Received: *20 July, 1999*, Accepted: *6 August, 1999*

The B_{s} oscillation amplitude analysis

Duccio Abbaneo and Gaëlle Boix

CERN, CH-1211,
Geneva 23, Switzerland
Email: duccio.abbaneo@cern.ch, gaelle.boix@cern.ch

Abstract: The properties of the amplitude method for B_{s} oscillation analyses are studied in detail. The world combination of measured amplitudes is converted into a likelihood profile as a function of oscillation frequency. A procedure is proposed to estimate the probability that the minimum observed is due to a statistical fluctuation. This method, applied to the data available at the time of 1999 Winter Conferences, gives $1 - \mathrm{C.L.} \approx 0.03$.

Keywords: e+ e- Experiments.

JHEP08(1999)004

Contents

1. Introduction

The first direct search for $B_{\rm s} - \overline{B_{\rm s}}$ oscillations was published by ALEPH in 1994 [1]. Since then, many sophisticated analyses have been developed by the LEP experiments, SLD and CDF [2]. None of these analyses has yet been able to measure the oscillation frequency, but can exclude a range for the mass difference between the two $B_{\rm s}$ mass eigenstates, $\Delta m_{\rm s}$, which drives the oscillation.

In order to combine the information provided by the different analyses in the absence of a measurement of $\Delta m_{\rm s}$, a new technique [3], known as the *amplitude method*, was proposed. The fit to the reconstructed proper time distribution of events tagged as mixed or unmixed is performed with a fixed frequency ω of the oscillating term, while its amplitude $\mathcal{A}$ is left as the free parameter. A scan in ω is performed and at each value the amplitude is measured. Averaging values from different analyses is straightforward. The expected value of the amplitude is unity when $\omega = \Delta m_{\rm s}$ (throughout this paper ω stands for the frequency folded in the fitting function, Δm or $\Delta m_{\rm s}$ indicate the frequency of the oscillations in the sample analysed).

The range of ω for which the amplitude is found to be compatible with zero and incompatible with unity can be excluded. An analysis (or a combination of analyses) has sensitivity in a given range of ω if the expected error on the measured amplitudes is small enough compared to unity, so that the two values $\mathcal{A} = 0.$ and $\mathcal{A} = 1.$ can be distinguished. In order to quote a sensitivity limit, it is normally chosen to determine the value of ω for which a measured value $\mathcal{A} = 0.$ implies that $\mathcal{A} = 1.$ is excluded at 95% C.L. This happens when $1.645 \times \sigma_{\mathcal{A}} = 1$.

The worldwide combination, at the time of the 1999 Winter Conferences, of amplitude analyses shows a deviation from $\mathcal{A} = 0$ at $\omega \approx 15\,\mathrm{ps}^{-1}$, close to the sensitivity limit of $14.3\,\mathrm{ps}^{-1}$ [4], which could suggest the presence of an oscillating signal. The combination includes many preliminary analyses, and some new analyses are still expected from SLD and the LEP experiments. The significance of the structure observed might therefore be reduced or further enhanced in the near future as ongoing analyses are completed and published; in any case a procedure to quantify the probability that the observed structure corresponds to a $B_{\rm s}$ oscillation signal is needed. A method is proposed here, based on the generation of *toy experiments* designed to be equivalent to the world combination of $\Delta m_{\rm s}$ analyses.

The paper is organized as follows:

In section 2 the combined amplitude at different values of ω is presented. The combined likelihood profile as a function of ω is extracted from the amplitude spectrum.

In section 3 the properties of the amplitude method are investigated. Analytical expressions for the expected shape of the measured amplitude and its error are derived. The small and large frequency limits are discussed, proposing an approximate interpretation in terms of Fourier transformations. The probability of observing statistical fluctuations which would fake a signal in a sample with frequency far beyond the sensitivity is also discussed.

In section 4 the structure and the features of the toy experiment generator used throughout the paper are described. A procedure to tune the parameters of the simulation in order to reproduce the observed errors is given.

In section 5 a procedure to extract a confidence level value from the likelihood function is presented and discussed. The uncertainty arising from the lack of a detailed simulation is investigated.

2. The experimental results

The combined amplitude measurements obtained from published and preliminary analyses available at the time of the 1999 Winter Conferences [4] is presented in figure 1.

JHEP08(1999)004

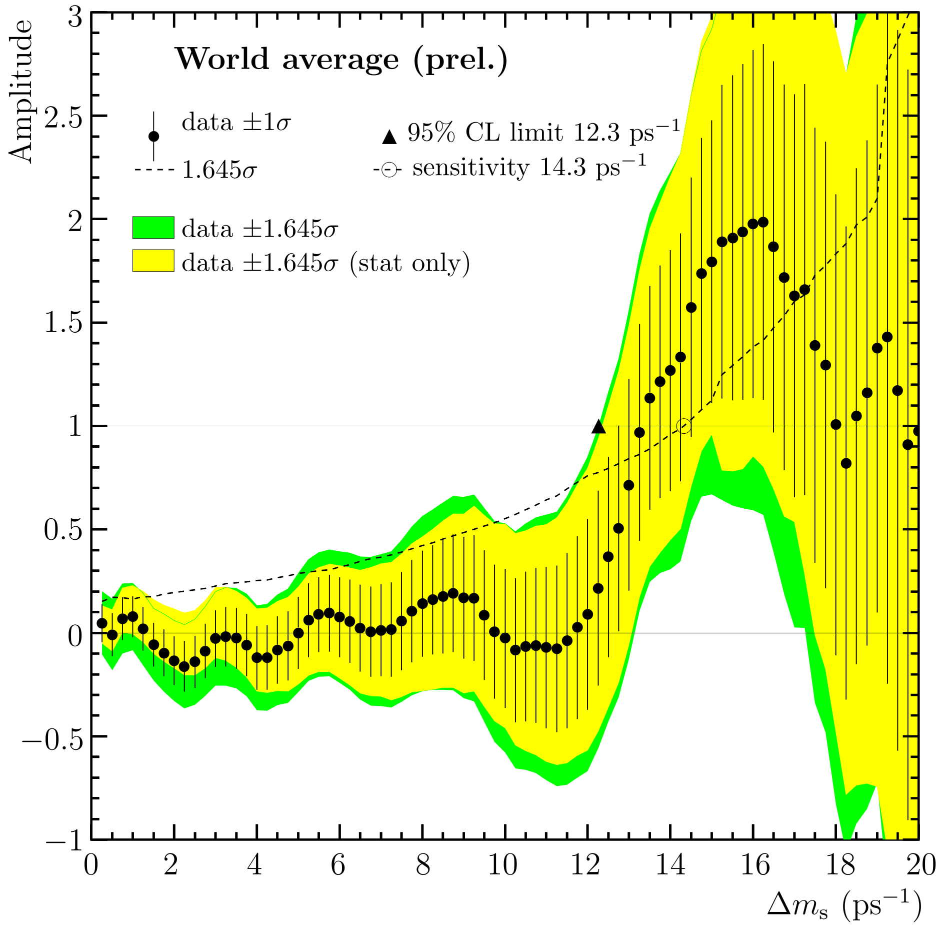

Figure 1: Current combined amplitude measurements as a function of ω, from the B Oscillation Working Group.

The observed limit is significantly smaller than the expected limit (i.e., that which would be obtained if Δm_s were infinitely large), due to positive amplitude values measured in the region close to the sensitivity limit.

As mentioned in the introduction, the amplitude measurements are obtained by maximizing the likelihood L of the proper time distributions of mixed and unmixed events with the amplitude of the oscillating term $\mathcal{A}$ as the free parameter, and ω fixed at a chosen value. Denote $\mathcal{L} = -\log L$, its expansion at second order around the minimum of $\mathcal{L}$, $\mathcal{L}_\omega(\bar{\mathcal{A}})$ can be approximated by

$$\mathcal{L}_\omega(\mathcal{A}) \simeq \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{\mathcal{A} - \bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega, \tag{2.1}$$

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

From eq. (2.1) it follows that, again for each value of ω

$$\mathcal{L}_\omega(\mathcal{A}=1) \simeq \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{1-\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega . \tag{2.2}$$

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

$$\mathcal{L}_{\omega\to\infty}(\text{any } \mathcal{A}) = \mathcal{L}_{\text{any } \omega}(\mathcal{A}=0) \ (\equiv \mathcal{L}_\infty) , \tag{2.3}$$

and therefore from eq. (2.1),

$$\mathcal{L}_\infty = \mathcal{L}_\omega(\bar{\mathcal{A}}) + \frac{1}{2}\left(\frac{\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2_\omega . \tag{2.4}$$

If eq. (2.4) is subtracted from eq. (2.2), the following formula is obtained

$$\Delta\mathcal{L}(\omega) \equiv \mathcal{L}_\omega(\mathcal{A}=1) - \mathcal{L}_\infty \simeq \left[\frac{1}{2}\left(\frac{1-\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2 - \frac{1}{2}\left(\frac{\bar{\mathcal{A}}}{\sigma_{\mathcal{A}}}\right)^2\right]_\omega , \tag{2.5}$$

which allows the value of $\Delta\mathcal{L}$ to be calculated, for each ω, from the fitted amplitude and its uncertainty. This formula was already given in ref. [3].

In the derivation which follows, the total uncertainties on the measured amplitudes are used, which is in principle not rigorous, since eq. (2.1) is valid only for the statistical part. This fact is not considered to be a problem, since most of the uncertainties quoted as systematic errors in the analyses combined are of statistical nature. Typically, for inclusive analyses, the largest sources of uncertainty on the amplitude measurement are the fraction f_{B_s} of B_s produced in the b quark hadronization, the b hadron lifetimes and their mean energy. All these parameters are measured, and their experimental errors are propagated as systematic uncertainties on $\mathcal{A}$. In some analyses the parameters considered as sources of systematic errors were even added as free parameters in the fit, with gaussian constraints coming from external measurements. The part of the error identified as systematic was thus found out *a posteriori* by running the minimization with those parameters fixed. In addition, in the region of interest ($\omega > 10\,\text{ps}^{-1}$) the statistical errors are dominant. The approach adopted should therefore be adequate.

The likelihood difference $\Delta\mathcal{L}(\omega)$ obtained for the data is shown in figure 2. A good parametrization for the shape of $\Delta\mathcal{L}$ is obtained with a function $f(\omega) \propto 1/\omega^\alpha$ with $\alpha = 1.64$, plus some gaussian functions to describe the deviations. A parabolic fit of the three lowest points of the plot gives a minimum for $\omega = 14.8\,\text{ps}^{-1}$, with a value $\Delta\mathcal{L}_{\text{min}} = -2.9$. The $\Delta\mathcal{L}_{\text{min}}+1/2$ and $\Delta\mathcal{L}_{\text{min}}+2$ levels are crossed at the values $(14.3-15.3)\,\text{ps}^{-1}$ and $(13.0-17.5)\,\text{ps}^{-1}$, respectively, giving

$$\begin{aligned} \omega &= 14.8 \pm 0.5 \ \text{ps}^{-1} \quad (\Delta\mathcal{L}_{\text{min}} + 1/2 \text{ interval}) , \\ \omega &= 14.8 \ ^{+2.7}_{-1.8}\,\text{ps}^{-1} \quad (\Delta\mathcal{L}_{\text{min}} + 2 \text{ interval}) . \end{aligned} \tag{2.6}$$

JHEP08(1999)004

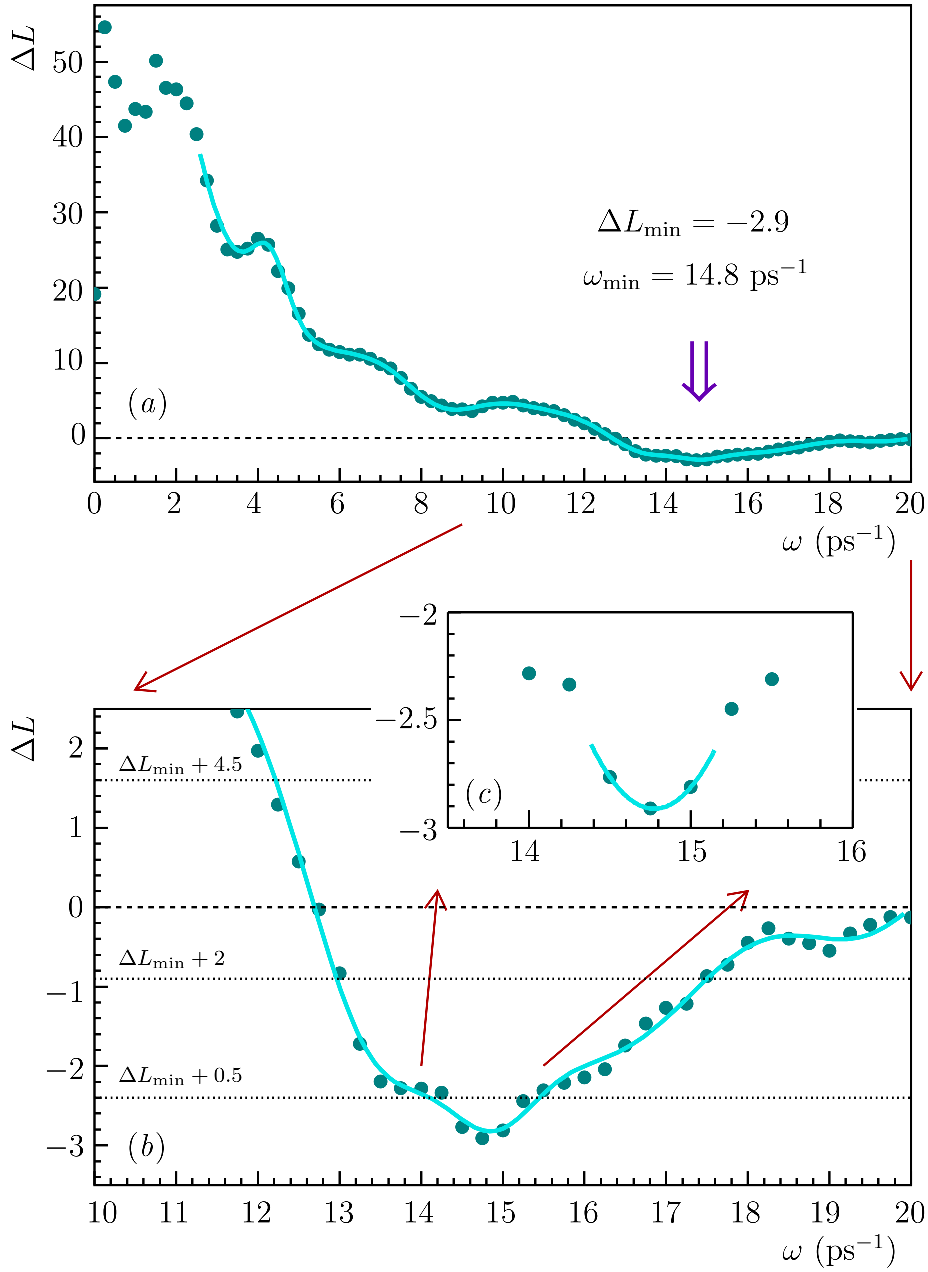

Figure 2: Likelihood as a function of ω derived from the combined amplitude measurements. A minimum is observed for $\omega = 14.8\,\mathrm{ps}^{-1}$. The parametrization described in the text is shown in (a) and (b); the parabolic fit to the three lowest points in (c).

These intervals would give the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties if the likelihood profile were parabolic in a range wide enough around the minimum. The $\Delta\mathcal{L}_{\rm min} + 9/2$ level is crossed on the lower side only, at $\omega = 12.1\,\mathrm{ps}^{-1}$.

As discussed in the following sections, the significance of this minimum cannot be extracted in an analytical way, but needs to be determined with toy experiments.

3. The amplitude analysis

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

$$\mathcal{P}^0_{\rm u,m}(t_0) = \Gamma \exp\left(-\Gamma t_0\right) \frac{1 \pm \cos \Delta m\, t_0}{2} \equiv \frac{E^0(t_0) \pm f^0_{\Delta m}(t_0)}{2}\,, \tag{3.1}$$

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

The reconstructed proper time distributions can then be written as

$$\mathcal{P}_{\rm u,m}(t) = \int_0^\infty dt_0\, \frac{E^0(t_0) \pm f^0_{\Delta m}(t_0)}{2}\, \mathcal{R}(t_0, t) \equiv \frac{E(t) \pm f_{\Delta m}(t)}{2}\,. \tag{3.2}$$

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

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

$$\begin{aligned}\mathcal{R}(t_0, t) &= \int_{-\infty}^{\infty} dp \frac{1}{\sqrt{2\pi}\,\sigma_p} \exp\left(-\frac{(p - p_0)^2}{2\,\sigma_p^2}\right) \frac{1}{\sqrt{2\pi}\,\sigma_l} \exp\left(-\frac{(pct - p_0 c t_0)^2}{2\,(m\sigma_l)^2}\right) \frac{pc}{m} \approx \\ &\approx \frac{1}{\sqrt{2\pi\,[\delta_l^2 + (\delta_p t)^2]}} \exp\left(-\frac{(t - t_0)^2}{2\,[\delta_l^2 + (\delta_p t)^2]}\right),\end{aligned} \tag{3.3}$$

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

A set of parameters is chosen here for the purpose of illustration. Resolution values of $\delta_p = 0.15$ and $\delta_l = 0.14\,\mathrm{ps}$ are used; the latter one would correspond to a monochromatic sample of $B_{\rm s}$ with $p_0 = 32\,\mathrm{GeV}/c$ and $\sigma_l = 250\,\mu\mathrm{m}$. In a real analysis the normalization of the non-oscillating component is the total number N of b decays (differences in lifetime are neglected), while the oscillation term is multiplied by $N f_{B_{\rm s}}(1-2\eta)$, $f_{B_{\rm s}}$ being the fractions of $B_{\rm s}$ in the sample and η the global mistag rate. For an inclusive analysis $f_{B_{\rm s}}(1-2\eta)$ is typically about 0.05. The curves obtained with these parameters, normalization factors omitted, are shown in figure 3. As the

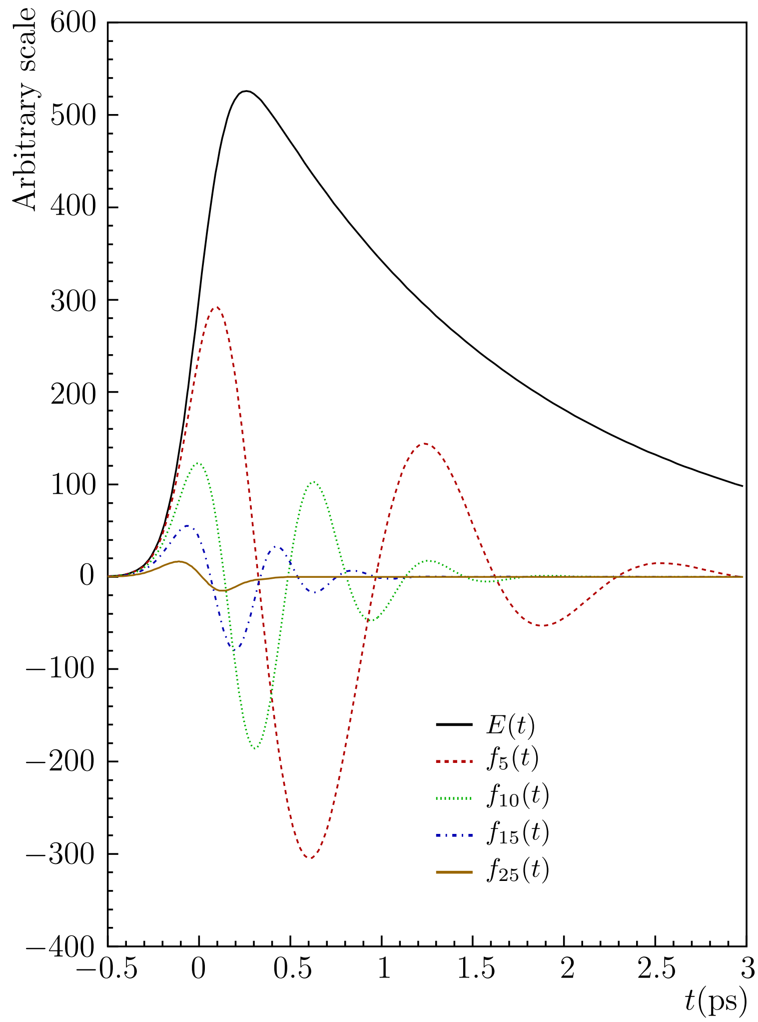

Figure 3: Reconstructed proper time distributions for the non–oscillating component, $E(t)$, and the oscillating component, $f_{\Delta m}(t)$, at different values of Δm. Resolutions of $\delta_p = 0.15$ and $\delta_l = 0.14\,\mathrm{ps}$ are assumed.

frequency increases, the oscillation amplitude is damped because of the resolution. For very large frequencies only the first period can be resolved.

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

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

$$-\log L = \frac{1}{2}\int_{-\infty}^{\infty} dt\Big[E(t) + f_{\Delta m}(t)\Big]\,\log\Big[E(t) + \mathcal{A} f_\omega(t)\Big] + \\ + \Big[E(t) - f_{\Delta m}(t)\Big]\,\log\Big[E(t) - \mathcal{A} f_\omega(t)\Big] + \text{Constant}\,, \qquad (3.4)$$

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

$$\int_{-\infty}^{\infty} dt \, \frac{f_\omega(t)\, f_{\Delta m}(t) - \mathcal{A} f_\omega^2(t)}{E(t)\,(1 - \mathcal{A}^2 f_\omega^2(t)/E^2(t))} = 0\,, \tag{3.5}$$

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

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

$$\chi^2 = \int_{-\infty}^{\infty} dt \, \frac{[f_{\Delta m}(t) - \mathcal{A} f_\omega(t)]^2}{E(t)}\,, \tag{3.6}$$

the minimization of which gives

$$\int_{-\infty}^{\infty} dt \, \frac{f_\omega(t)\, f_{\Delta m}(t) - \mathcal{A} f_\omega^2(t)}{E(t)} = 0\,. \tag{3.7}$$

Eqs. (3.5) and (3.7) both give $\mathcal{A} = 1$ for $\omega = \Delta m$. For $\omega \neq \Delta m$ they are equivalent if $\mathcal{A} f_\omega(t)$ is negligible compared to $E(t)$.

The expression of $\mathcal{A}_{\Delta m}(\omega)$ can be derived from eq. (3.7) as

$$\mathcal{A}_{\Delta m}(\omega) = \frac{\displaystyle\int_{-\infty}^{\infty} dt\, f_\omega(t)\, f_{\Delta m}(t)/E(t)}{\displaystyle\int_{-\infty}^{\infty} dt\, f_\omega^2(t)/E(t)}\,. \tag{3.8}$$

The resulting amplitude curves for $\Delta m = 5, 10, 15\,\mathrm{ps}^{-1}$ are shown in figure 4*a*. On top of the curves, values obtained from the likelihood fit eq. (3.5) are also shown. The two fitting methods are indeed equivalent for $\omega \approx \Delta m$, as expected, while some difference appears for $\omega \neq \Delta m$.

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

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

$$\chi^2(\mathcal{A} + \sigma_{\mathcal{A}}) - \chi^2(\mathcal{A}) = 1\,, \tag{3.9}$$

which in turn gives

$$\sigma_{\mathcal{A}}(\omega) = \frac{1}{\sqrt{\displaystyle\int_{-\infty}^{\infty} dt f_\omega^2(t)/E(t)}}\,. \tag{3.10}$$

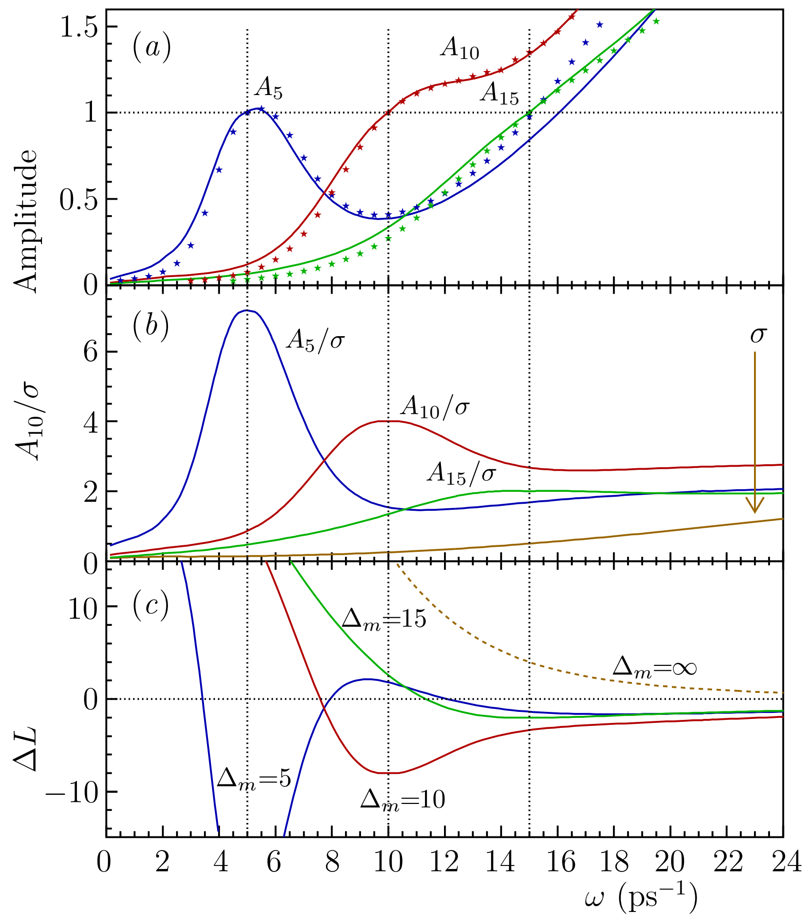

Figure 4: (a) Expected amplitude values for $\Delta m = 5, 10, 15, \mathrm{ps}^{-1}$. The curves refer to the χ^2 minimization, the points to the likelihood fit. (b) Amplitude significance curves (χ^2 fit). The expected shape of $\sigma(\omega)$ is also shown. (c) Expected shape of the likelihood, derived from the amplitude and its error. The dashed line corresponds to the limit $\Delta m \to \infty$. Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu m$ are assumed for the 3 plots.

The significance of the measured amplitude is therefore

$$S_{\Delta m}(\omega) = \frac{\mathcal{A}_{\Delta m}(\omega)}{\sigma_{\mathcal{A}}^{\Delta m}(\omega)} = \frac{\displaystyle\int_{-\infty}^{\infty} dt\, f_\omega(t)\, f_{\Delta m}(t)/E(t)}{\sqrt{\displaystyle\int_{-\infty}^{\infty} dt f_\omega^2(t)/E(t)}}\,. \tag{3.11}$$

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

The amplitude significance curves for $\Delta m = 5,\ 10,\ 15\ \mathrm{ps}^{-1}$ are shown in figure 4b. The normalization of the error, in the same figure, is arbitrarily chosen to have $\sigma_{\mathcal{A}} = 0.5$ at $\omega = 15\,\mathrm{ps}^{-1}$.

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

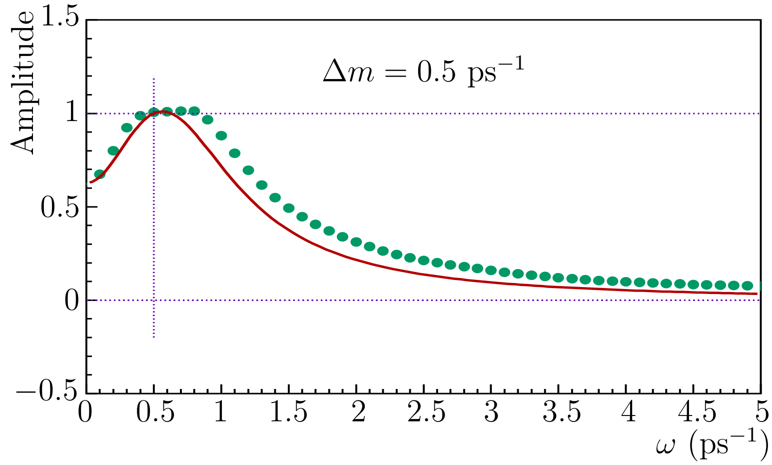

Figure 5: The full curve gives the expected shape of the amplitude for a signal at $\Delta m = 0.5\,\text{ps}^{-1}$ when all resolution effects are neglected. The dots are obtained with a toy experiment in which resolution effects are simulated (where $\delta_p = 0.15$ and $\sigma_l = 250\,\mu m$). The two shapes are in agreement.

The expected shape of the likelihood, as calculated from the amplitude and its error using eq. (2.5), is shown in figure 4*c*.

3.1 Limits for small and large Δm

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

Small Δm. If $\delta_l \ll 1/\Delta m, \delta_p/\Gamma \ll 1/\Delta m$, the oscillation is slow and marginally affected by the resolution. This limit holds in the case of B_d oscillations. If the resolution effects are neglected, eq. (3.8) can be rewritten as

$$\mathcal{A}_{\Delta m}(\omega) = \frac{\displaystyle\int_0^\infty dt\,\Gamma\exp(-\Gamma t)\cos\omega t\cos\Delta m t}{\displaystyle\int_0^\infty dt\,\Gamma\exp(-\Gamma t)\cos^2\omega t}\,, \tag{3.12}$$

which gives

$$\mathcal{A}_{\Delta m}(\omega) \approx \frac{\Gamma^2/\big(\Gamma^2 + (\omega+\Delta m)^2\big) + \Gamma^2/\big(\Gamma^2 + (\omega-\Delta m)^2\big)}{1 + \Gamma^2/\big(\Gamma^2 + 4\omega^2\big)}\,. \tag{3.13}$$

The resulting shape is shown in figure 5. The dots superimposed are obtained with toy experiments generated at the same value of the frequency, including resolution effects (for details on the simulation see section 4; the parameters used in the generation are those of samples **S** there defined). The two shapes are in qualitatively good agreement.

Large Δm. In this limit, which corresponds to the regime of B_s oscillations, the resolution effects dominate. If $\Delta m_\mathrm{s} \approx 15\,\mathrm{ps}^{-1}$ and $\delta_p = 0.15$, then $\delta_p/\Gamma \simeq 0.23\,\mathrm{ps}$, which is larger than $1/\Delta m_\mathrm{s} \approx 0.07\,\mathrm{ps}$ and therefore implies that only events with small proper time contribute to the sensitivity. Similarly, taking $\delta_l = 0.14\,\mathrm{ps}$ gives $\delta_l > 1/\Delta m_\mathrm{s}$, which implies a substantial damping of the amplitude of the oscillating term due to the decay length resolution. In this case, a useful approximation is to assume that the term $E(t)$ in eq. (3.8) varies slowly compared to the fast oscillating term $f_\omega(t)$, which is nonzero in a limited time range (figure 3), and take it out of the integral. In this way the expression can be simplified and rewritten in terms of the Fourier transformation of the oscillating components,

$$\mathcal{A}_{\Delta m}(\omega) \approx \frac{\displaystyle\int_{-\infty}^{\infty} dt\, f_\omega(t)\, f_{\Delta m}(t)}{\displaystyle\int_{-\infty}^{\infty} dt\, f_\omega^2(t)} = \frac{\displaystyle\int_{-\infty}^{\infty} d\nu\, \tilde{\mathrm{f}}_\omega(\nu)\, \tilde{\mathrm{f}}_{\Delta m}(\nu)}{\displaystyle\int_{-\infty}^{\infty} d\nu\, \tilde{\mathrm{f}}_\omega^{\,2}(\nu)}\,. \tag{3.14}$$

The approximation is valid only if both ω and Δm are large. The functions $\tilde{\mathrm{f}}_\omega$ are shown in figure 6*a* for a few different values of $\omega \geq 10\,\mathrm{ps}^{-1}$. Figure 6*c* shows the product of two of these Fourier transformations to illustrate the behaviour of the ratio in eq. (3.14).

With increasing Δm, the frequency spectra, $\tilde{\mathrm{f}}_\omega$, become broader and smaller in amplitude. High *true* frequencies, Δm, have their spectrum damped faster than low frequencies, and the peak at $\omega \approx \Delta m$ disappears for Δm well beyond the sensitivity (figure 6*b*). Due to the broadening of the spectra, the product $\tilde{\mathrm{f}}_{\omega_1}(\nu)\tilde{\mathrm{f}}_{\omega_2}(\nu)$ is peaked around the smallest between ω_1 and ω_2 (figure 6*c*). This fact implies that when a sample with oscillations at frequency Δm is analysed with a function containing a frequency $\omega < \Delta m$, the measured amplitude is dominated by the frequencies around ω; therefore the shape of $\mathcal{A}_{\Delta m}(\omega)$ for $\omega < \Delta m$ resembles that of $\tilde{\mathrm{f}}_{\Delta m}(\omega)$. For $\omega > \Delta m$ the frequencies around Δm are always tested, with a normalization factor which increases with ω (eq. (3.14), the denominator decreases very fast), therefore $\mathcal{A}_{\Delta m}(\omega)$ increases monotonically.

In order to understand better the effect of the decay length and proper time resolution, it is useful to study them separately. Setting $\delta_p = 0$ in eq. (3.3), the following simplified expression can be obtained,

$$\tilde{\mathrm{f}}_{\Delta m}(\omega) = \frac{1}{2}\left[\frac{\Gamma^2}{\Gamma^2 + (\omega + \Delta m)^2} + \frac{\Gamma^2}{\Gamma^2 + (\omega - \Delta m)^2}\right] \exp\left(-\frac{\delta_l^2\omega^2}{2}\right), \tag{3.15}$$

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

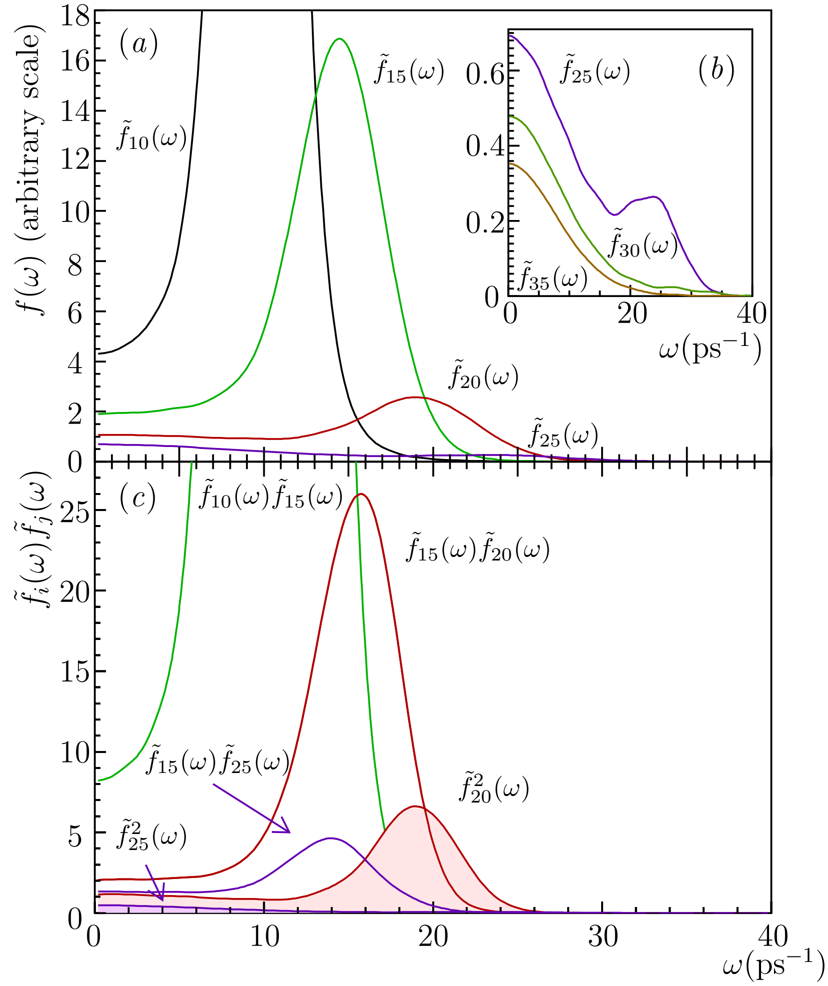

Figure 6: (a) Expected shapes of the Fourier spectra $\tilde{\mathrm{f}}_{\Delta m}$ for different values of Δm. The spectra become broader and lower in amplitude when Δm increases. (b) Detail of the spectra for high Δm. (c) Products of pairs of Fourier spectra. The resulting functions are peaked around the smallest of the two frequency values. Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu m$ are assumed.

Considering the momentum resolution alone the following expression is obtained,

$$\tilde{\mathrm{f}}_{\Delta m}(\omega) = \int_{-\infty}^{\infty} d\nu \frac{1}{2}\left[\frac{\Gamma^2}{\Gamma^2 + (\nu + \Delta m)^2} + \frac{\Gamma^2}{\Gamma^2 + (\nu - \Delta m)^2}\right] \exp\left(-\frac{(\omega - \nu)^2}{2\,(\delta_p \omega)^2}\right), \tag{3.16}$$

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

A broader frequency spectrum corresponds to a broader structure in the amplitude spectrum, or, equivalently, to higher correlations between values of the

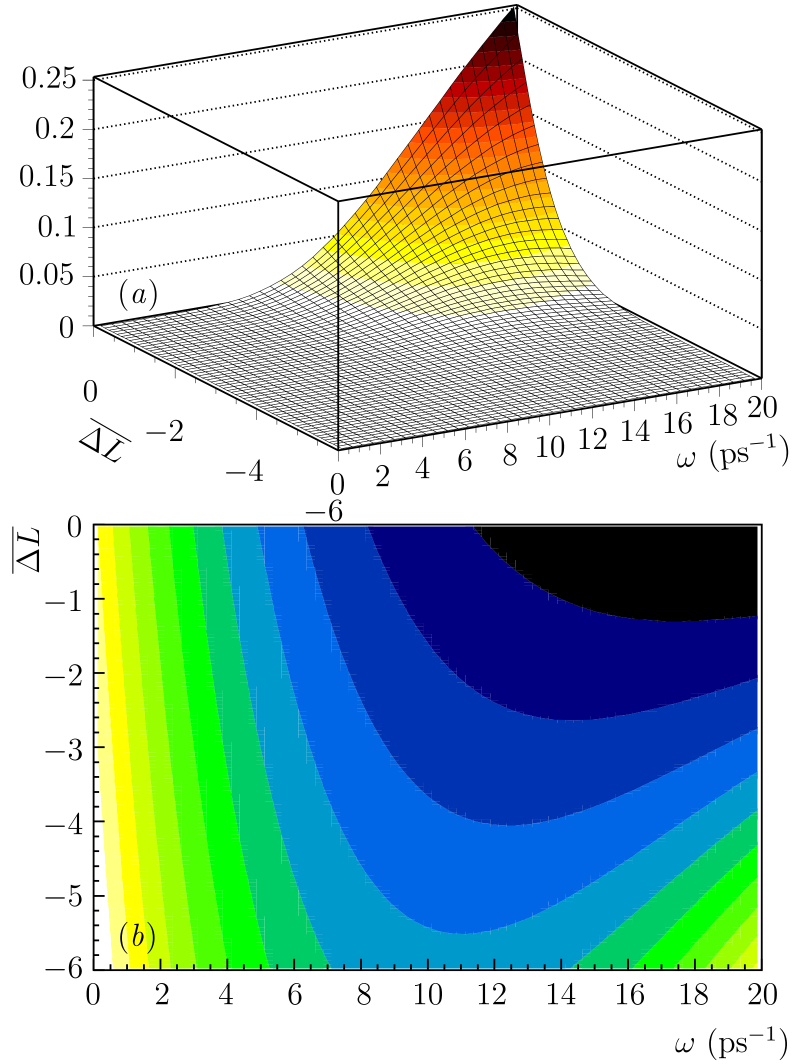

Figure 7: (a) Probability to find a $\Delta\mathcal{L}$ value lower than $\overline{\Delta\mathcal{L}}$ at a frequency ω. (b) Contours of equal probability in the $(\overline{\Delta\mathcal{L}}, \omega)$ plane. Values between -3 and -5 are found with higher probability when the error on the amplitude is from 0.5 (at $\omega = 15\,\mathrm{ps}^{-1}$ in this example) to 0.3 ($\omega \approx 11\,\mathrm{ps}^{-1}$). The contours shown are equidistant on a logarithmic scale. Resolutions of $\delta_p = 0.15$ and $\delta_l = 0.14\,\mathrm{ps}$ are assumed.

amplitude measured at different frequencies. This property is relevant for the confidence level estimation as explained in section 5.

3.2 Fluctuations

The expected shape of the likelihood for a sample with oscillations at a frequency far beyond the sensitivity is shown in figure 4c. In a given frequency range, statistical fluctuations of the likelihood can produce values below 0 which can fake a signal. The probability of observing that $\Delta\mathcal{L}$ is lower than a given value $\overline{\Delta\mathcal{L}}$ at a given frequency ω can be estimated from eq. (2.5), using the fact that the errors on the measured amplitudes are found to be gaussian with high precision

$$\mathcal{P}(\Delta\mathcal{L},\omega) \equiv \mathcal{P}(\Delta\mathcal{L} < \overline{\Delta\mathcal{L}})_\omega = \frac{1}{2}\,\mathrm{erfc}\left[\frac{-\overline{\Delta\mathcal{L}}\,\sigma_{\mathcal{A}}(\omega) + 1/(2\sigma_{\mathcal{A}}(\omega))}{\sqrt{2}}\right]. \tag{3.17}$$

The function $\mathcal{P}(\Delta\mathcal{L},\omega)$ is shown in figure 7*a*, where the same parameters and normalization as for figure 4 are used. This function can be used as an estimator of the signal-ness of a given sample. Estimator contours, equidistant on a logarithmic scale, are drawn in figure 7*b*. Small negative values are most probable at high frequencies, while for larger negative values the maximum of the probability is found at lower frequencies.

4. The toy experiments

The probability that the minimum observed in the likelihood (figure 2) is caused by a fluctuation can be evaluated by means of toy experiments with the above estimator eq. (3.17).

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

At each frequency point, the probability that the measured $\Delta\mathcal{L}$ is lower than a given value $\overline{\Delta\mathcal{L}}$ can be calculated as explained in section 3.2, starting from the errors on the measured amplitude. This procedure cannot be applied to the minimum, since $\Delta\mathcal{L}_{\mathrm{min}}$ is not an "unbiased" value, but it is chosen as the lowest value found over a certain frequency range explored.

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

The only viable possibility is to calibrate the significance of the structure observed with the help of toy experiments. The worldwide combination includes many analyses, and a detailed simulation of each of them is highly impractical. The procedure adopted here is to choose a set of parameters for the generation of the toy experiments such that each experiment is equivalent to the world average. The set of parameters cannot be uniquely determined from the data: it turns out that some parameters need to be fixed *a priori*, and therefore the dependence of the result obtained upon the particular choice adopted needs to be understood. The possible effects of the lack of a detailed simulation are investigated in section 5.1 by studying the dependence of the correlations in the amplitude measurements upon the parameters chosen to generate the toy experiments.

JHEP08(1999)004

4.1 Generation

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

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

$$l_0 = \frac{t_0 p_0}{m} c \,. \tag{4.1}$$

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

$$t = \frac{lm}{pc} \,. \tag{4.2}$$

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

4.2 The choice of the parameters

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

The errors on the measured amplitudes can be formally written as (see also eq. (3.10))

$$\sigma_{\mathcal{A}}^{-1}(\omega) = \sqrt{N} f_{B_s}(1-2\eta)\Sigma(\delta_p, \sigma_l, \omega)\,. \tag{4.3}$$

The factor $\kappa = \sqrt{N}\, f_{B_s}\,(1 - 2\,\eta)$ gives the normalization of the error distribution, without affecting the shape, and obviously the three parameters can not be disentangled. It is chosen to fix f_{B_s} and η to some "typical" values (namely $f_{B_s} = 0.15$, $\eta = 0.25$), and adjust N to fit the data errors. The effect of a different choice which yields the same κ value is investigated in section 5.1.

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

Samples are generated at three starting points for σ_l, which are chosen to be $200\,\mu$m, $250\,\mu$m and $300\,\mu$m, each with 30000 events and the other parameters as described above, and reported in more detail in section 4.3.

For each value of σ_l, the number of events is tuned by minimizing the sum of the differences with the data errors,

$$\sum_i \left(\sigma_{\mathcal{A}}^{\text{data}} - \sigma_{\mathcal{A}}^{\text{toy}}\right)_i^2, \tag{4.4}$$

where the scaling low of eq. (4.3) is used (figure 8*a*). The three minima found are then compared and interpolated with a parabolic fit (figure 8*b*) to find the absolute minimum, which turns out to be very close to $250\,\mu$m. The number of events needed at this point is 34000.

4.3 Samples description

On the basis of the procedure described in section 4.2, a set of parameters **S** is defined as follows

- a single *purity* class: 15% B_s 38% B_d 38% B^+ 9% Λ_b;
- a single *tagging* class: mistag rate $\eta = 25\%$ for all species;
- a single *resolution* class

$$\left.\begin{aligned} \sigma_l &= 250\ \mu\text{m}\,, \\ \frac{\sigma_p}{p_0} &= 0.15 \end{aligned}\right\} \quad \text{both gaussian with no tails}\,; \tag{4.5}$$

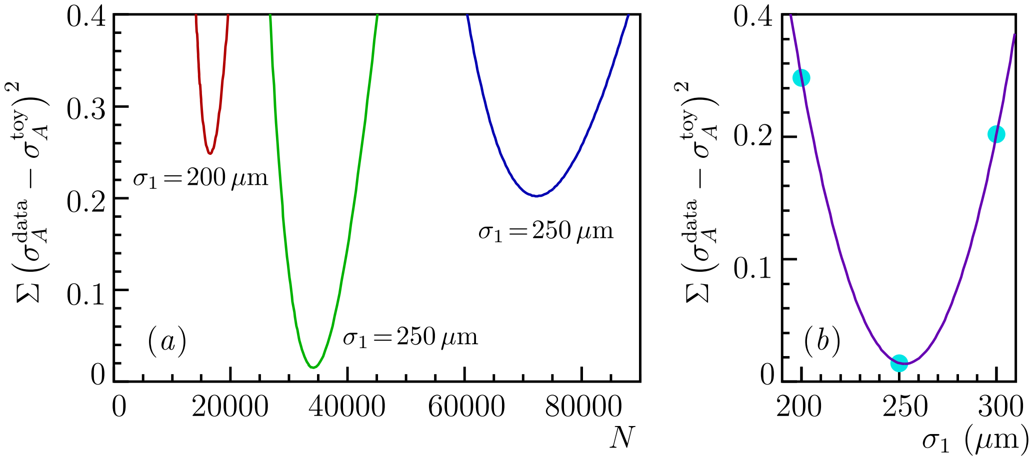

Figure 8: (a) Optimization of the number of events for three different values of the decay length resolution. (b) Choice of the optimal decay length resolution.

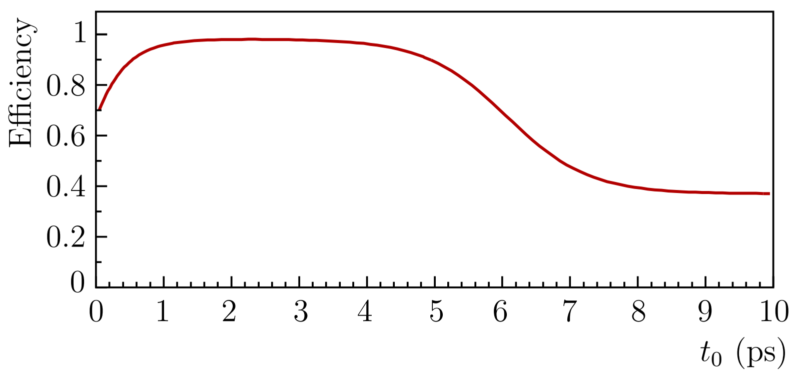

Figure 9: Shape of the reconstruction efficiency as a function of true proper time. The normalisation of the vertical scale is arbitrary.

- Monte Carlo parametrized efficiency (taken from the analysis of ref. [5]). The curve is shown in figure 9;
- b hadron masses and lifetimes, and $\Delta m_{\rm d}$ from ref. [6];
- $\langle x_b \rangle = 0.7$;
- $\Delta m_{\rm s}$ fixed at different values, according to the study considered;
- statistics of 34000 b hadron decays.

A second set of parameters $\mathbf{S}'$ is defined to generate a second family of toy experiments. The momentum resolution is chosen to be $\sigma_p/p_0 = 0.07$, which is significantly better than what is typically achieved in inclusive analyses. In order to keep the agreement with the world average errors on the measured amplitudes, the number of events is reduced to 29000 (obtained with the procedure described

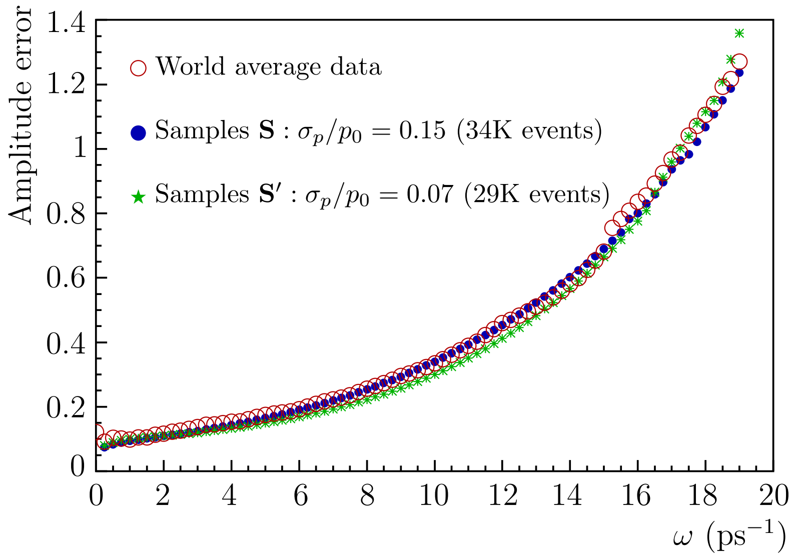

Figure 10: Amplitude errors comparison: simulated experiments *versus* world average data.

in section 4.2). The other parameters are left unchanged. These experiments are used in the following to investigate the dependence of the confidence level upon the momentum resolution.

The errors on the amplitude, $\sigma_{\mathcal{A}}$, obtained with these two sets of experiments are compared to the errors from the combined data in figure 10. The step at $\omega = 15\,\mathrm{ps}^{-1}$ could be reproduced by averaging, for each "experiment", two "analyses", of which one has its scan stopped at that point. No attempt was made in this direction.

A third set of samples $\mathbf{S}''$ with $\delta_p = 0.15$, $\sigma_l = 200\,\mu$m and statistics of 16500 decays (which correspond to the optimization of figure 8) is used to investigate the dependence of the confidence level upon the decay length resolution.

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

In figure 11 the expected shape of the amplitude and the likelihood is shown, as obtained by averaging 1000 samples of type **S**, generated with $\Delta m_\mathrm{s} = 150\,\mathrm{ps}^{-1}$. The expected value is consistently zero, and the errors on $\mathcal{A}$ are gaussian, which confirms the validity of the amplitude method to set limits on the oscillation frequency.

5. The estimate of the significance

As demonstrated in section 3.2, the probability that, at a given point in the frequency scan, a value of the likelihood $\Delta\mathcal{L} < \overline{\Delta\mathcal{L}}$ be found, can be calculated, given $\overline{\Delta\mathcal{L}}$, from the error on the measured amplitude, which is available from the data.

JHEP08(1999)004

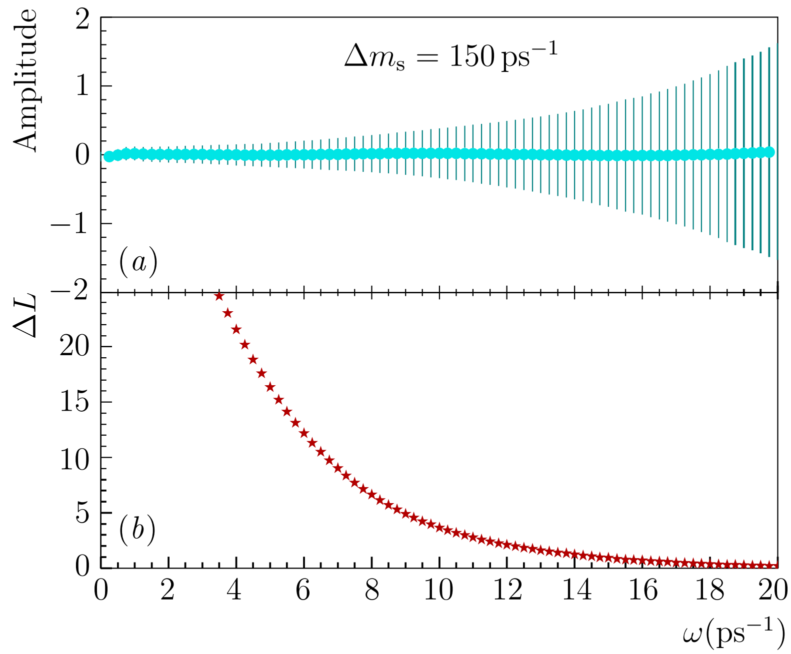

Figure 11: (a) Expected amplitude and error for samples of type **S**, with $\Delta m_{\mathrm{s}} = 15\mathrm{ps}^{-1}$, as a function of ω. (b) Expected likelihood shape. The plots are obtained by averaging 2000 samples. Resolutions of $\delta_p = 0.15$ and $\sigma_l = 250\,\mu m$ are assumed.

For the purpose of establishing the significance of the minimum, however, this probability is not enough, since what is needed is the probability that anywhere in the range explored a configuration more unlikely than the one observed may appear (in the hypothesis of large Δm_{s}). This significance is driven not only by the errors, but also by the correlations between the amplitude measurements at different frequencies, which are not controlled from the data, and might depend on the particular combination of parameters chosen for the simulation. It is therefore mandatory to identify the most relevant sources of systematic uncertainty which might affect the extraction of the confidence level. This point is investigated in the next section.

5.1 Correlations

From the discussion of section 3, it turns out that the momentum resolution is the most critical parameter to determine the point-to-point correlation in the amplitude scan. In a sample with better momentum resolution, correlations are smaller and therefore the probability of having significant deviations from $\mathcal{A} = 0$ in a sample with no signal is larger, in a given frequency range explored.

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

$$\langle \mid \mathcal{A}_i - \mathcal{A}_j \mid \rangle = \sqrt{\frac{2}{\pi}} \sqrt{{\sigma^i_{\mathcal{A}}}^2 + {\sigma^j_{\mathcal{A}}}^2} \,. \tag{5.1}$$

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

For each of the four set of parameters, $\mathbf{S}$, $\mathbf{S}'$, $\mathbf{S}''$ and $\mathbf{s}$ defined in section 4.3, 150 samples are produced, and the quantity $\langle \mid \mathcal{A}_i - \mathcal{A}_j \mid \rangle$ is calculated, for $i - j = 1, 4, 7, 10$. The results are shown in figure 12, where they are compared with the expectation for no point-to-point correlation.

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

Samples $\mathbf{S}'$ can be used to estimate a "systematic uncertainty" on the confidence level obtained, coming from the specific choice of the parameters used in the simulation.

5.2 The confidence level

The significance of the minimum observed in the $\Delta\mathcal{L}$ distribution (figure 2) is estimated by computing the probability that a structure as or more unlikely is observed in a sample with Δm_{s} far beyond the sensitivity.

In order to do that, it is taken into account that the probability of observing a given value of $\Delta\mathcal{L}$ is a non-trivial function of ω. Probability contours in the $(\Delta\mathcal{L}, \omega)$ plane (as in figure 7*b*) are built from the data errors. The contour corresponding to the data sample is computed. $N_{\mathrm{exp}} = 2000$ samples of type $\mathbf{S}$ with $\Delta m_{\mathrm{s}} = 150\,\mathrm{ps}^{-1}$ are analysed and the number $N^{\mathrm{out}}_{\mathrm{exp}}$ of those that give a minimum $\Delta\mathcal{L}_{\mathrm{min}} < 0$ outside the contour corresponding to the data is recorded. Since the expected value of the likelihood is positive for all frequencies (see figure 11*b*), occasionally the minimum in the range $0 - 19\,\mathrm{ps}^{-1}$ is also positive. These minima are not counted, independently of the frequency at which they occur, since thay can not be interpreted as a signal of oscillations.

The population of the toy experiments in the $(\Delta\mathcal{L}, \omega)$ plane along with the point corresponding to the data sample, is shown in figure 13.

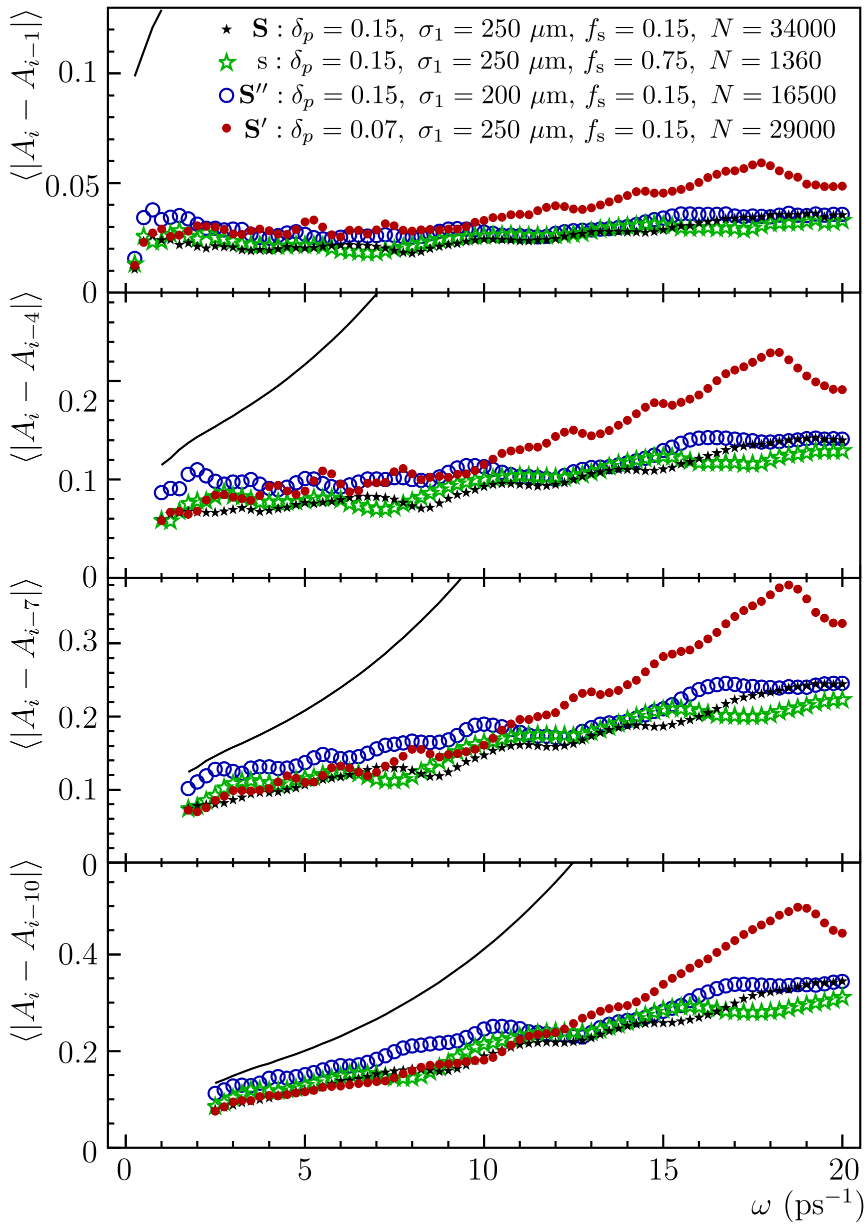

Figure 12: Point-to-point fluctuations for four sets of samples. From top to bottom, the average difference between points distant 1–4–7–10 steps in the amplitude scan are shown. The line corresponds to the limit of no correlation between points.

The confidence level is computed as

$$1 - \text{C.L.} \equiv \frac{N_{\text{exp}}^{\text{out}}}{N_{\text{exp}}} = 0.021 \pm 0.003\,. \tag{5.2}$$

The study is repeated with 2000 samples of type $\mathbf{S}'$, and yields

$$1 - \text{C.L.} = 0.033 \pm 0.004\,. \tag{5.3}$$

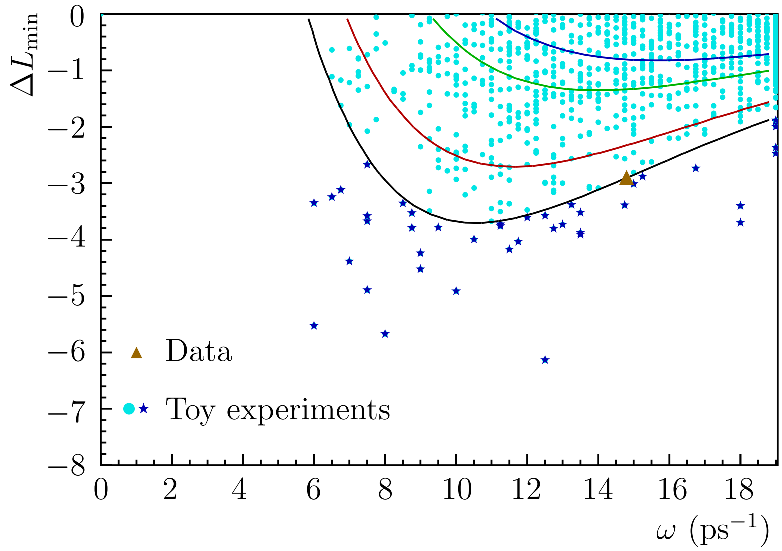

Figure 13: Minima of $\Delta\mathcal{L}$ for 2000 samples of type **S**, with $\Delta m_s = 150\,\mathrm{ps}^{-1}$. The curves represent contours of equal probability of observing a value of $\Delta\mathcal{L}$ smaller than $\overline{\Delta\mathcal{L}}$, as a function of ω (as in figure 7).

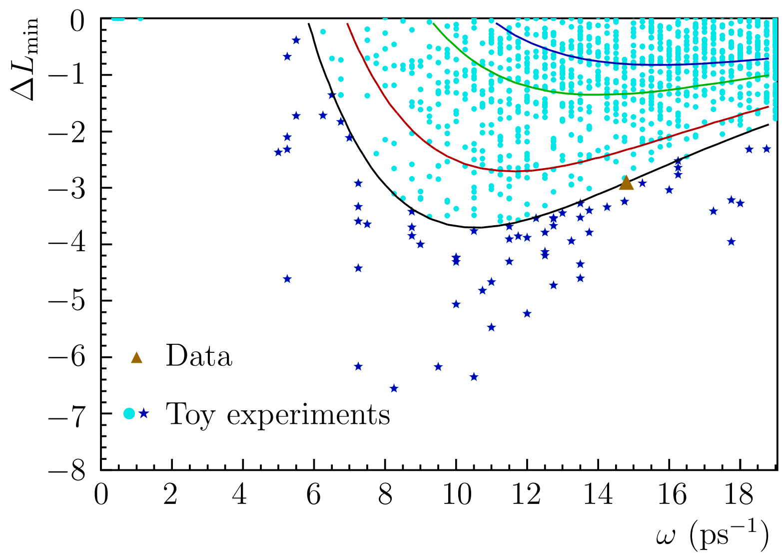

Figure 14: Minima of $\Delta\mathcal{L}$ for 2000 samples of type $\mathbf{S}'$ with $\Delta m_s = 150\,\mathrm{ps}^{-1}$. The same curves as in figure 13 are shown.

This value has to be understood as a conservative estimate of the probability of statistical fluctuations, since it is obtained with experiments built to have lower point-to-point correlations than that expected for the average of real analyses. The distribution of the minima for this case is shown in figure 14. The difference between the values of eq. (5.2) and (5.3) gives an upper limit for the uncertainty coming from the lack of a detailed simulation.

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

5.3 Comparison with the oscillation hypothesis

In order to check that the amplitude spectrum observed in the data is in qualitative agreement with the hypothesis of oscillations, 500 samples of type **S** have been produced, with $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$. The expected amplitude and error at each frequency value are shown in figure 15, with the data points superimposed. The agreement is good over the whole frequency range.

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

A simple check is performed instead. The 500 samples with oscillation at a value $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$ are analysed in terms of their incompatibility with the no-oscillation hypothesis. The scatter plot of the likelihood minima in the $(\overline{\Delta\mathcal{L}}, \omega)$ plane, as for the samples with $\Delta m_s = 150\,\mathrm{ps}^{-1}$, is presented in figure 16.

An enhanced density in the region $14\,\mathrm{ps}^{-1} < \omega < 16\,\mathrm{ps}^{-1}$, $-3 < \Delta\mathcal{L}_{\min} < -1$ is seen in the plot. A cluster of experiments with minima at $\omega = 19\,\mathrm{ps}^{-1}$ is also clearly visible: for these experiments the lowest point of the likelihood was at the boundary

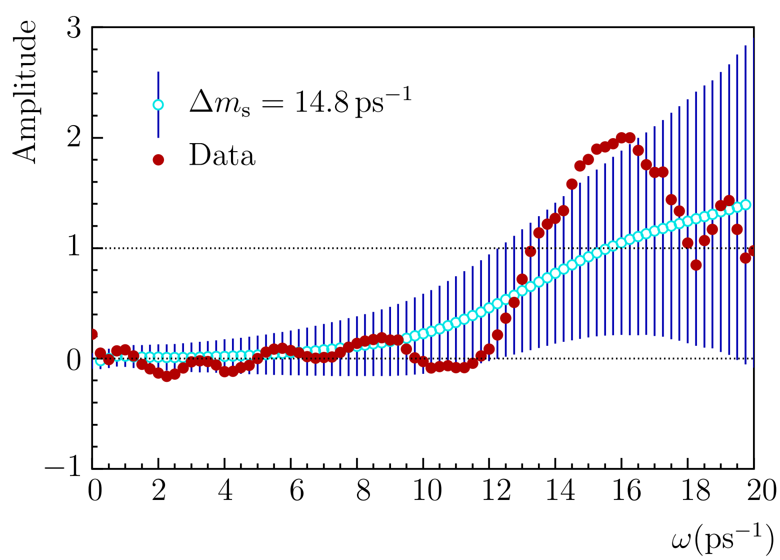

Figure 15: Average amplitude and expected error as a function of ω for a signal at $\Delta m_s = 14.8\,\mathrm{ps}^{-1}$. The amplitude values, obtained by averaging 500 toy experiments, are in good agreement with the data measurements (solid points).

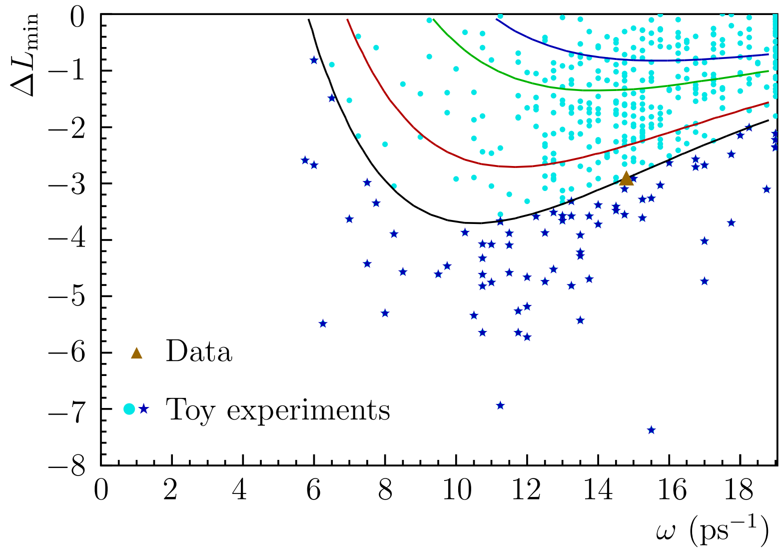

Figure 16: Minima of $\Delta\mathcal{L}$ for 500 samples of type **S**, with $\Delta m_{\mathrm{s}} = 14.8\,\mathrm{ps}^{-1}$. The same curves as in figure 13 are shown.

of the region analysed. Experiments with $\Delta\mathcal{L}_{\mathrm{min}} < -5$ appear at frequencies lower than the true one, where fluctuations which can produce deep minima are more likely.

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

6. Conclusion

The likelihood profile as a function of ω derived from the combined amplitude measurements available at the time of the 1999 Winter Conferences shows a minimum at $\omega = 14.8\,\mathrm{ps}^{-1}$. The depth of the minimum compared to the asymptotic value for $\omega \to \infty$ is $\Delta\mathcal{L}_{\mathrm{min}} = -2.9$. The intervals at which the $\Delta\mathcal{L}_{\mathrm{min}} + 1/2$ and $\Delta\mathcal{L}_{\mathrm{min}} + 2$ levels are crossed are

$$\omega = 14.8 \pm 0.5 \text{ ps}^{-1} \quad (\Delta\mathcal{L}_{\mathrm{min}} + 1/2 \text{ interval})\,, \tag{6.1}$$

$$\omega = 14.8 \; ^{+2.7}_{-1.8} \text{ ps}^{-1} \quad (\Delta\mathcal{L}_{\mathrm{min}} + 2 \text{ interval})\,. \tag{6.2}$$

The significance of the minimum observed in the likelihood cannot be calculated analytically, but needs to be calibrated using toy experiments. With the method proposed here, the probability that the result observed is produced by a statistical fluctuation in a sample with no signal is found to be

$$1 - \text{C.L.} \approx 3\,\%\,. \tag{6.3}$$

The uncertainty on this estimate coming from the lack of a detailed simulation of the individual analyses contributing to the average is estimated to be below 1 %.

Acknowledgments

We would like to thank Hans-Günther Moser for following our work and giving constructive feedback, and Olivier Leroy for many lively discussions and for checking all calculations throughout the paper. We are grateful to Patrick Janot and Gigi Rolandi for their advice and help all along the development of the analysis and their careful reading of the manuscript.

References

[1] ALEPH collaboration, *An investigation of B_d^0 and B_s^0 oscillation*, *Phys. Lett.* **B 322** (1994) 441.

[2] References to all B_s oscillation analyses can be found in `http://www.cern.ch/LEPBOSC/references`:

ALEPH collaboration, *Study of the $B_s^0\bar{B}_s^0$ oscillation frequency using $D_{\rm s}^- l^+$ combinations in Z decays*, *Phys. Lett.* **B 377** (1996) 205; *Study of B_s^0 oscillations and lifetime using fully reconstructed $D_{\rm s}^-$ decays*, *Eur. Phys. J.* **C 4** (1998) 367; *B_s^0 oscillations study with the ALEPH detector at LEP*, *Eur. Phys. J.* **C 7** (1999) 553;
DELPHI collaboration, *Search for $B_s^0 - \bar{B}_s^0$ oscillations*, *Phys. Lett.* **B 414** (1997) 382; *Search for $B_s^0 - \bar{B}_s^0$ oscillations and measurement of the B_s^0 lifetime*, EPS-HEP Jerusalem contribution 457;
OPAL collaboration, *A study of B_s^0 Meson oscillation using hadronic Z^0 decays containing leptons*, `hep-ex/9907061`, to appear on *Eur. Phys. J. C*;
CDF collaboration, *A search for $B_s^0 - \bar{B}_s^0$ oscillations using the semileptonic decay $B_s^0 \to \phi\ell^+ X\nu$*, *Phys. Rev. Lett.* **82** (1999) 3576;
SLD collaboration, *Time dependence of $B_s^0 - \bar{B}_s^0$-mixing using inclusive and semileptonic B decays at SLD*, SLAC-PUB-7885, Contributed paper PA08-181 ICHEP98, Vancouver, July 1998.

[3] H.-G. Moser and A. Roussarie, *Nucl. Instrum. Methods* **A 384** (1997) 491.

[4] LEP B Oscillation Working Group, LEPBOSC note 99/1. Available from: `http://www.cern.ch/LEPBOSC/combined_results/may_1999/`.

[5] ALEPH collaboration, *Search for B_s^0 oscillations using inclusive lepton events*, *Eur. Phys. J.* **C 7** (1999) 553 [`hep-ex/9811018`].

[6] Particle Data Group, *Review of particle physics*, *Eur. Phys. J.* **C 3** (1998) 1.