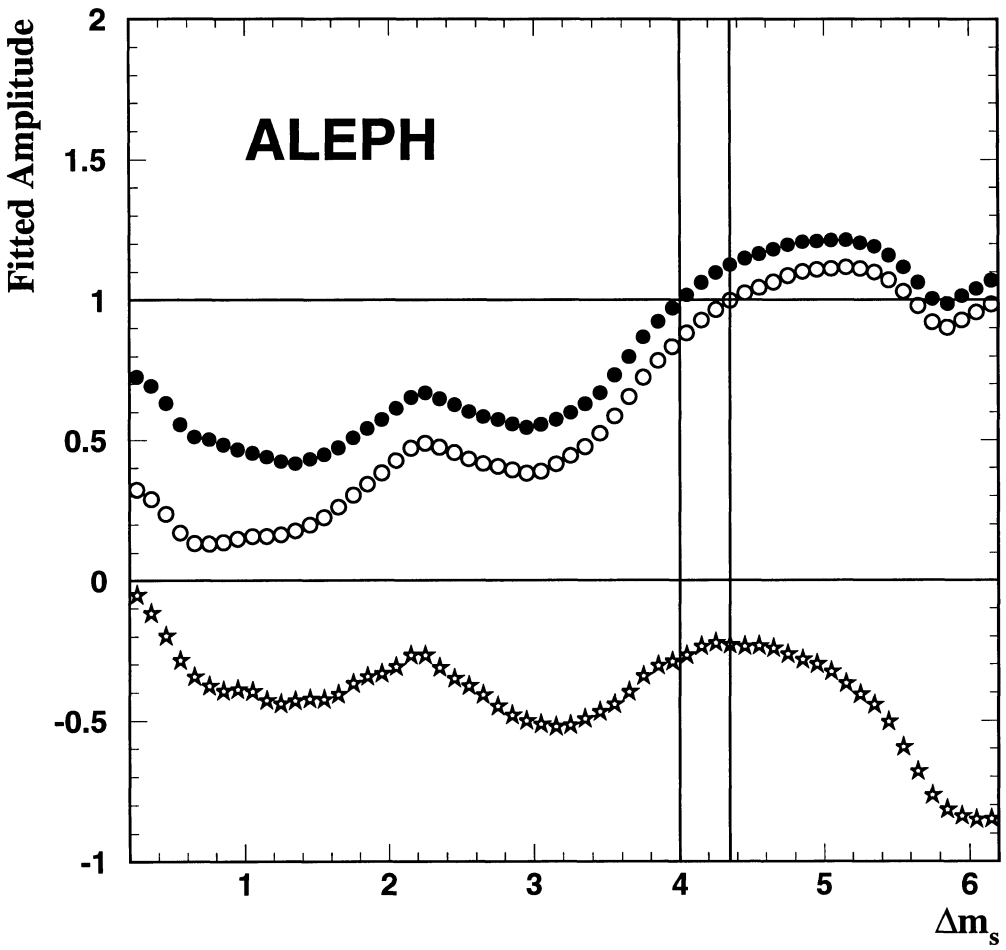

Figure 7: Fitted amplitude $\mathcal{B}$ versus Δm_s stars: data, open circles: $\mathcal{B}$ 95 dots: $\mathcal{B}$ 95 % C.L. limit (stat. + syst.). A value of 0 corresponds $\Delta m_s = \infty$ and 1 to full mixing at this value of Δm_s

Let i be the index labelling a given analysis. The averaging requires that each experiment has tabulated, for given values of Δm_s, four numbers: $\mathcal{B}_i$ the measured fitted amplitude, σ_i^{stat} the statistical error, and the systematic errors σ_i^{syst} which should have been separated into two terms, σ_i^{syst-1} the quadratic sum of the systematic errors which are specific of this analysis, and σ_i^{syst-2} the part(s) of the systematic errors which are in common in the analyses which are to combine. The reason to separate these two kind of systematics is that the later type is to be summed coherently when combining results. The first type of systematic error σ_i^{syst-1} has to be added quadratically with the statistical error σ_i^{stat} to get the total error, σ_i^{tot}, of each experiment, independent of the others. A weight G_i, for each analysis, is defined as the inverse of the square of this error. Let define the sum of the weights, $G = \sum_{i=1} G_i$. The averaged amplitude is:

$$\langle \mathcal{B} \rangle = \frac{\sum_{i=1} G_i \cdot \mathcal{B}_i}{G}$$

and the global error on this amplitude is:

$$\langle \sigma^{tot} \rangle = 1/\sqrt{G}$$

while the global common systematic error is obtained from the average:

$$\langle \sigma^{syst} \rangle = \frac{\sum_{i=1} G_i \cdot \sigma_i^{syst-2}}{G}$$

Having combined amplitudes and errors, the 95% CL limit can be computed as a function of Δm_s and the common limit derived.

In table 8 are presented the results of our analysis tabulated in the Δm_s range from 3 ps^{-1} up to 10 ps^{-1} by step of 0.5 ps^{-1}. The following values are given: $\mathcal{B}$ the measured fitted amplitude, σ^{stat} the statistical error, and the systematic error of type 1, σ_i^{syst-1}, the quadratic sum of the systematic errors which are specific of our analysis. There may be several parameters creating the common systematics. In the case of our analysis, only two of them are significant. The main one is due to the f_s parameter and the other, which is marginally significant, is due to f_b. All other possibility of common systematic are negligible. The values of these two common systematic errors on the amplitude are given in table 8.

Table 8: Amplitude fit method results as a function of Δm_s. The fitted amplitude is given with the statistical error, the systematic error specific to this analysis and two systematic errors which may be in common with other analysis (see text).

Δm_s	$\mathcal{B}$	$\sigma^{stat}[\mathcal{B}]$	σ^{syst-1}	σ^{syst-2}, f_s	σ^{syst-2}, f_b
3.00	-0.503	0.535	0.221	0.240	0.084
3.50	-0.457	0.615	0.222	0.244	0.090
4.00	-0.280	0.692	0.222	0.243	0.092
4.50	-0.235	0.773	0.224	0.249	0.093
5.00	-0.305	0.858	0.221	0.245	0.089
5.50	-0.549	0.973	0.228	0.245	0.096
6.00	-0.846	1.087	0.229	0.247	0.106
6.50	-0.669	1.205	0.234	0.248	0.116
7.00	-0.705	1.332	0.234	0.251	0.120
7.50	-1.550	1.495	0.236	0.249	0.126
8.00	-2.291	1.663	0.244	0.250	0.132
8.50	-2.090	1.837	0.245	0.247	0.133
9.00	-0.762	2.021	0.255	0.242	0.134
9.50	1.387	2.240	0.263	0.240	0.137
10.00	3.504	2.470	0.269	0.235	0.140

4.3 Identity of the two methods

We are going to show that, in fact, the two previous methods coincide. They are mathematically related by a given relation: For a given value of Δm_s and a log-likelihood value $\Delta\mathcal{L}$ we define the following linear transformation

$$\mathcal{A} = \frac{\Delta\mathcal{L}^{\infty}(\Delta m_s) - \Delta\mathcal{L}}{\Delta\mathcal{L}^{\infty}(\Delta m_s) - \langle \Delta\mathcal{L}(\Delta m_s) \rangle}$$

where the coefficients, $\Delta\mathcal{L}^{\infty}(\Delta m_s)$ and $\langle\Delta\mathcal{L}(\Delta m_s)\rangle$, have been defined in section 4.1. By definition, this relation transforms, the likelihood curves of figure 5 associated to a given Δm_s (resp. $\Delta m_s = \infty$) to the corresponding value of the amplitude $\mathcal{A} = 1$ (resp. $\mathcal{A} = 0$). But this relation also transforms, with a relative accuracy better than 1%, the likelihood data curve of figure 5 to the corresponding amplitude data curve on figure 7. To the same level of accuracy the likelihood errors (both statistical and systematic) are transformed in the corresponding errors in amplitude. The 95 % CL curves of figure 5 (log-likelihood difference) are therefore transformed to the curves of figure 7. The crossing of curves, therefore the Δm_s limits, are the same.

From the analytical formulae on likelihood given in section 4.1.1 and from the relation between likelihood and amplitude given above, a simple relation can be given which relates likelihood and amplitude rms.

$$\sigma^{stat}[\mathcal{A}] = 1/\sigma^{stat}[\Delta\mathcal{L}(\Delta m_s)]$$

5 Conclusion

From the 1991 to 1994 data collected by the ALEPH detector a new limit on the B_s^0 mixing has been presented. 4436 lepton-fragmentation kaon-jet correlations have been measured. For B_s^0 events, the mistag rate and the B_s selection enrichment have been determined by Monte-Carlo to be respectively $(19.0 \pm 1.3)\%$ and $(35 \pm 4)\%$. A fast new maximum likelihood technique has been derived where the log-likelihood is calculated from the time distribution probabilities using analytic formulae. A new method, the amplitude fit method, is also described. It gives identical results to the log-likelihood difference method, but is more pedagogical and will allow an easy combination of mixing results originating from various analyses and experiments. Taking the fraction f_s of the b quarks that form B_s to be $(12 \pm 3)\%$, a lower limit has been set on the B_s oscillation parameter: $\Delta m_s > 4.0 ps^{-1}$ at 95% confidence level.

References

[1] A. Ali and D. London, Z. Phys. C65 (1995) 431.

[2] D. Buskulic et al. (ALEPH Collaboration), Phys. Lett. B322 (1994) 441. Contributed paper to Int. Conf. on HEP,Glasgow, Scotland (1994), GLS 0584.

[3] R. Akers et al. (OPAL Collaboration), CERN-PPE/95-012.

[4] D. Buskulic et al. (ALEPH Collaboration), CERN-PPE/95-084.

[5] D. Decamp et al. (ALEPH Collaboration), Nucl. Inst Methods A294 (1990) 121.

[6] D. Buskulic et al. (ALEPH Collaboration), CERN-PPE/94-170.

[7] D. Buskulic et al. (ALEPH Collaboration), Phys. Lett. B322 (1994) 275.

[8] V. Sharma, Measurement of Beauty Lifetime, 6^{th} International Symposium on Heavy Flavor Physics, Pisa, Italy, June 1995.

[9] S. Emery and W. Kozanecki, measurement of the B_d^0 - $\bar{B}_d^0$ oscillation frequency using a Jet-Charge Method, ALEPH note Physics 95-030.

[10] M. Jimack, b-physics at LEP, Proc. of XXX Rencontres de Moriond (Les Arcs, Savoie, France), March 1995.

[11] S. Emery, Time Dependent $B_0\bar{B}_0$ oscillation at LEP, Proc. of Rencontres de Moriond QCD (Les Arcs, Savoie, France), March 1995.

[12] D. Buskulic et al, (ALEPH Collab.). A measurement of $|V_{cb}|$ from $\bar{B}^0 \to D^{*+}l^-\nu_l$, in preparation.