HOW TO INCLUDE THE INFORMATION FROM THE B_s^0-$\bar{B}_s^0$ OSCILLATION FREQUENCY IN THE CKM FITS

F. Parodi[1)], *P. Roudeau*[2)], *A. Stocchi*[2)], *A. Villa*[3)]
[1)] Dipartimento di Fisica, Università di Genova and INFN, Via Dodecaneso 33, 16146 Genova, Italy
[2)] Laboratoire de l'Accélérateur Linéaire IN2P3-CNRS et Université de Paris-Sud, BP 34, F-91898 Orsay Cedex, France
[3)] Dipartimento di Fisica, Università di Milano, Via Celoria, 16 I-20133 Milano, Italy

Abstract
In this paper we discuss how to include the Δm_s information in the CKM fits starting from the amplitude spectrum given by the LEP Oscillation Working Group. The number usually quoted by the Working Group is the 95% C.L. limit. The possibility of using the Likelihood Ratio method is discussed and defended.

1 A brief introduction on the CKM Fits

The aim of the CKM fit is to extract the best estimate of the parameters of the Cabibbo Kobayashi Maskawa (CKM) matrix from the available experimental measurements and theoretical computations.

The elements of the CKM matrix are usually expressed in term of four parameters [1]: λ, $|V_{cb}|$, $\bar{\rho}$, $\bar{\eta}$.

In the Bayesian framework, adopted in this work, for the case of N parameters (except for $\bar{\rho}$ and $\bar{\eta}$) and M constraints the following relation holds:

$$f(\bar{\rho}, \bar{\eta}, x_1, x_2, ..., x_N | \hat{c}_1, \hat{c}_2, ..., \hat{c}_M) \propto \prod_{j=1,M} f_j(\hat{c}_j | \bar{\rho}, \bar{\eta}, x_1, x_2, ..., x_N) \prod_{i=1,N} f_i(x_i) \times f_o(\bar{\rho}, \bar{\eta})$$

where $f_{j(i)}$ are the p.d.f. for the parameters, $\hat{c}_j$ are the constraints (likelihoods) coming from experimental measurements and f_o is the a-priori p.d.f. for $(\bar{\rho}, \bar{\eta})$.

The p.d.f. for $(\bar{\rho}, \bar{\eta})$ (the two least known parameters) is obtained by integrating over the parameter space:

$$f(\bar{\rho}, \bar{\eta}) \propto \int \prod_{j=1,M} f_j(\hat{c}_j | \bar{\rho}, \bar{\eta}, x_1, x_2, ..., x_N) \prod_{j=1,N} f_i(x_i) dx_i \times f_o(\bar{\rho}, \bar{\eta}) \tag{1}$$

The prior p.d.f. f_o is chosen to be uniform over the $\bar{\rho}, \bar{\eta}$ plane.

The measurements of the oscillation frequencies Δm_d and Δm_s in the neutral B_d^0, B_s^0 meson systems are important constraints in the ($\bar{\rho}$-$\bar{\eta}$) plane:

$$\begin{array}{lcl} \Delta m_d & \propto & f_{B_d^0}^2 B_{B_d^0}((1-\bar{\rho})^2 + \eta^2) \\ \Delta m_d / \Delta m_s & \propto & (f_{B_d^0}^2 B_{B_d^0} / f_{B_s^0}^2 B_{B_s^0})((1-\bar{\rho})^2 + \eta^2) \end{array}$$

$f_{B_q}^2 B_{B_q}$ is a theoretical parameter which encapsulates the non-perturbative effects. Since the value of the ratio of these parameters is theoretically better determined than their absolute values, the second constraint is in principle more effective.

Unfortunately, whereas Δm_d has been precisely measured, no evidence has been found, until now, for B_s^0-$\bar{B}_s^0$ oscillations. Nevertheless the data sample used to extract limits on Δm_s contains more information than is summarized in the limit itself. In this work we propose and test a method to fully exploit this information.

2 B_s^0 oscillations and the Amplitude method

The search for a $B_s^0 - \bar{B}_s^0$ oscillations signal is usually pursued at LEP, SLD and CDF by using the Amplitude method [3]. Events are divided in two classes corresponding to oscillating and non-oscillating candidates, according to the information obtained by tagging the presence of a B or a $\bar{B}$ meson at the beam interaction time and at the b-hadron decay time. The expected decay-time distribution of events is obtained by using a function which contains the time distributions for all components, taking into account the imperfect event tagging, the different behaviour of b-hadron decays which depend on their type ($\bar{B}_d^0$, B^-, $\bar{B}_s^0$ or b-baryon), and the background components originating from charm and light flavours. The time distribution for the oscillating (non-oscillating) signal is given by

$$\mathcal{P}(B_s^0 \to \bar{B}_s^0(B_s^0)) = \frac{1}{2\tau_{B_s^0}} e^{-t/\tau_{B_s^0}} [1 - (+) \cos(\Delta m_s t)]. \tag{2}$$

Such theoretical time distributions are convoluted with the expected time resolution, the mistag probability and all the non-B_s^0 background distributions are included to get the p.d.f. for like and unlike events.

$$\mathcal{L}(\Delta m_s) = \prod_{i=1}^{N_{like}} \mathcal{P}^{like}(t_i, \Delta m_s) \prod_{i=1}^{N_{unlike}} \mathcal{P}^{unlike}(t_i, \Delta m_s)$$

The amplitude method consists of changing the p.d.f. for B_s^0 oscillations by multiplying the cosine term by $\mathcal{A}$

$$\begin{aligned} 1 - (+) \cos(\Delta m_s t) &\to 1 - (+) \mathcal{A} \cos(\Delta m_s t) \\ \mathcal{L}(\Delta m_s) &\to \mathcal{L}(\Delta m_s, \mathcal{A}) \end{aligned}$$

$\mathcal{L}(\Delta m_s, \mathcal{A})$ is then maximized w.r.t. $\mathcal{A}$ at different values of Δm_s; the typical output of an analysis is a set of values $\mathcal{A}(\Delta m_s)$ with their errors $\sigma_{\mathcal{A}}(\Delta m_s)$.

Fitted amplitude values from different analyses are combined [7] using a common set of determinations for external parameters and taking into account possible correlations between the different analyses. In this respect the use of the oscillation amplitude provides a simple framework for incorporating all these effects. The 95% C.L. upper limit on Δm_s for excluding oscillations corresponds to the value Δm_s^{lim} for which $\mathcal{A} + 1.645\ \sigma_{\mathcal{A}} = 1$, whereas the sensitivity is the value Δm_s^{sens} at which the 95% C.L. limit would be set if $\mathcal{A} = 0$ and is obtained from: $1.645\ \sigma_{\mathcal{A}} = 1$. The World Average computed by the LEP Oscillation Working Group [7] sets a lower limit at 15.0 ps^{-1} with a sensitivity at 18.0 ps^{-1} (see Figure 2).

3 Δm_s Information for CKM fits

The 95% C.L. limit is useful, as well as the sensitivity, to summarize in one single number the results of the analysis. However to include Δm_s in a CKM fit and to determine probability regions for the variables $\bar{\rho}$ and $\bar{\eta}$, continuous information about the degree of exclusion is needed.

Two different methods to include Δm_s in CKM fit will be reviewed and compared. The requirements for an optimal method are:

- the method should work independently of the significance of the signal: this criterion is important to avoid switching from one method to another because of the presence (absence) of a significant signal (whose definition is arbitrary);
- the probability regions derived should have correct coverage[1].

[1] In the Bayesian framework correct coverage is not guaranteed. In particular it depends on the choice of priors. Nevertheless the variations of the coverage, because of different choice of the priors, have been evaluated in the case of study and found to be negligible compared with the variations due to different methods.

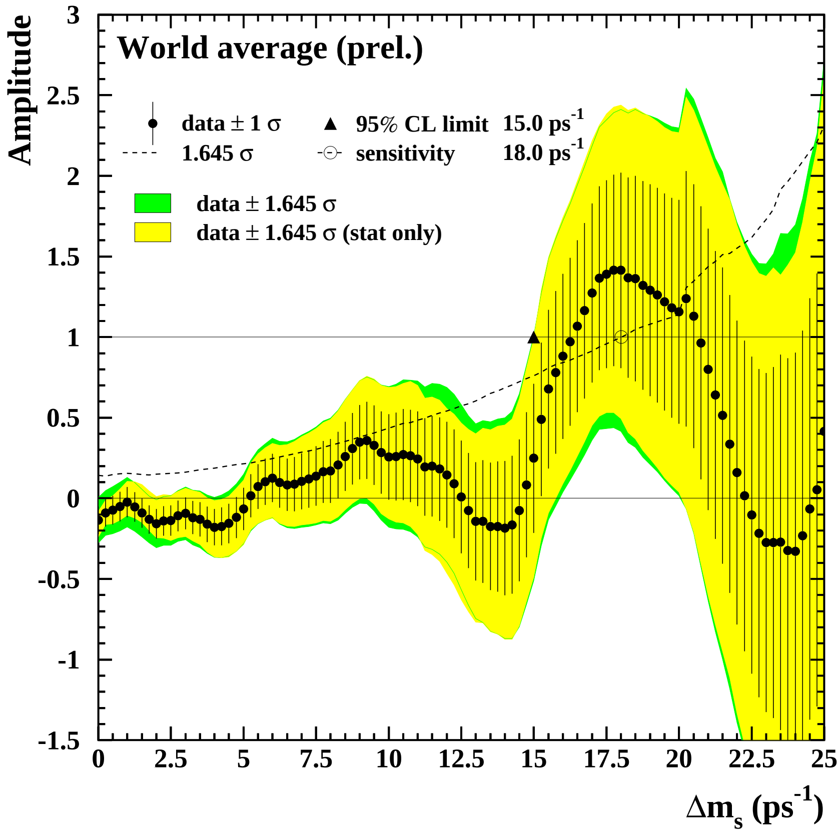

Fig. 1: World Average Amplitude spectrum as a function of the B_s^0 oscillation frequency (Δm_s).

3.1 Likelihood Ratio method $\mathcal{R}$

Recently it has been proposed to use the log-likelihood function $\Delta \log \mathcal{L}^{\infty}(\Delta m_s)$ referenced to its value obtained for $\Delta m_s = \infty$ [8, 2]. The log-likelihood values can easily be deduced from $\mathcal{A}$ and $\sigma_{\mathcal{A}}$, in the Gaussian approximation, by using expressions given in [3]

$$\Delta \log \mathcal{L}^{\infty}(\Delta m_s) = \log \mathcal{L}(\infty) - \log \mathcal{L}(\Delta m_s) = \frac{1}{2}\left[\left(\frac{\mathcal{A}-1}{\sigma_{\mathcal{A}}}\right)^2 - \left(\frac{\mathcal{A}}{\sigma_{\mathcal{A}}}\right)^2\right] = \left(\frac{1}{2} - \mathcal{A}\right)\frac{1}{\sigma_{\mathcal{A}}^2}\,, \quad (3)$$

$$\Delta \log \mathcal{L}^{\infty}(\Delta m_s)_{mix} = -\frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2}\,, \quad (4)$$

$$\Delta \log \mathcal{L}^{\infty}(\Delta m_s)_{nomix} = \frac{1}{2}\frac{1}{\sigma_{\mathcal{A}}^2}\,. \quad (5)$$

The last two equations give the average log-likelihood value when Δm_s corresponds to the true (Δm_s^{true}) oscillation frequency (*mixing* case) and when Δm_s is far from the true oscillation frequency ($|\Delta m_s - \Delta m_s^{true}| \gg \Gamma/2$, *no-mixing* case). Γ is the full width at half maximum of the amplitude distribution in the case of a signal; typically $\Gamma \simeq 1/\tau_{B_s^0}$.

The Likelihood Ratio R, defined as:

$$\mathrm{R}(\Delta m_s) = e^{-\Delta \log \mathcal{L}^{\infty}(\Delta m_s)} = \frac{\mathcal{L}(\Delta m_s)}{\mathcal{L}(\Delta m_s = \infty)}\,, \quad (6)$$

has been adopted in [2] to incorporate the Δm_s constraint in Eq. 1.

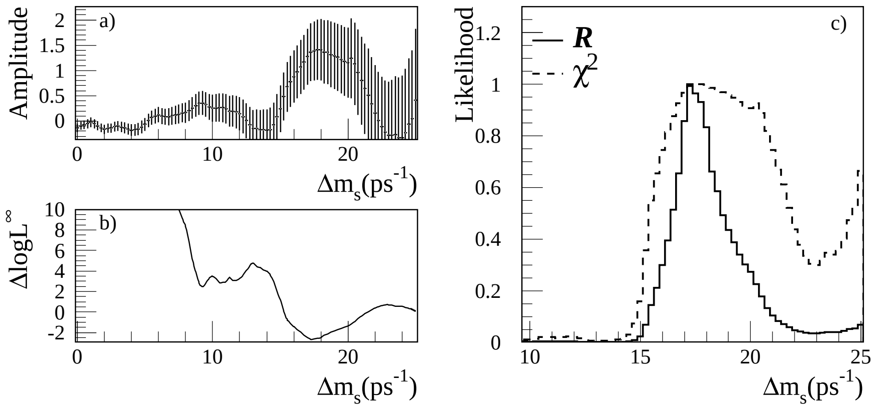

Fig. 2: World Average amplitude analysis: a) amplitude spectrum, b) $\Delta \log \mathcal{L}^{\infty}(\Delta m_s)$, c) comparison between the Likelihood Ratio method ($\mathcal{R}$) and the Modified χ^2 method (χ^2). The information in b) and in the solid histogram in c) is identical.

3.2 Modified χ^2 method

In the first CKM fits [4, 5, 6] the χ^2 of the complete amplitude spectrum w.r.t. 1 was used:

$$\chi^2 = \left(\frac{1-\mathcal{A}}{\sigma_{\mathcal{A}}}\right)^2 \tag{7}$$

This method has two main drawbacks:

- the sign of the deviation of the amplitude with respect to the value $\mathcal{A} = 1$ was not used, whereas it is expected that evidence for a signal would manifest itself by giving an amplitude value which is simultaneously compatible with $\mathcal{A} = 1$ and incompatible with $\mathcal{A} = 0$;
- values of $\mathcal{A} > 1$ are disfavoured w.r.t. $\mathcal{A} = 1$, while it is expected that, because of statistical fluctuations, the amplitude value corresponding to the "true" Δm_s value could be higher than 1. This problem was solved, in the early days of the use of Δm_s in CKM fits, by taking $\mathcal{A} = 1$ whenever it was in fact higher.

A modified χ^2 has been introduced in an "ad hoc" way in [10] to solve the second problem:

$$\chi^2 = 2 \cdot \left[Erfc^{-1}\left(\frac{1}{2}Erfc\left(\frac{1-A}{\sqrt{2}\sigma_{\mathcal{A}}}\right)\right)\right]^2 \tag{8}$$

4 Comparison of the two methods using the World Average Amplitude Spectrum

The variation of the amplitude as a function of Δm_s and the corresponding $\Delta \log \mathcal{L}^{\infty}(\Delta m_s)$ value are shown in Figure 2-a),b). The constraints obtained using the Likelihood Ratio method ($\mathcal{R}$) and the Modified χ^2 method (χ^2) are shown in Figure 2-c). For this comparison the Modified χ^2 method has been converted to a likelihood function using $\mathcal{L} \propto e^{-\frac{1}{2}\chi^2}$. This is possible, with the present data, since all the amplitude measurements are Gaussian.

It is clear that the two methods ($\mathcal{R}$ and χ^2) give very different constraints. In particular the Modified χ^2 method, with the present World Average, corresponds to a less tight constraint for the CKM fits (and in particular for the determination of the $\bar{\rho}$ and γ parameters).

It is thus important to establish which is the correct method, especially as doubts were formulated in [10] concerning the Likelihood Ratio approach.

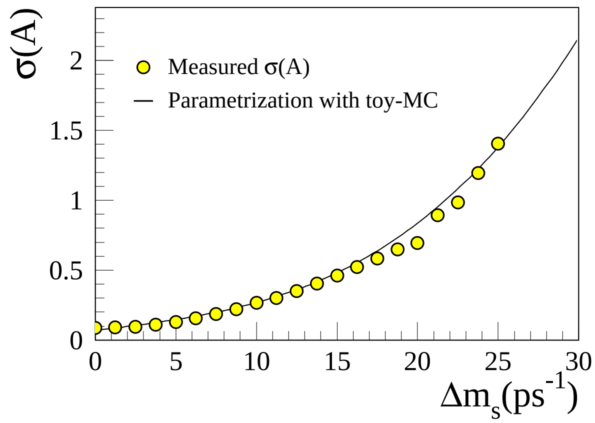

Fig. 3: Comparison between the error distribution computed with the toy-MC (solid line) and the measured amplitude errors (circles).

5 The Toy Monte Carlo

In order to test the two methods it is necessary to generate several experiments having similar characteristics to the data used for the World Average. We will call equivalent those experiments having the same $\sigma_{\mathcal{A}}$ behaviour.

The $\sigma_{\mathcal{A}}$ dependence on Δm_s can be reproduced by tuning the parameters of a fast simulation (toy-MC). The method used here is similar to the one presented in [9]. The errors on the amplitude can be written as:

$$\sigma_{\mathcal{A}}^{-1} = \sqrt{N}\,\eta_{B_s^0}\,(2\epsilon_d - 1)\,(2\epsilon_p - 1)\,W(\sigma_L, \sigma_P, \Delta m_s)$$

where N is the total number of events, $\eta_{B_s^0}$ the purity of the sample of B_s^0 decays, $\epsilon_{d(p)}$ the purity of the tagging at the decay (production) time, σ_L is the B_s^0 flight length uncertainty and σ_P the relative uncertainty on its momentum, W is the function that accounts for the damping of the oscillation due to the finite proper time resolution.

The parameters σ_L, σ_P and the global factor that multiply the W function are obtained by adjusting the simulated error distribution to the one measured with real events.

Figure 3 shows the agreement between the toy-MC calculation and real data up to Δm_s=25 ps^{-1} (the upper Δm_s value at which amplitudes are given in the combined plot).

An additional problem when using the World Average amplitude spectrum to construct the likelihood function for Δm_s is that, in principle, one would like to define the likelihood within the interval $[0, \infty]$ whereas the amplitude spectrum is defined only up to a certain value (for the present World Average the value is $25\ ps^{-1}$). A procedure has to be introduced to continue $\sigma_{\mathcal{A}}$ and $\mathcal{A}$.

The continuation for $\sigma_{\mathcal{A}}$ is shown in Figure 3. It has been verified (with an independent toy-MC) that the extrapolated errors approximate well the simulated error distribution.

The continuation of $\mathcal{A}$ is more delicate. In particular it is more sensitive to the real amplitude spectrum. Nevertheless if $\Delta m_s^{sens} << \Delta m_s^{last}$, the significance S ($S = \mathcal{A}/\sigma_{\mathcal{A}}$) is approximately constant. It is then a good approximation to take $\mathcal{A}(\Delta m_s) = \frac{\mathcal{A}(\Delta m_s^{last})}{\sigma_{\mathcal{A}}(\Delta m_s^{last})}\sigma_{\mathcal{A}}(\Delta m_s)$. Although this procedure is reasonable, it should be stressed that it is very desirable to have all the amplitudes (with errors) up to the Δm_s value where the significance is reasonably stable.

6 Comparison of the methods in case of an oscillation signal

In this section we test the behaviour of the two methods in the presence of a clear Δm_s oscillation signal.

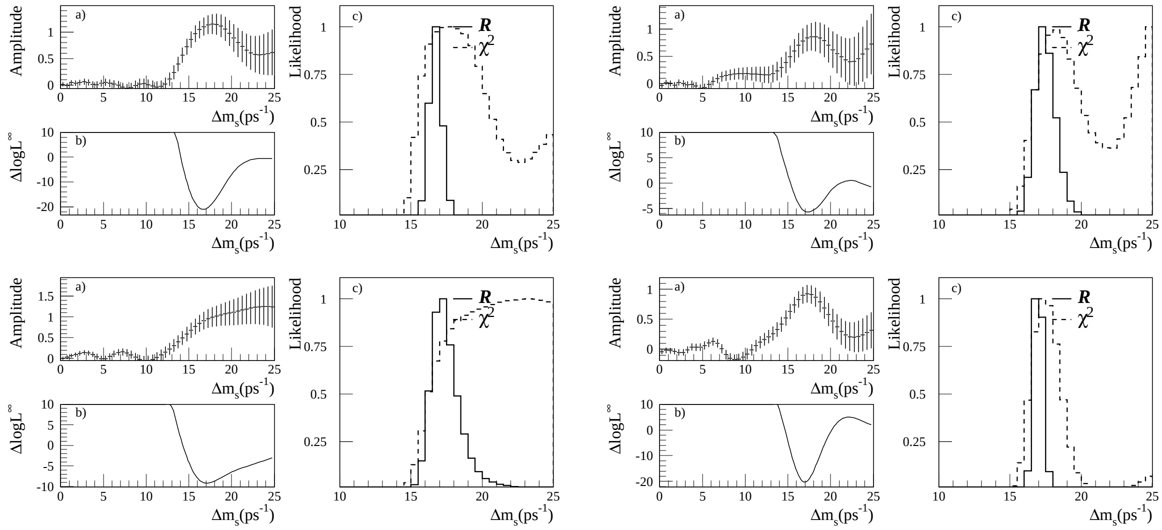

Fig. 4: Toy-MC analyses with Δm_s generated at 17 ps^{-1} corresponding to four virtual experiments. Each experiment is summarized in three plots: a) amplitude spectrum, b) $\Delta \log \mathcal{L}^{\infty}(\Delta m_s)$, c) comparison between Likelihood Ratio method ($\mathcal{R}$) and the Modified χ^2 method (χ^2).

For this reason we perform several Δm_s Toy-MC analyses with the same $\sigma_{\mathcal{A}}$ versus Δm_s behaviour as the World Average analysis. For this study we have generated a Δm_s signal at 17 ps^{-1} (corresponding to the value where there is the bump in the World Average amplitude spectrum).

Results are shown in Figure 4. It is clear that the Likelihood Ratio method is able to see the signal at the correct Δm_s value, whereas the Modified χ^2 method failed. The same exercise has been repeated for different generated values of Δm_s, always giving the same result.

7 Test of the coverage of the two methods applied to CKM Fits

In absence of a clear signal, the Likelihood Ratio method results in a Δm_s range which extends to infinity at any C.L. A criticism was made in [10] claiming that it is dangerous to use such information in CKM fits. We think that the best way to answer this remark is to test the coverage of the probability regions computed by the fit.

To do this we have prepared a simplified CKM fit where we measure the quantity $R_t(=\sqrt{(1-\rho)^2+\eta^2})$, which is one side of the Unitarity Triangle, using only the Δm_d and the $\Delta m_d/\Delta m_s$ constraints.

The set of constraints on the quantity R_t is:

$$\begin{array}{l} \Delta m_d = a^2 R_t^2 \\ \Delta m_d/\Delta m_s = b^2 R_t^2 \quad (\text{or} \quad \Delta m_s = a^2/b^2) \end{array}$$

where a and b are Gaussian distributed parameters with errors $\sigma_a = 20\%$ and $\sigma_b = 10\%$ (simulating the theoretical uncertainties involved in the computation).

Several experiments have been generated, each of them characterized by the following set of parameters:

R_t	
a_{theo}	extracted from the a distribution
b_{theo}	extracted from the b distribution
$\Delta m_d(theo)$	computed from R_t and a
$\Delta m_s(theo)$	computed from R_t and b
$\Delta m_d(exp)$	from $\Delta m_d(theo)$ smeared by the experimental resolution
Amplitude spectrum	from a toy-experiment generated with $\Delta m_s(theo)$

For each experiment a fit for R_t was performed and it was checked whether the generated value of R_t was inside the 68%, 95% and 99% probability regions defined by the Likelihood Ratio and by the Modified χ^2 methods. This exercise was repeated 1000 times, with similar and independent fits.

The frequency for experiments found inside the specified intervals obtained with the two methods are given in Table 1 and 2 respectively.

The Likelihood Ratio method always has close to the correct coverage. This is not the case of the Modified χ^2 method.

	68%	95%	99%
$\Delta m_s = 10$	67.5 ± 1.5	93.1 ± 0.8	98.1 ± 0.4
$\Delta m_s = 18.2$	71.4 ± 1.4	96.1 ± 0.6	99.6 ± 0.2
$\Delta m_s = 25$	69.5 ± 1.5	96.4 ± 0.6	99.3 ± 0.3

Table 1: Results obtained with the Likelihood Ratio method. For three different values of generated Δm_s (left column) we indicate the percentage of "experiments" for which the generated true value of R_t falls inside the 68%, 95% and 99% probability interval.

	68%	95%	99 %
$\Delta m_s = 10$	48.6 ± 1.6	83.8 ± 1.2	94.3 ± 0.7
$\Delta m_s = 18.2$	64.6 ± 1.5	93.0 ± 0.8	99.2 ± 0.3
$\Delta m_s = 25$	77.5 ± 1.5	98.2 ± 0.4	99.7 ± 0.2

Table 2: As for Table 1, but for the Modified χ^2 method.

8 Conclusions

In this paper we have studied the problem of including, in CKM fits, the information contained in the Δm_s World Average analysis.

We have demonstrated that the Likelihood Ratio method, proposed in [8, 2], is the optimal method because it gives probability intervals with correct coverage. As expected, it also gives the correct measurement of Δm_s, in the case of a signal.

References

[1] A. J. Buras, M. E. Lautenbacher and G. Ostermaier, *Phys. Rev.* **D 50** (1994) 3433.

[2] M. Ciuchini, G. D'Agostini, E. Franco, V. Lubicz, G. Martinelli, F. Parodi, P. Roudeau and A. Stocchi, *JHEP* **0107**(2001) 013 (hep-ph/0012308)

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

[4] P. Paganini, F. Parodi, P. Roudeau and A. Stocchi, *Phys. Scripta* **68** (1998) 556

[5] Y. Grossman, Y. Nir, S. Plaszczynski and M. H. Schune, *Nucl. Phys.* **B511** (1998) 69

[6] F. Parodi, P. Roudeau and A. Stocchi *Nuovo Cim.* **A112** (1999) 833

[7] The LEP B Oscillation Working Group, http://lepbosc.web.cern.ch/LEPBOSC/

[8] P. Checchia, E. Piotto and F. Simonetto, *Tampere 1999, High energy physics* 526 (hep-ph/9907300).

[9] G. Boix and D. Abbaneo, *Journal of High Energy Physics* **9908** (1999) 004.

[10] A. Hoeker, H. Lacker, S. Laplace and F. Le Diberder, *Eur. Phys. J.* **C21** (2001) 225 (hep-ph/0104062)