Chapter 4

CKM ELEMENTS FROM K AND B MESON MIXING

Conveners : J.M. Flynn, M. Paulini, S. Willocq.
Contributors: D. Abbaneo, C. Bozzi, A.J. Buras, R. Forty, R. Gupta, R. Hawkings, A. Hoecker, M. Jamin, P. Kluit, A. Kronfeld, V. Lacker, F. Le Diberder, L. Lellouch, C. Leonidopoulos, D. Lin, V. Lubicz, H.G. Moser, U. Nierste, J. Ocariz, F. Parodi, C. Paus, P. Roudeau, Y. Sakai, O. Schneider, A. Stocchi, C. Weiser, N. Yamada.

1. Basic formulae for particle–antiparticle mixing

1.1. K sector: basic formula for ε_K

In the $\mathrm{K}^0 - \overline{\mathrm{K}}^0$ system, to lowest order in electroweak interactions $\Delta S = 2$ transitions are induced through the box diagrams of Fig. 4.1. Including leading and next-to-leading QCD corrections in renormalization group improved perturbation theory the effective Hamiltonian for the $\Delta S = 2$ transitions for scales $\mu < \mu_c = \mathcal{O}(m_c)$ is given by

$$\begin{aligned}\mathcal{H}_{\mathrm{eff}}^{\Delta S=2} &= \frac{G_{\mathrm{F}}^2}{16\pi^2} M_{\mathrm{W}}^2 \left[\lambda_c^2 \eta_1 S_0(x_c) + \lambda_t^2 \eta_2 S_0(x_t) + 2\lambda_c\lambda_t \eta_3 S_0(x_c, x_t)\right] \times \\ &\quad \times \left[\alpha_s^{(3)}(\mu)\right]^{-2/9} \left[1 + \frac{\alpha_s^{(3)}(\mu)}{4\pi} J_3\right] Q(\Delta S = 2) + \mathrm{h.c.} \qquad (1)\end{aligned}$$

where $\lambda_i = V_{is}^* V_{id}$, $\alpha_s^{(3)}$ is the strong coupling constant in an effective three flavour theory and $J_3 = 307/162 = 1.895$ in the NDR scheme [1]. In (1), the relevant operator

$$Q(\Delta S = 2) = (\bar{s}\gamma_\mu(1-\gamma_5)d)(\bar{s}\gamma^\mu(1-\gamma_5)d), \qquad (2)$$

is multiplied by the corresponding Wilson coefficient function. This function is decomposed into a charm-, a top- and a mixed charm-top contribution. The functions $S_0(x_i)$ and $S_0(x_c, x_t)$ are given by ($x_i = m_i^2/M_{\mathrm{W}}^2$):

$$S_0(x_t) = \frac{4x_t - 11x_t^2 + x_t^3}{4(1-x_t)^2} - \frac{3x_t^3 \ln x_t}{2(1-x_t)^3}, \qquad S_0(x_c) = x_c, \qquad (3)$$

$$S_0(x_c, x_t) = x_c \left[\ln \frac{x_t}{x_c} - \frac{3x_t}{4(1-x_t)} - \frac{3x_t^2 \ln x_t}{4(1-x_t)^2}\right], \qquad (4)$$

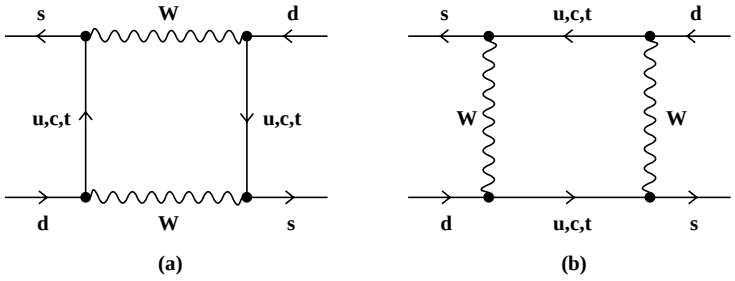

Fig. 4.1: *Box diagrams contributing to* $\mathrm{K}^0 - \overline{\mathrm{K}}^0$ *mixing in the SM.*

where we keep only linear terms in $x_c \ll 1$, but of course all orders in x_t. The exact expression can be found in [2].

Short-distance QCD effects are described through the correction factors η_1, η_2, η_3 and the explicitly $\alpha_{\rm s}$-dependent terms in (1). The NLO values of η_i are given as follows [1,3–6]:

$$\eta_1 = (1.32 \pm 0.32) \left(\frac{1.30\text{GeV}}{m_c(m_c)}\right)^{1.1}, \qquad \eta_2 = 0.57 \pm 0.01, \qquad \eta_3 = 0.47 \pm 0.05\ . \tag{5}$$

It should be emphasized that the values of η_i depend on the definition of the quark masses m_i. The ones in (5) correspond to $m_t \equiv m_t(m_t)$ and $m_c \equiv m_c(m_c)$. With this definition the dependences of η_2 on m_t and of η_3 on m_t and m_c are fully negligible but the dependence of η_1 on m_c turns out to be significant. It can be well approximated by the formula in (5). The scale dependence in $m_i(\mu_t)$, where $\mu_t = \mathcal{O}(m_t)$, present generally in the functions $S_0(x_t)$ and $S_0(x_t, x_c)$ is canceled to an excellent accuracy in the products $\eta_2 S_0(x_t)$ and $\eta_3 S_0(x_t, x_c)$. The corresponding scale dependence in $m_{\rm c}(\mu_c)$, where $\mu_c = \mathcal{O}(m_c)$, is cancelled to a large extent in the product $\eta_3 S_0(x_t, x_c)$ but remains still sizable in $\eta_1 S_0(x_c)$. As we use $m_c(m_c)$ and $m_t(m_t)$ we have included the left-over scale uncertainties due to μ_c and μ_t present in (1) in the errors of η_i that also include the uncertainties due to $\Lambda_{\overline{\rm MS}}$, the scale in the QCD running coupling. The small changes in η_1 and η_3 relative to the original papers are due to changes in $\alpha_s(M_Z)$.

Now, ε_K is defined by

$$\varepsilon_K = \frac{A(\mathrm{K_L} \to (\pi\pi)_{I=0})}{A(\mathrm{K_S} \to (\pi\pi)_{I=0})} \tag{6}$$

with I denoting isospin. From (6) one finds

$$\varepsilon_K = \frac{\exp(i\pi/4)}{\sqrt{2}\Delta M_K} \left(\mathrm{Im}M_{12} + 2\bar{\xi}\mathrm{Re}M_{12}\right), \qquad \bar{\xi} = \frac{\mathrm{Im}A_0}{\mathrm{Re}A_0} \tag{7}$$

with the off-diagonal element M_{12} in the neutral K-meson mass matrix representing K^0-$\overline{\mathrm{K}}^0$ mixing given by

$$2M_K M_{12}^* = \langle \overline{\mathrm{K}}^0 | \mathcal{H}_{\rm eff}(\Delta S = 2) | \mathrm{K}^0 \rangle\, . \tag{8}$$

The factor $2M_K$ reflects our normalization of external states and A_0 is the isospin amplitude. ΔM_K is the $\mathrm{K}_L - \mathrm{K}_S$ mass difference that is taken from experiment as it cannot be reliably calculated due to long distance contributions. The expression in (7) neglects higher order CP-violating terms: see the discussion in the review article in reference [7].

Defining the renormalization group invariant parameter $\hat{B}_K$ by [1]

$$\hat{B}_K = B_K(\mu) \left[\alpha_{\rm s}^{(3)}(\mu)\right]^{-2/9} \left[1 + \frac{\alpha_{\rm s}^{(3)}(\mu)}{4\pi} J_3\right], \tag{9}$$

$$\langle \overline{\mathrm{K}}^0 | Q(\Delta S = 2) | \mathrm{K}^0 \rangle \equiv \frac{8}{3} B_K(\mu) F_K^2 M_K^2 \tag{10}$$

and using (8) and (1) one finds

$$M_{12} = \frac{G_{\mathrm{F}}^2}{12\pi^2} F_K^2 \hat{B}_K M_K M_{\mathrm{W}}^2 \left[\lambda_c^{*2} \eta_1 S_0(x_c) + \lambda_t^{*2} \eta_2 S_0(x_t) + 2\lambda_c^* \lambda_t^* \eta_3 S_0(x_c, x_t) \right], \tag{11}$$

where $F_K = 160$ MeV is the K-meson decay constant and M_K the K-meson mass.

To proceed further we neglect the last term in (7) as in the standard CKM phase convention it constitutes at most a $2\,\%$ correction to ε_K. This is justified in view of other uncertainties, in particular those connected with $\hat{B}_K$. Inserting (11) into (7) we find

$$\varepsilon_K = C_\varepsilon \hat{B}_K \mathrm{Im}\lambda_t \left\{ \mathrm{Re}\lambda_c \left[\eta_1 S_0(x_c) - \eta_3 S_0(x_c, x_t) \right] - \mathrm{Re}\lambda_t \eta_2 S_0(x_t) \right\} \exp(i\pi/4), \tag{12}$$

where we have used the unitarity relation $\mathrm{Im}\lambda_c^* = \mathrm{Im}\lambda_t$ and have neglected $\mathrm{Re}\lambda_t/\mathrm{Re}\lambda_c = \mathcal{O}(\lambda^4)$ in evaluating $\mathrm{Im}(\lambda_c^* \lambda_t^*)$. The numerical constant C_ε is given by

$$C_\varepsilon = \frac{G_{\mathrm{F}}^2 F_K^2 M_K M_{\mathrm{W}}^2}{6\sqrt{2}\pi^2 \Delta M_K} = 3.837 \cdot 10^4 . \tag{13}$$

To this end we have used the experimental value of $\Delta M_K = 3.837 \cdot 10^{-15}$ GeV and $M_{\mathrm{W}} = 80.4$ GeV.

The main uncertainty in (12) resides in the parameter $\hat{B}_K$. The present status of $\hat{B}_K$ is discussed in Sec. 2.2. Here we note only that when $\hat{B}_K > 0$, as found by all non-perturbative methods, the formula (12) combined with the experimental value for ε_K implies $0 < \delta < \pi$ in the standard parametrization or equivalently $\bar{\eta} > 0$ in the Wolfenstein parametrization.

1.2. B sector: basic formulae for $\Delta M_{d,s}$ oscillation frequencies

The strengths of the $\mathrm{B}_{d,s}^0 - \overline{\mathrm{B}}_{d,s}^0$ mixings are described by the mass differences

$$\Delta M_{d,s} = M_H^{d,s} - M_L^{d,s} \tag{14}$$

where the subscripts H and L denote the heavy and light mass eigenstates respectively. The long distance contributions are estimated to be very small, in contrast to the situation for ΔM_K, and $\Delta M_{d,s}$ are very well approximated by the relevant box diagrams. Moreover, since $m_{u,c} \ll m_t$ only the top sector can contribute significantly to $\Delta M_{d,s}$. The charm and mixed top-charm contributions are entirely negligible.

$\Delta M_{d,s}$ can be expressed in terms of the off-diagonal element in the neutral B-meson mass matrix as follows

$$\Delta M_q = 2|M_{12}^{(q)}|, \qquad q = d, s \tag{15}$$

with M_{12} given by a formula analogous to (8)

$$2M_{B_q}|M_{12}^{(q)}| = |\langle \overline{\mathrm{B}}_q^0 | \mathcal{H}_{\mathrm{eff}}(\Delta B = 2) | \mathrm{B}_q^0 \rangle|. \tag{16}$$

In the case of $\mathrm{B}_d^0 - \overline{\mathrm{B}}_d^0$ mixing

$$\begin{aligned} \mathcal{H}_{\mathrm{eff}}^{\Delta B=2} &= \frac{G_{\mathrm{F}}^2}{16\pi^2} M_W^2 \left(V_{tb}^* V_{td}\right)^2 \eta_B S_0(x_t) \times \\ &\quad \times \left[\alpha_s^{(5)}(\mu_b)\right]^{-6/23} \left[1 + \frac{\alpha_s^{(5)}(\mu_b)}{4\pi} J_5\right] Q(\Delta B = 2) + h.c. \end{aligned} \tag{17}$$

Here $\mu_b = \mathcal{O}(m_b)$, $J_5 = 5165/3174 = 1.627$ in the NDR scheme [1],

$$Q(\Delta B = 2) = (\bar{b}\gamma_\mu(1-\gamma_5)d)(\bar{b}\gamma^\mu(1-\gamma_5)d) \tag{18}$$

and

$$\eta_B = 0.55 \pm 0.01 \tag{19}$$

summarizes the NLO QCD corrections [1,8]. In the case of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ mixing one should simply replace $d \to s$ in (17) and (18) with all other quantities and numerical values unchanged. Again $m_t \equiv m_t(m_t)$.

Defining the renormalization group invariant parameters $\hat{B}_{B_q}$ in analogy to (9) and (10)

$$\hat{B}_{B_q} = B_{B_q}(\mu)\left[\alpha_{\mathrm{s}}^{(5)}(\mu)\right]^{-6/23}\left[1 + \frac{\alpha_{\mathrm{s}}^{(5)}(\mu)}{4\pi}J_5\right], \tag{20}$$

$$\langle \overline{\mathrm{B}}_q^0 | Q(\Delta B = 2) | \mathrm{B}_q^0 \rangle \equiv \frac{8}{3} B_{B_q}(\mu) F_{B_q}^2 M_{B_q}^2 \tag{21}$$

one finds using (17)

$$\Delta M_q = \frac{G_{\mathrm{F}}^2}{6\pi^2}\eta_B M_{B_q}(\hat{B}_{B_q} F_{B_q}^2) M_{\mathrm{W}}^2 S_0(x_t)|V_{tq}|^2, \tag{22}$$

where F_{B_q} is the B_q-meson decay constant. This implies two approximate but rather accurate formulae

$$\Delta M_d = 0.50/\mathrm{ps} \cdot \left[\frac{\sqrt{\hat{B}_{B_d}} F_{B_d}}{230\mathrm{MeV}}\right]^2 \left[\frac{\overline{m}_{\mathrm{t}}(m_{\mathrm{t}})}{167\mathrm{GeV}}\right]^{1.52} \left[\frac{|V_{td}|}{7.8 \cdot 10^{-3}}\right]^2 \left[\frac{\eta_B}{0.55}\right] \tag{23}$$

and

$$\Delta M_s = 17.2/\mathrm{ps} \cdot \left[\frac{\sqrt{\hat{B}_{B_s}} F_{B_s}}{260\mathrm{MeV}}\right]^2 \left[\frac{\overline{m}_{\mathrm{t}}(m_{\mathrm{t}})}{167\mathrm{GeV}}\right]^{1.52} \left[\frac{|V_{ts}|}{0.040}\right]^2 \left[\frac{\eta_B}{0.55}\right]. \tag{24}$$

The main uncertainty here stems from the parameters $F_{B_{d,s}}$ and $\hat{B}_{B_{d,s}}$. The most recent lattice and QCD sum rule results are summarized in Sec. 2.1.

1.3. Basic formulae for B oscillation probabilities

The probability $\mathcal{P}$ for a B_q^0 meson ($q = d, s$) produced at time $t = 0$ to decay as B_q^0 at proper time t is given as

$$\mathcal{P}(\mathrm{B}_q^0 \to \mathrm{B}_q^0) = \frac{1}{2}\Gamma_q\, e^{-\Gamma_q t}\left[\cosh(\frac{\Delta\Gamma_q}{2}t) \; + \; \cos(\Delta M_q t)\right]. \tag{25}$$

Here we neglect effects from CP violation, while $\Gamma_q = \frac{\Gamma_q^H + \Gamma_q^L}{2}$, $\Delta\Gamma_q = \Gamma_q^H - \Gamma_q^L$ and ΔM_q is defined in Eq. (14). The Standard Model predicts $\Delta\Gamma_q \ll \Delta M_q$. Neglecting a possible lifetime difference between the heavy and light mass eigenstates of the B_q^0, the above expression simplifies to:

$$\mathcal{P}_{B_q^0}^{\mathrm{unmix}} = \mathcal{P}(\mathrm{B}_q^0 \to \mathrm{B}_q^0) \; = \; \frac{1}{2}\Gamma_q\, e^{-\Gamma_q t}\left[1 + \cos(\Delta M_q t)\right] \tag{26}$$

Similarly, the probability for the B_q^0 to decay as $\overline{\mathrm{B}}_q^0$ is given by

$$\mathcal{P}_{B_q^0}^{\mathrm{mix}} = \mathcal{P}(\mathrm{B}_q^0 \to \overline{\mathrm{B}}_q^0) \; = \; \frac{1}{2}\Gamma_q\, e^{-\Gamma_q t}[1 - \cos(\Delta M_q t)]. \tag{27}$$

Thus, a measurement of the oscillation frequency gives a direct measurement of the mass difference between the two physical B meson states*.

Figure 4.2 shows the time evolution of $B^0 - \overline{\mathrm{B}}^0$ oscillations displaying the unmixed (solid) and mixed (dashed) contributions for two different oscillation frequencies ΔM. The sum of $\mathcal{P}^{\mathrm{mix}}$ and $\mathcal{P}^{\mathrm{unmix}}$ is just the exponential particle decay $\Gamma_q\, \mathrm{e}^{-\Gamma_q t}$ and is shown by the dotted line in Fig. 4.2.

*ΔM_q is usually given in ps^{-1}, where 1 ps^{-1} corresponds to 6.58 10^{-4}eV.

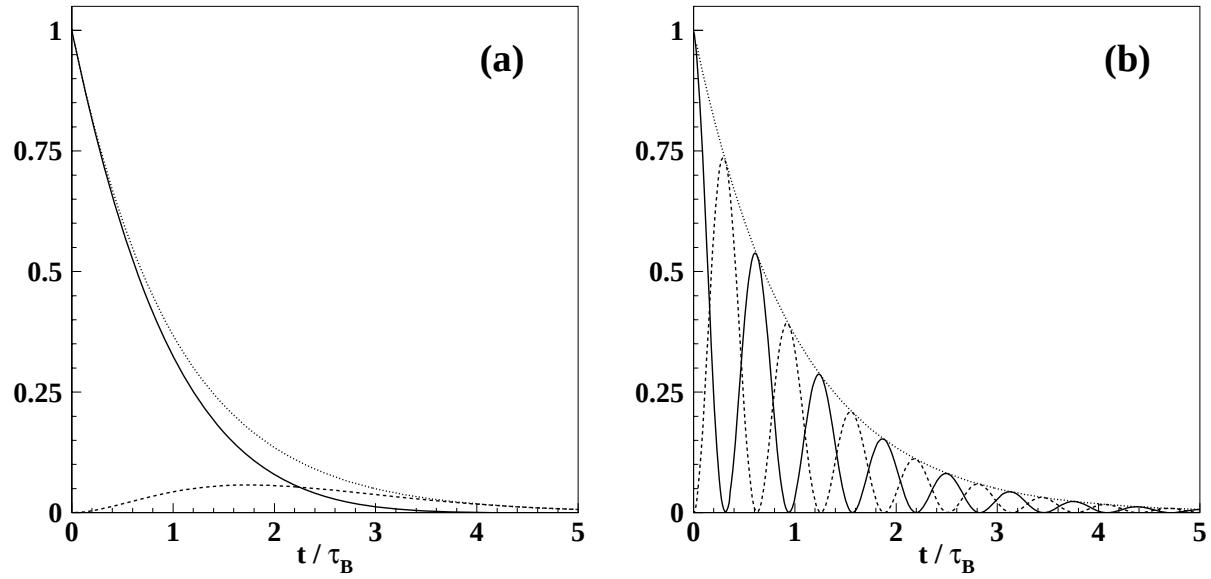

Fig. 4.2: *Time evolution of* B^0–$\overline{\mathrm{B}}^0$ *oscillations displaying the unmixed (solid) and the mixed (dashed) contribution as well as the sum of the two (dotted) for (a) slow and (b) fast oscillation frequencies* ΔM_q.

The integral of the probability $\mathcal{P}^{\rm mix}_{B^0_q}$ defined in Eq. (27) gives the mixing parameter:

$$\chi_q = \frac{x_q^2}{2\,(1+x_q^2)} \qquad \text{with} \quad x_q = \Delta M_q\,\tau_{B_q}, \tag{28}$$

where the lifetime $\tau_{B_q} = 1/\Gamma_q$.

2. Theoretical issues

2.1. Non-perturbative parameters for B meson mixing

From the discussion in Sec. 1.2. above, the main uncertainty in determining $|V_{td}|$ from ΔM_d comes from the factor $F_{B_d}\sqrt{\hat{B}_{B_d}}$ in Eq. 23. In the standard analysis of the Unitarity Triangle (see Chapter 5), ΔM_s is used in a ratio with ΔM_d, so that the important quantity is ξ, that is crucial for the determination of $|V_{td}|/|V_{ts}|$:

$$\frac{|V_{td}|}{|V_{ts}|} = \xi \sqrt{\frac{M_{\mathrm{B}_s}}{M_{\mathrm{B}_d}}}\sqrt{\frac{\Delta M_d}{\Delta M_s}}, \qquad \xi = \frac{F_{B_s}\sqrt{\hat{B}_{B_s}}}{F_{B_d}\sqrt{\hat{B}_{B_d}}}. \tag{29}$$

Although the quantities $F_{B_q}\sqrt{\hat{B}_{B_q}}$ for $q = d, s$ are needed for UT fits, it is common to find F_{B_q} and $\hat{B}_{B_q}$ separately.

2.1.1. F_{B_q} and ξ from lattice QCD

Lattice calculations are based on a first-principles evaluation of the path integral for QCD on a discrete space-time lattice. They have statistical errors arising from the stochastic (Monte Carlo) techniques used to evaluate the integral. They also have systematic errors from discretization effects, finite volume effects, the treatment of heavy quarks, chiral extrapolations and quenching (or partial quenching). We now briefly discuss these different sources of error.

Statistical, discretization and finite volume errors can all be addressed by brute-force improvement of numerical simulations. We can also use improved discretization procedures (to reduce discretization effects at a given lattice spacing) and understand (and even make use of) the finite volume effects.

Lattice results need to be matched either directly to physical quantities, or perhaps to quantities defined in some continuum renormalization scheme. On the lattice side this can be done using lattice perturbation theory, but with the development of non-perturbative renormalization methods, the uncertainty from the lattice can be systematically reduced. For a physical quantity (such as the decay constant F_{B_q}) this is the end of the story. If matching is made to a quantity in a continuum scheme (such as B_{B_q} in $\overline{\rm MS}$), the remaining uncertainty comes from the *continuum* perturbation theory: see for example the discussion in [9].

There are a number of ways to treat the heavy b-quark on the lattice. Results for $\rm B^0$–$\overline{\rm B}^0$ mixing obtained using different approaches broadly agree, suggesting that the heavy quark mass dependence is under control.

This leaves chiral extrapolations and quenching to consider. We will start with quenching. Recall that the QCD path integral is over both gauge and fermion fields. However, since the fermions appear quadratically in the action, the fermion integral can be done exactly to leave a determinant (actually a determinant for each flavour of quark). The calculation of the determinant is extremely intensive numerically, so the so-called *quenched approximation* replaces it with a constant, together with a shift in the bare couplings. This is not a controlled approximation, but today more and more lattice simulations are being done including the determinant for at least some of the quarks. The first dynamical quark algorithms produced sea quarks in degenerate pairs (in order to get a positive weight function for the Monte Carlo generation of the gauge field ensemble) and two-flavour ($N_f = 2$) dynamical simulations are still the most commonly encountered. However, methods are being developed to cope with single flavours of dynamical quark and $N_f = 2 + 1$ simulation results, with two degenerate light flavours and one heavy flavour, are beginning to appear, although there are still questions about the validity of some steps in the algorithm.

Each quark whose determinant is evaluated is labeled as a 'dynamical' or 'sea' quark in lattice parlance. A typical lattice calculation of a hadronic correlation function (from which masses and/or matrix elements may be extracted) involves an average over an ensemble of gauge fields of a combination of quark propagators. These propagators are evaluated on the background of each gauge field in the ensemble by means of a matrix inversion. The set of masses used for the propagators define the 'valence' masses of the simulation, which may or may not be the same as the dynamical masses which were incorporated (via determinant factors) when generating the gauge field ensemble. Usually the valence and sea masses are different and we talk of a 'partially quenched' calculation.

Results for F_B from quenched calculations have remained stable for a number of years. Numerical simulations using two flavours of dynamical quarks show an increase in F_B compared to quenched results. The latest developments have seen the first 3-flavour dynamical results [10,11], where two flavours are 'light' and one is heavier, around the strange quark mass. For the future, the development of more realistic dynamical simulations will continue.

Another important (and related) issue is that of chiral extrapolations, the subject of a panel discussion [12] at the Lattice 2002 conference. It is difficult to simulate realistically light (valence or sea) quarks, so that calculations of F_{B_q}, say, are made for a a set of (valence) quark masses m_q, typically in a range from about $m_s/2$ to $2m_s$ and the results are interpolated or extrapolated as required. Likewise, in partially quenched calculations, results from simulations with a range of sea quark masses need to be extrapolated. The control of these extrapolations is a serious issue for UT fits because of their effect on the final values of F_{B_d} and F_{B_s} and hence on the impact of the $\Delta M_s/\Delta M_d$ constraint. As far back as late 1994 Booth noted the striking difference between the quenched and QCD chiral logarithms [13] and posted a warning that F_{B_s}/F_{B_d} in QCD would be larger than in the quenched approximation. Recently, this issue has attracted much more attention [14–18].

Consider an idealized lattice calculation of the decay constant of a heavy-light pseudoscalar meson with valence content $Q\bar{q}$, where Q is the heavy quark and $\bar{q}$ a light quark. Imagine that the simulation is performed either with or without the presence of N_f flavours of (degenerate) sea quarks f and let ΔF_{B_q} be the correction to F_{B_q} depending on the mass(es) of the valence (q) and sea (f) quarks. With no sea quark effects included, the calculation is quenched. When $m_q \neq m_f$ the calculation is partially quenched and when $m_q = m_f$ it is QCD(-like). The dependence of ΔF_{B_q} on the valence and sea quark masses can be calculated in quenched (Q), partially quenched (PQ) or ordinary chiral perturbation theory, and shows up as dependence on the masses m_{qq}, m_{qf} and m_{ff} of pseudoscalar mesons made from the corresponding quarks [19]. The expressions are as follows

$$(\Delta F_{B_q})^{\rm QQCD} = \frac{1}{(4\pi f)^2}(X m_{qq}^2 + Y m_0^2)\ln\Big(\frac{m_{qq}^2}{\Lambda^2}\Big) \tag{30}$$

$$(\Delta F_{B_q})^{\rm PQQCD} = -\frac{(1+3g^2)}{(4\pi f)^2}\left[\frac{N_f}{2} m_{qf}^2 \ln\Big(\frac{m_{qf}^2}{\Lambda^2}\Big) + \frac{(m_{ff}^2 - 2m_{qq}^2)}{2N_f}\ln\Big(\frac{m_{qq}^2}{\Lambda^2}\Big)\right] \tag{31}$$

$$(\Delta F_{B_q})^{\rm QCD} = -\frac{(1+3g^2)}{(4\pi f)^2}\Big(\frac{N_f}{2} - \frac{1}{2N_f}\Big) m_{qq}^2 \ln\Big(\frac{m_{qq}^2}{\Lambda^2}\Big) \tag{32}$$

with $m_{qf}^2 = (m_{qq}^2 + m_{ff}^2)/2$ (at this order of calculation). In the factor $1/(4\pi f)^2$, f is equal to the common light pseudoscalar meson decay constant at leading order, while X, Y and m_0 are also built from coefficients of the effective Lagrangian. The dependence on the ultraviolet cutoff Λ is canceled by that of 'analytic terms' not shown here. The coupling g comes from the leading interaction term in the heavy meson chiral Lagrangian (see the textbook by Manohar and Wise [20] for details and original references) and fixes the $\mathrm{B}^*\mathrm{B}\pi$ coupling in the limit $M_B \to \infty$ by

$$g_{B^*B\pi} = \frac{2\,g\,M_B}{f} \tag{33}$$

where

$$\langle \mathrm{B}^+(p)\pi^-(q)|\mathrm{B}^*(\epsilon,p')\rangle = g_{B^*B\pi}\epsilon\cdot q. \tag{34}$$

The decay $\mathrm{B}^* \to \mathrm{B}\pi$ is not kinematically allowed, but g can be estimated using CLEO results [21] for $\mathrm{D}^* \to \mathrm{D}\pi$, or from a lattice QCD calculation of the matrix element of the light-quark axial current between B and B^* mesons [22] (or D and D^* [23]). The CLEO results lead to $g = 0.6$, consistent with the recent lattice calculation [23].

The expressions in Eqs. (30), (31) and (32) show that both the quenched and partially quenched 'chiral logarithms' diverge as the valence quark mass and hence m_{qq} vanishes while the sea quark mass is held fixed. In contrast, there are no divergences when the sea quark masses vanish with the valence masses held fixed. For the QCD-like case, things also remain finite as the joint valence and sea quark mass vanishes. The problem for lattice practitioners is how best to perform the chiral extrapolations from results calculated with sets of m_q and m_f values, particularly since it is very difficult to make the masses small enough to see the logarithmic dependence.

For F_{B_d} the situation is like the 'QCD' case above where the valence d quark in the B_d meson and (some of) the sea quarks are very light. For F_{B_s}, the valence mass is fixed at m_s and the sea quark masses are extrapolated to small values (more like the partially quenched case above). The JLQCD collaboration find [24] that these different extrapolations tend to decrease the value of F_{B_d} relative to F_{B_s}, and therefore increase ξ. However, a number of caveats must be kept in mind [25]. Although the data is consistent with the chiral logarithmic forms, all the data points are at masses beyond the region of strong variation in the logarithms. Moreover, at these larger masses, higher order terms in the chiral expansion may be required. Furthermore, in dynamical simulations the lattice spacing changes as the sea quark mass changes at fixed lattice coupling (β), so that care is needed not to interpret lattice-spacing

(and volume) dependence as sea-quark mass dependence. An added twist is that JLQCD find that their results for F_π are *not* consistent with the expected logarithmic behaviour.

The MILC collaboration have also estimated chiral logarithm effects as part of their extensive analysis of $N_f = 2$ simulations for heavy-light decay constants [17]. Their method is based on extrapolation of the ratio of the light-light to the heavy-light decay constant, where the chiral logarithmic terms cancel to a large extent. MILC's conclusion is that these effects do tend to increase the value of the ratio F_{B_s}/F_{B_d} and MILC ascribe a systematic error of $+0.04$ from chiral logarithms to a central value of 1.16 for F_{B_s}/F_{B_d}.

Kronfeld and Ryan (KR) [14] consider the ratios $\xi_f = F_{B_s}/F_{B_q}$ and $\xi_B = B_{B_s}/B_{B_q}$ as the mass of the quark q varies from the strange mass down to that of the light quarks u and d and match ChPT to lattice data for m_q not too far from m_s. Their analysis gives $\xi = 1.32(10)$. Another more recent phenomenological analysis (BFPZ) [18] supports the increase in ξ coming from chiral logarithms and leads to a consistent result $\xi = 1.22(7)$. This value is extracted using the double ratio

$$R = \frac{(F_{B_s}\sqrt{M_{B_s}})/(F_{B_d}\sqrt{M_{B_d}})}{F_K/F_\pi}. \tag{35}$$

An expression for R in leading order heavy meson and pion chiral perturbation theory (in full, 3-flavour QCD) is combined with the experimental ratio $(F_K/F_\pi)_{\rm expt} = 1.22(1)$ to extract F_{B_s}/F_{B_d}. Systematic error in both analyses arises from the uncertain values of parameters in the heavy meson and pion chiral Lagrangian, namely the coupling g in the leading interaction term, already encountered above, together with sums of coefficients of higher-order terms in the heavy meson chiral Lagrangian. In addition the analysis using R depends on L_5, the coefficient of a higher-order term in the pion chiral Lagrangian through its use of the ratio F_K/F_π.

In conclusion, lattice results for F_B can show significant light-quark mass dependence and more work is needed to understand to what extent this dependence is physical. At present a reasonable conservative view [25] is to allow a *decrease* of up to -10% in F_{B_d} with a negligible change in F_{B_s} as added systematic errors. These are included in the final estimates presented in Eq. (37).

A summary of lattice calculations for the decay constants, published after 1996, is given in Fig. 4.3 (taken from the review by Lellouch [25]), which shows results for F_{B_d} and the ratio F_{B_s}/F_{B_d}. The 'summary' numbers at the bottom of the plots give quenched averages for F_{B_d} and F_{B_s}/F_{B_d}, together with ratios of these quantities for $N_f = 2$ and $N_f = 0$:

$$\begin{aligned} F_{B_d}^{N_f=0} &= 178(20)\text{MeV} & \frac{F_{B_d}^{N_f=2}}{F_{B_d}^{N_f=0}} &= 1.09(6) \\ (F_{B_s}/F_{B_d})^{N_f=0} &= 1.14(3) & \frac{(F_{B_s}/F_{B_d})^{N_f=2}}{(F_{B_s}/F_{B_d})^{N_f=0}} &= 1.02(2) \end{aligned} \tag{36}$$

For the mixing parameter B_{B_q}, the situation with quenching and chiral extrapolation looks more favourable. Very little variation is observed between quenched ($N_f = 0$) and $N_f = 2$ results. The partially quenched chiral logarithm for B_{B_q} has a coefficient containing $1 - 3g^2 \simeq -0.1$ compared to $1 + 3g^2 \simeq 2.1$ in the F_{B_q} case (using $g = 0.6$ as discussed above) so the chiral extrapolation is better-controlled and leads to a small error in $\hat{B}_{B_s}/\hat{B}_{B_d}$ [14,25,16]. The heavy quark mass dependence is mild and different formulations agree at the physical point for B-mesons.

There is, however, an issue concerning lattice results for ξ which are normally quoted by combining results for F_B and $\hat{B}_B$. Of course, it is also possible to evaluate ξ directly from the ratio of $\Delta B = 2$ matrix elements. In this case ξ turns out to be larger, although with large errors [35,27]. Clearly the two procedures should give consistent answers, so this issue will need to be resolved.

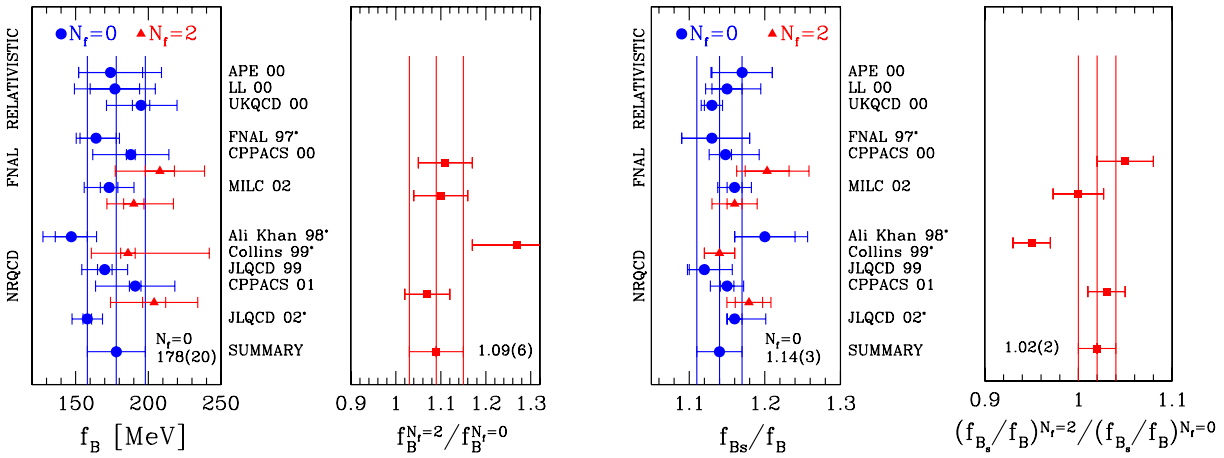

Fig. 4.3: *From left to right: lattice results published after 1996 for (a) F_{B_d} in quenched ($N_f = 0$) and two-flavour ($N_f = 2$) QCD, (b) the ratio $F_{B_d}^{N_F=2}/F_{B_d}^{N_f=0}$, (c) F_{B_s}/F_{B_d} in quenched ($N_f = 0$) and two-flavour ($N_f = 2$) QCD, (d) the ratio $(F_{B_s}/F_{B_d})^{N_f=2}/(F_{B_s}/F_{B_d})^{N_f=0}$. The results are grouped according to the formulation used to treat the heavy quark and the references are: APE 00 [26], LL 00 [27], UKQCD 00 [28], FNAL 97 [29], CPPACS 00 [30], MILC 02 [17], Ali Khan 98 [31], Collins 99 [32], JLQCD 99 [33], CPPACS 01 [34] and JLQCD 02 [16]. Figs. taken from [25].*

2.1.2. Summary on F_{B_q} and ξ from the lattice QCD

Using the quenched averages as a starting point together with the ratios of $N_f = 2$ to $N_f = 0$ results allows an extrapolation to $N_f = 3$ [25]. An additional systematic error equal to the shift from 2 to 3 flavours is added to account for the uncertainty in this procedure[‡]. This leads to:

$$\boxed{\begin{array}{lll} F_{B_d} = 203(27)(^{0}_{20})\text{MeV} & F_{B_s} = 238(31)\text{MeV} & \dfrac{F_{B_s}}{F_{B_d}} = 1.18(4)(^{12}_{0}) \\ \hat{B}_{B_d} = 1.34(12) & \hat{B}_{B_s} = 1.34(12) & \dfrac{\hat{B}_{B_s}}{\hat{B}_{B_d}} = 1.00(3) \\ F_{B_d}\sqrt{\hat{B}_{B_d}} = 235(33)(^{0}_{24})\text{MeV} & F_{B_s}\sqrt{\hat{B}_{B_s}} = 276(38)\text{MeV} & \xi = 1.18(4)(^{12}_{0}) \end{array}} \quad . \tag{37}$$

Here, the last, asymmetric, error, where present, is due the uncertainty in the chiral extrapolation discussed above. The first error combines statistical and all other systematic errors. In UT analyses, the value of ξ given above should be understood as

$$\xi = 1.24(4)(6) \tag{38}$$

and likewise for other quantities affected by this asymmetric error. Note that this does not apply for F_{B_s} and $\hat{B}_{B_s}$, for which the chiral logarithmic uncertainties appear small compared to other systematic errors. The result for ξ in Eq. (38) is consistent with the KR [14] and BFPZ [18] analyses mentioned above.

2.1.3. F_{B_d} and F_{B_s} from QCD sum rules

Within the framework of QCD sum rules [36,37], the decay constants F_{B_d} and F_{B_s} can be calculated by equating phenomenological and theoretical spectral functions for the pseudoscalar B_d and B_s mesons,

[‡]An alternative way to quote the final answer would be to use the $N_f = 2$ results extracted from Eq. (36) and add a systematic error for the extrapolation to $N_f = 3$. In this case, the final central value for F_{B_s}/F_{B_d} would be 1.16. The value of 1.18, however, is consistent with the latest preliminary MILC results for $N_f = 3$, which give $(F_{B_s}/F_{B_d})^{N_f=3} = 1.18(1)(^{4}_{1})$

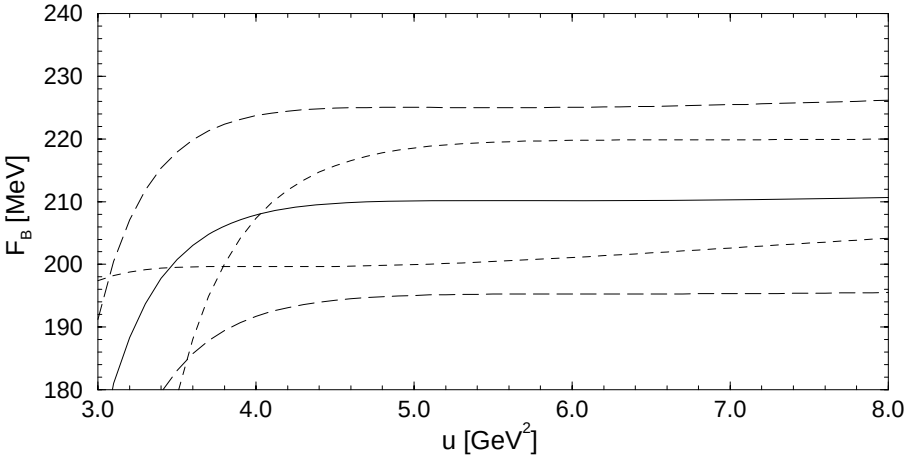

Fig. 4.4: F_{B_d} *as a function of the sum rule scale* u *for different sets of input parameters. Solid line: central values of Table 4.1; long-dashed line:* $m_b(m_b) = 4.16$ GeV *(upper line),* $m_b(m_b) = 4.26$ GeV *(lower line); dashed line:* $\mu_m = 3$ GeV *(lower line),* $\mu_m = 6$ GeV *(upper line).*

which leads to the relation [38–40]‡

$$M_B^4 F_{B_d}^2 = \int_0^{s_0} e^{(M_B^2 - s)/u} \rho(s)\, ds \tag{39}$$

for the B_d meson and analogously for B_s. Eq. (39) is the central relation for the sum rule analysis. The theoretical spectral function $\rho(s) \equiv \Im\Psi(s)/\pi$ can be obtained by calculating the two-point correlator of hadronic currents

$$\Psi(p^2) \equiv i \int dx\, e^{ipx}\, \langle 0|\, T\{\, j_5(x)\, j_5(0)^\dagger\}|0\rangle \tag{40}$$

in perturbative QCD, including corrections from the operator product expansion. For the B meson, the pseudoscalar current $j_5(x)$ takes the form

$$j_5(x) = (m_b + m_u)\, :\bar{u}(x)\, i\gamma_5 b(x):\ . \tag{41}$$

The parameter s_0 in Eq. (39) indicates the energy range up to which experimental knowledge of the phenomenological spectral function is available. This parameter will be further discussed below.

Substantial progress in determining the theoretical spectral function has been achieved very recently through a calculation of the perturbative three-loop order α_s^2 corrections [42,43]. These are important because the size of higher-order corrections depends on the renormalization scheme employed for the quark masses. As can be inferred from refs. [42,43], the α_s^2 term turns out to be of similar order to the leading contribution if pole quark masses are used, whereas good convergence of the perturbative series emerges for quark masses defined in the $\overline{\mathrm{MS}}$ scheme. Nevertheless, these scheme dependences influence only the theoretical uncertainties, since F_{B_d} and F_{B_s} are physical quantities which certainly should not depend on the quark mass definitions. Higher-dimensional operator corrections to the sum rule are known up to dimension six [39] and are also under good theoretical control.

Figure 4.4 shows numerical results for F_{B_d} of Ref. [39], plotted as a function of the sum rule scale u, after evaluating the sum rule of Eq. (39). Reliable values of F_{B_q} can be extracted from the sum rule if an energy region exists in which the physical quantity is only weakly dependent on u. For F_{B_d} this

‡A review of the procedure and further original references can be found in [41].

Parameter	Value	s_0	u_0	ΔF_{B_d}
$m_b(m_b)$	4.21 ± 0.08 GeV	$\begin{smallmatrix}32.8\\34.6\end{smallmatrix}$	$\begin{smallmatrix}6.5\\5.0\end{smallmatrix}$	∓ 24
μ_m	$3.0 - 6.0$ GeV	$\begin{smallmatrix}33.5\\34.4\end{smallmatrix}$	$\begin{smallmatrix}6.8\\4.0\end{smallmatrix}$	± 10
$\langle \bar{u}u \rangle (2\text{ GeV})$	$-(267 \pm 17\text{ MeV})^3$	$\begin{smallmatrix}33.9\\33.3\end{smallmatrix}$	$\begin{smallmatrix}5.7\\5.5\end{smallmatrix}$	± 6
$\mathcal{O}(\alpha_s^2)$	$\begin{smallmatrix}2\times\mathcal{O}(\alpha_s^2)\\ \text{no } \mathcal{O}(\alpha_s^2)\end{smallmatrix}$	33.6	5.6	± 2
$\alpha_s(M_Z)$	0.1185 ± 0.020	33.6	5.6	± 1

Table 4.1: *Values for the dominant input parameters, continuum thresholds* s_0 [GeV^2], *points of maximal stability* u_0 [GeV^2], *and corresponding uncertainties for* F_{B_d} [MeV].

Parameter	Value	s_0	u_0	ΔF_{B_s}
$m_b(m_b)$	4.21 ± 0.08 GeV	$\begin{smallmatrix}34.3\\36.9\end{smallmatrix}$	$\begin{smallmatrix}5.8\\4.6\end{smallmatrix}$	∓ 26
μ_m	$3.0 - 6.0$ GeV	$\begin{smallmatrix}35.2\\37.2\end{smallmatrix}$	$\begin{smallmatrix}6.2\\3.6\end{smallmatrix}$	$\begin{smallmatrix}+8\\-9\end{smallmatrix}$
$\langle \bar{s}s \rangle / \langle \bar{u}u \rangle$	0.8 ± 0.3	$\begin{smallmatrix}35.9\\35.2\end{smallmatrix}$	$\begin{smallmatrix}5.3\\4.7\end{smallmatrix}$	± 8
$\langle \bar{u}u \rangle (2\text{ GeV})$	$-(267 \pm 17\text{ MeV})^3$	$\begin{smallmatrix}35.7\\35.3\end{smallmatrix}$	$\begin{smallmatrix}5.2\\4.9\end{smallmatrix}$	$\begin{smallmatrix}+5\\-4\end{smallmatrix}$
$m_s(2\text{ GeV})$	100 ± 15 MeV	35.5	5.1	± 2
$\mathcal{O}(\alpha_s^2)$	$\begin{smallmatrix}2\times\mathcal{O}(\alpha_s^2)\\ \text{no } \mathcal{O}(\alpha_s^2)\end{smallmatrix}$	35.5	5.1	± 3
$\alpha_s(M_Z)$	0.1185 ± 0.020	35.5	5.1	± 1

Table 4.2: *Values for the dominant input parameters, continuum thresholds* s_0 [GeV^2], *points of maximal stability* u_0 [GeV^2], *and corresponding uncertainties for* F_{B_s} [MeV].

turns out to be the case in the range $4\text{ GeV}^2 \lesssim u \lesssim 6\text{ GeV}^2$. Averaging the results of refs. [38,39] in this energy range, one extracts the central results $F_{B_d} = 208$ MeV and $F_{B_s} = 242$ MeV.[§]

The dominant uncertainties in the sum rule determination of F_{B_d} and F_{B_s} arise from the strong dependence on the value of the bottom quark mass m_b and correspondingly on the scale μ_m at which the quark masses are renormalized. The ranges for the variation of these parameters and the corresponding variations of F_{B_d} and F_{B_s} have been collected in Tables 4.1 and 4.2 respectively. The reader should note that compared to Ref. [39], the error on $m_b(m_b)$ has been enlarged, in order to coincide with the value employed throughout this report, although the larger uncertainty should be considered very conservative. The Tables also list the values u_0 at which the sum rule displays optimal stability, as well as the parameters s_0 which can be determined consistently from an independent sum rule for the B_d and B_s meson masses. Additional smaller uncertainties are due to: variation of the strong coupling constant α_s; higher order QCD corrections; the value of the quark condensate $\langle \bar{u}u \rangle$ [44] which is the leading contribution from higher-dimensional operators; the strange condensate $\langle \bar{s}s \rangle$ and the strange quark mass m_s in the case of F_{B_s}. Ranges for these inputs together with the variations of F_{B_d} and F_{B_s} are also collected in Tables 4.1 and 4.2. For further details of the numerical analysis, the reader is referred to Ref. [39].

[§]Owing to the criticism put forward in Ref. [39], the result of Ref. [40] has not been included in the average, despite the apparent agreement for the numerical values.

Adding all errors for the various input parameters in quadrature, the final results for the B_d and B_s meson leptonic decay constants from QCD sum rules are:

$$F_{B_d} = 208 \pm 27 \text{ MeV} \qquad \text{and} \qquad F_{B_s} = 242 \pm 29 \text{ MeV}. \tag{42}$$

Owing to the strong sensitivity of these results on the bottom quark mass, one should note that for example using the very recent average $m_b(m_b) = 4.24$ GeV [45], the resulting values for F_{B_d} and F_{B_s} are lowered by almost 10 MeV.

2.1.4. B_{B_d} and B_{B_s} from QCD sum rules

The status of the determination of the hadronic B-parameters B_{B_d} and B_{B_s} from QCD sum rules is less satisfactory than for the decay constants. In principle, the B-parameters can be calculated from two different types of sum rules: namely three-point function sum rules with the insertion of two pseudoscalar currents and one four-quark operator [46,47], or two-point function sum rules with the insertion of two local four-quark operators [48,49]. However, both approaches are plagued with difficulties¶.

The first determinations of the hadronic B-parameters [46,47] employed three-point function sum rules and found a value of $B_{B_d}(m_b) = 0.95 \pm 0.10$, slightly lower than the factorization approximation which results in $B_{B_d} = 1$. The dominant non-factorizable contribution due to the gluon condensate turned out to be negative, thus lowering the B-parameter. However, the perturbative part was only considered at the leading order, and thus the scale and scheme dependences of B_{B_d} were not under control. Besides, the analytic structure of three-point function sum rules is more delicate than for two-point correlators, and therefore great care has to be taken to properly extract the quantity in question [41].

For the case of the two-point function sum rules, next-to-leading order QCD corrections have been calculated in Ref. [48], which provides better control over the renormalization dependence of B_B. This analysis resulted in $B_{B_d}(m_b) = 1.0 \pm 0.15$. However, here the phenomenological parametrization of the spectral function is more complicated, since contributions from intermediate states containing B* mesons have to be taken into account in addition to the B meson. Steps in this direction have recently been taken in Ref. [49] were the value $B_{B_d}(m_b) = 1.15 \pm 0.11$ was obtained, now indicating a positive correction.

Although averaging the results of the two approaches might appear problematic, we nevertheless decided to quote a common value for the B meson B-parameter from QCD sum rules:

$$B_{B_d}(m_b) = 1.10 \pm 0.15 \qquad \text{and} \qquad \hat{B}_{B_d} = 1.67 \pm 0.23, \tag{43}$$

which covers the outcome of both methods within the uncertainties. On the other hand, general agreement exists for the flavour dependence of the B-parameter. In all present sum rule approaches it was found to be negligible, thus yielding $B_{B_s}/B_{B_d} = 1$ to a good approximation.

2.2. K^0–$\overline{K}^0$ mixing: determination of B_K

2.2.1. B_K from lattice QCD

The most commonly used method to calculate the matrix element $\langle \overline{\mathrm{K}}^0 \mid Z\ (\bar{s}d)_{V-A}(\bar{s}d)_{V-A}(\mu) \mid \mathrm{K}^0 \rangle$ is to evaluate the three point correlation function shown in Fig. 4.5. This corresponds to creating a K^0 at some time t_1 using a zero-momentum source; allowing it to propagate for time $t_{\mathcal{O}} - t_1$ to isolate the lowest state; inserting the four-fermion operator at time $t_{\mathcal{O}}$ to convert the K^0 to a $\overline{\mathrm{K}}^0$; and finally allowing the $\overline{\mathrm{K}}^0$ to propagate for long time $t_2 - t_{\mathcal{O}}$. To cancel the K^0 ($\overline{\mathrm{K}}^0$) source normalization at times t_1 and t_2 and the time evolution factors $e^{-E_K t}$ for times $t_2 - t_{\mathcal{O}}$ and $t_{\mathcal{O}} - t_1$ it is customary to divide this three-point function by the product of two 2-point functions as shown in Fig 1. If, in the 2-point functions, the

¶For a different approach see also Ref. [50].

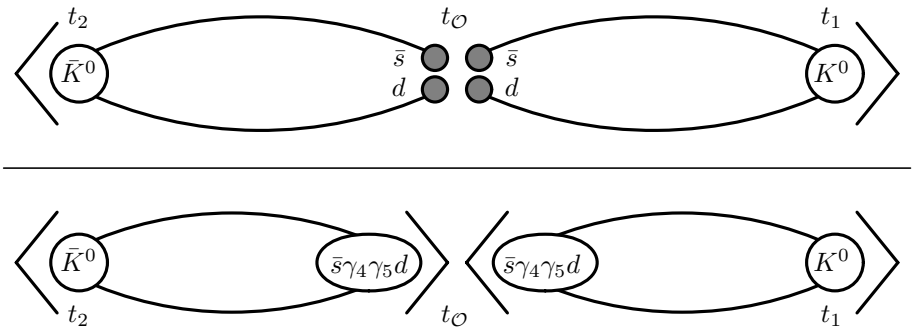

Fig. 4.5: *Ratio of lattice correlation functions used to calculate B_K.*

bilinear operator used to annihilate (create) the K^0 ($\overline{\mathrm{K}}^0$) at time $t_{\mathcal{O}}$ is the axial density $\bar{s}\gamma_4\gamma_5 d$, then the ratio of the 3-point correlation function to the two 2-point functions is $(8/3)B_K$.

B_K is defined to be the value of the matrix element at the physical kaon and normalized by the Vacuum Saturation Approximation value $8/3M_K^2F_K^2$

$$\langle \mathrm{K}^0 \mid Z\ (\bar{s}d)_{V-A}(\bar{s}d)_{V-A}(\mu) \mid \overline{\mathrm{K}}^0 \rangle \quad = \quad (8/3)B_K M_K^2 F_K^2\,.$$

The earliest calculations of B_K were done using Wilson fermions and showed significant deviations from this behaviour. It was soon recognized that these lattice artifacts are due to the explicit breaking of chiral symmetry in the Wilson formulation [51–55]. Until 1998, the only formulation that preserved sufficient chiral symmetry to give the right chiral behaviour was Staggered fermions. First calculations using this approach in 1989 gave the quenched estimate $B_K(\mathrm{NDR}, 2\mathrm{GeV}) = 0.70 \pm 0.01 \pm 0.03$. In hindsight, the error estimates were highly optimistic, however, the central value was only 10% off the current best estimate, and most of this difference was due to the unresolved $O(a^2)$ discretization errors.

In 1997, the staggered collaboration refined its calculation and obtained $0.62(2)(2)$ [56], again the error estimate was optimistic as a number of systematic effects were not fully included. The state-of-the-art quenched calculation using Staggered fermions was done by the JLQCD collaboration in 1997 and gave $B_K(2\mathrm{GeV}) = 0.63 \pm 0.04$ [57]. This estimate was obtained using six values of the lattice spacing between 0.15 and 0.04 fermi, thus allowing much better control over the continuum extrapolation as shown in Fig. 4.6 along with other published results. This is still the benchmark against which all results are evaluated and is the value exported to phenomenologists. This result has three limitations: (i) It is in the quenched approximation. (ii) All quenched calculations use kaons composed of two quarks of roughly half the "strange" quark mass and the final value is obtained by interpolation to a kaon made up of $(m_s/2, m_s/2)$ instead of the physical point (m_s, m_d). Thus, SU(3) breaking effects ($m_s \neq m_d$) have not been incorporated. (iii) There are large $O(a^2)$ discretization artifacts, both for a given transcription of the $\Delta S = 2$ operator on the lattice and for different transcriptions at a given value of the lattice spacing, so extrapolation to the continuum limit is not as robust as one would like. These limitations are discussed after a brief summary of the recent work.

In the last four years a number of new methods have been developed and the corresponding results are summarized in Table 4.3.

- The Rome collaboration has shown that the correct chiral behaviour can be obtained using $O(a)$ improved Wilson fermions provided non-perturbative renormalization constants are used. Their latest results, with two different "operators", are $B_K(2\mathrm{GeV}) = 0.63(10)$ and $0.70(12)$ [58]. These, while demonstrating the efficacy of this method, do not supplant the staggered result, as the continuum extrapolation is based on only three points and the data have larger errors. The discretization errors can be characterized as $B_K(a) = B_K(1 + a\Lambda)$ with $\Lambda \approx 400$MeV and are similar in magnitude to those with staggered fermions at $1/a = 2$ GeV, as are the differences in

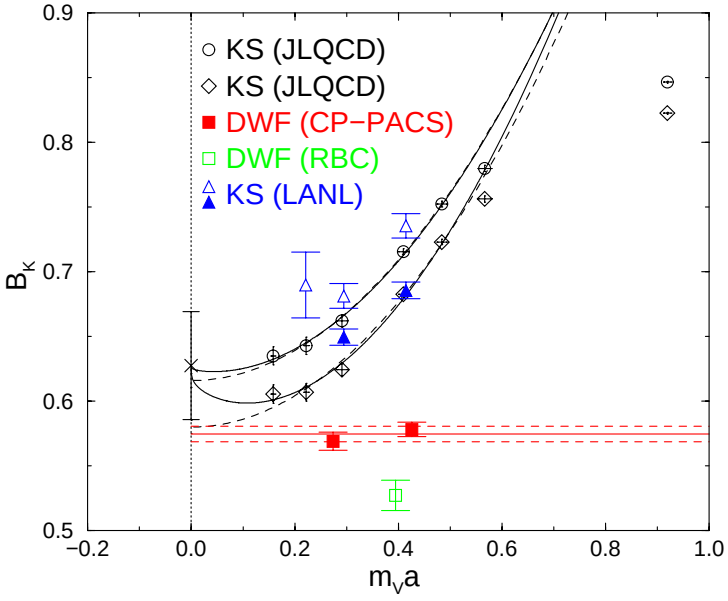

Fig. 4.6: *Published estimates of B_K with fermion formulations that respect chiral symmetry. All results are in the quenched approximation.*

Collaboration	year	$B_K(2\text{GeV})$	Formulation	Renormalization	a^{-1} (GeV)
Staggered [56]	1997	0.62(2)(2)	staggered	1-loop	∞
JLQCD [57]	1997	0.63(4)	staggered	1-loop	∞
Rome [58]	2002	0.63(10)	Improved Wilson	NP	∞
Rome [58]	2002	0.70(12)	Improved Wilson	NP	∞
CP-PACS [59]	2001	0.58(1)	Domain Wall	1-loop	1.8 GeV
CP-PACS [59]	2001	0.57(1)	Domain Wall	1-loop	2.8 GeV
RBC [60]	2002	0.53(1)	Domain Wall	NP	1.9 GeV
DeGrand [61]	2002	0.66(3)	Overlap	1-loop	1.6 GeV
DeGrand [61]	2002	0.66(4)	Overlap	1-loop	2.2 GeV
GGHLR [62]	2002	0.61(7)	Overlap	NP	2.1 GeV

Table 4.3: *Quenched estimates for B_K evaluated in the NDR scheme at* 2GeV. *The fermion formulation used in the calculation, the method used for renormalizing the operators, and the lattice scale at which the calculation was done are also given. NP indicates non-perturbative renormalization using the RI/MOM scheme and $a^{-1} = \infty$ implies that the quoted result is after a continuum extrapolation.*

estimates with using different operators. In the staggered formulation, the artifacts are, however, $O(a^2\Lambda^2)$ and $O(\alpha_s^2)$ and the data suggest an unexpectedly large $\Lambda \sim 900$MeV.

- Four collaborations have new results using domain wall and overlap fermions as shown in Table 4.3 [63,60,59,61,62]. Both formulations have built in chiral symmetry at finite a and $O(a)$ improvement. Each of these collaborations have used slightly different methodology, so they cannot be compared head on, or combined to do a continuum extrapolation. Thus, the results are quoted with reference to the lattice spacing at which the calculation was done. The differences reflect $O(a^2)$ (and $O(\alpha_s^2)$ in cases where perturbative renormalization constants have been used) artifacts.
- Calculations are in progress [64] using another method with good chiral behaviour, twisted mass QCD.

Deriving an estimate for the physical $\hat{B}_K$, starting from the current best quenched lattice estimate, the JLQCD staggered result $B_K(2\mathrm{GeV}) = 0.63(4)$, requires consideration of the following issues.

- The $O(a^2)$ errors in the staggered formulation are large. Nevertheless, the error 0.04 obtained by the JLQCD collaboration on including both $O(a^2)$ and $O(\alpha_s^2)$ terms in the extrapolation is a reasonable 1σ estimate of both the statistical and the extrapolation to continuum limit errors.
- A choice for α_s and the number of flavours in the perturbative expression has to be made to convert $B_K \to \hat{B}_K$. It turns out that the result is insensitive to whether one uses quenched or full QCD values. Using the 2-loop expression, the result for the central value is $\hat{B}_K = 0.86(6)$.
- An estimate of the systematic uncertainty associated with the quenched approximation and SU(3) breaking. Preliminary numerical estimates suggest that dynamical quarks would increase the value by about 5% [65,66]. Sharpe estimates, using ChPT, that unquenching would increase B_K by 1.05 ± 0.15, and SU(3) breaking effects would also increase it by 1.05 ± 0.05 [67]. This analysis of systematic errors is not robust and, furthermore, the two uncertainties are not totally independent. So one can take an aggressive and a conservative approach when quoting the final result for $\hat{B}_K$. In the aggressive approach, the error estimate is given by combining in quadrature the offset of the central values with respect to unity. This gives a 7% uncertainty and

$$\hat{B}_K = 0.86 \pm 0.06 \pm 0.06\,. \tag{44}$$

  In the conservative approach, advocated by Sharpe [67], one combines the uncertainty in quadrature to get a 16% uncertainty. The final result in this case is

$$\boxed{\hat{B}_K = 0.86 \pm 0.06 \pm 0.14} \tag{45}$$

Given the lack of a robust determination of the systematic error, it is important to decide how to fold these errors in a phenomenological analysis. One recommendation is to assume a flat distribution for the systematic error and add to it a Gaussian distribution with $\sigma = 0.06$ on either end, and do a separate analysis for the aggressive and conservative estimates. In other words, a flat distribution between 0.72 and 1.0 for a conservative estimate of $\hat{B}_K$ (or from 0.80 to 0.92 for the aggressive estimate) to account for systematic errors due to quenching and SU(3) breaking. Since this is the largest uncertainty, current calculations are focused on reducing it.

Finally, the reasons why the quenched lattice estimate of B_K has been stable over time and considered reliable within the error estimates quoted above are worth reemphasizing:

- The numerical signal is clean and accurate results are obtained with a statistical sample of even 50 decorrelated lattices.
- Finite size effects for quark masses $\geq m_s/2$ are insignificant compared to statistical errors once the quenched lattices are larger the 2 fermi.
- In lattice formulations with chiral symmetry, the renormalization constant connecting the lattice and continuum schemes is small ($< 15\%$), and reasonably well estimated by one-loop perturbation theory.

- For degenerate quarks, the chiral expansion for the matrix element has no singular quenched logarithms (they cancel between the AA and VV terms) that produce large artifacts at small quark masses in observables like M_π^2, f_π, etc. Also, the chiral expansions have the same form in the quenched and full theories [68–71].
- ChPT estimates of quenching and SU(3) breaking systematic errors are at the 7–16% level [70,65,66].

2.2.2. B_K from non-lattice approaches

The parameter B_K can also be calculated using other non-perturbative approaches to QCD, like QCD sum rules, the large-N_c expansion or the chiral quark model. As for the parameter B_B in the B-meson system, B_K can be obtained from sum rules by considering two-point [72–74] or three-point [75,76] correlation functions. However, both methods suffer from the same inadequacies as in the case of B_B. For the two-point function sum rule, the phenomenological spectral function is difficult to parametrise reliably, whereas for the three-point function sum rule no next-to-leading order QCD corrections are available and thus a proper matching with the Wilson coefficient function is at present not possible. For these reasons, we shall concentrate below on existing results in the large-N_c expansion [77–80], which in our opinion are developed furthest. After commenting on the large-N_c approach in more detail, the calculation of B_K within the chiral quark model [81] will also be briefly discussed.

Calculations of weak hadronic matrix elements in the framework of the large-N_c expansion were developed by Bardeen, Buras and Gérard in the nineteen-eighties. For B_K, at the next-to-leading order in $1/N_c$, this method resulted in $B_K = 0.7 \pm 0.1$ [77], to be compared with $B_K = 0.75$ in the strict large-N_c limit. However, at that time the next-to-leading order correction to the Wilson coefficient function [82] was not available, and anyhow it is debatable whether the result of [77] can be properly matched to the short distance coefficient. The proper matching of the scale and scheme dependencies in matrix elements as well as Wilson coefficients is, however, a crucial aspect for all approaches to weak hadronic matrix elements.

In the approach of [78] a significant dependence on the matching scale is still present, resulting in sizable uncertainties for B_K. Explicit cancellation of scale and scheme dependences was demonstrated in Ref. [79] within the chiral limit, and, to a lesser extent in Ref. [80], also for a physical strange quark. The main ingredients in the approaches of [79,80] are: the large-N_c expansion; chiral perturbation theory to control the low-energy end of the Green function required for the calculation of the matrix elements; the operator product expansion to control the higher-energy region of the Green function above roughly 1 GeV; a model which connects the low- and high energy regimes. To this end, in [79] the relevant Green function was saturated by the lowest lying vector meson, the ρ, whereas in [80] the extended Nambu-Jona-Lasinio model was applied which, however, does not display the correct QCD high-energy behaviour. The dependence on these models constitutes the dominant uncertainty for the latter approaches.

In the chiral limit, the findings $\hat{B}_K = 0.38 \pm 0.11$ [79] as well as $\hat{B}_K = 0.32 \pm 0.13$ [80] are in very good agreement with the current algebra result $\hat{B}_K = 0.33$ [83], obtained by relating $\hat{B}_K$ to the $K^+ \to \pi^+\pi^0$ decay rate. In fact, this agreement could be interpreted as a successful description of the $K^+ \to \pi^+\pi^0$ decay from large-N_c. The authors of Ref. [80] have also extended their calculation beyond the chiral limit with the result $\hat{B}_K = 0.77 \pm 0.07$. The smaller error compared to the chiral limit case is due to a reduced model dependence for a physical strange quark. However, as is obvious from these results, the chiral corrections amount to more than 100%, and it remains to be seen whether $\hat{B}_K$ of [80] incorporates all such corrections. Nevertheless, it is interesting to observe that the final result of Ref. [80] is again very close to the strict large-N_c prediction, and is also in good agreement with the average from lattice QCD quoted above.

An independent approach to hadronic matrix elements and to B_K in particular is the chiral quark model [81]. The chiral quark model provides a link between QCD and chiral perturbation theory and

bears some similarity to the extended Nambu-Jona-Lasinio model already mentioned above. In this framework, the hadronic matrix elements depend on the values of quark and gluon condensates, also present in the QCD sum rule approach, as well as constituent masses for the quarks. For values of these parameters which fit the $\Delta I = 1/2$ rule for $\mathrm{K} \to \pi\pi$ decays, the authors of [81] then obtain $\hat{B}_K = 1.1 \pm 0.2$, where the error is dominated by the variation of constituent quark mass and gluon condensate. However, owing to a poor matching between long- and short-distance contributions in the case of B_K, an additional systematic uncertainty of the order of 15% could be present in the result of Ref. [81].

3. Experimental methods for the study of B^0 and $\overline{\mathrm{B}}^0$ mixing

The system of neutral B mesons, B^0 and $\overline{\mathrm{B}}^0$, can be described in terms of states with well defined mass and lifetime exhibiting the phenomenon of particle-antiparticle oscillations. The frequency of B^0_d and B^0_s mixing can be described by the mass difference $\Delta M_{d,s}$ as defined in Eq. (14). This mass difference between the two mass eigenstates leads to a time-dependent phase difference between the particle wave functions. In the Standard Model, B^0–$\overline{\mathrm{B}}^0$ mixing is described via second order weak processes, as displayed for the case of K^0–$\overline{\mathrm{K}}^0$ mixing in Fig. 4.1. The mass difference $\Delta M_{d,s}$ can be determined by computing the electroweak box diagram, where the dominant contribution is through top quark exchange as can be seen in Eq. (22). A measurement of ΔM_d or ΔM_s in principle allows the determination of the Cabibbo-Kobayashi-Maskawa matrix elements $|V_{td}|$ or $|V_{ts}|$ as indicated by the relations in Eq. (23) and (24). The main uncertainty in relating measurements of the mixing frequency to the CKM matrix elements originates from the parameters $F_{B_{d,s}}$ and $\hat{B}_{B_{d,s}}$ as discussed in Sec. 2.1.. However, in the ratio $\Delta M_d/\Delta M_s$ several of the theoretical uncertainties cancel as is obvious from Eq. (29). Thus, the ratio $\Delta M_d/\Delta M_s$ is related to the ratio of CKM matrix elements $|V_{td}|/|V_{ts}|$ and will ultimately determine one of the sides of the CKM unitarity triangle.

3.1. Time integrated oscillation analyses and determination of B hadron production rates

At the $\Upsilon(4S)$, only B^0_d and B^+ mesons are produced, whereas at high energy colliders B^0_s mesons and b-baryons are also present. In the latter case, B^0_d and B^0_s mesons contribute to time integrated mixing measurements with a weight proportional to their relative production fractions:

$$\bar{\chi} = f_{\mathrm{B}^0_d}\,\chi_d \;+\; f_{\mathrm{B}^0_s}\,\chi_s. \tag{46}$$

Here, $f_{\mathrm{B}^0_d}$ and $f_{\mathrm{B}^0_s}$ are the production rates of B^0_d and B^0_s mesons in b quark jets, while $\chi_{d,s}$ are the respective mixing parameters defined in Eq. (28)[‖]. The non-linear relation between x and χ (see Eq. (28)) implies that χ becomes insensitive to x for values greater than $x \sim 5$. Thus, a time dependent oscillation analysis is necessary to observe fast oscillations as expected for B^0_s mesons. At the $\Upsilon(4S)$ resonance, a measurement of χ_d allows to directly extract x_d because only slowly oscillating B^0_d mesons are produced. A time integrated mixing analysis is, however, important to determine the hadron production fractions $f_{\mathrm{B}^0_d}$ and $f_{\mathrm{B}^0_s}$. For example, $f_{\mathrm{B}^0_d}$ is an essential input for a measurement of V_{cb} using $\overline{\mathrm{B}}^0_d \to D^{*+}\ell^-\bar{\nu}_\ell$ decays and the source of an important systematic error in ΔM_d measurements at high energy colliders. Furthermore, the sensitivity to B^0_s–$\overline{\mathrm{B}}^0_s$ oscillations in inclusive analyses depends on the B^0_s production rate $f_{\mathrm{B}^0_s}$.

The production rates of B hadrons in b quark jets can be obtained from the measured integrated oscillation rates of B mesons (see Eq. (46)). When measuring the time integrated oscillation parameter in a semileptonic sample, the mixing probability can be written as

$$\bar{\chi} = g_{\mathrm{B}^0_s}\chi_s + g_{\mathrm{B}^0_d}\chi_d, \tag{47}$$

[‖] The world average for the time integrated mixing parameter is $\bar{\chi}$=0.1194 $\pm$ 0.0043 [84].

b-hadron fractions	direct measurement	direct plus mixing
$f_{\mathrm{B}_s^0}$	$(9.2 \pm 2.4)\%$	$(9.3 \pm 1.1)\%$
$f_{b-\mathrm{baryon}}$	$(10.5 \pm 2.0)\%$	$(10.5 \pm 1.8)\%$
$f_{\mathrm{B}_d^0} = f_{\mathrm{B}^+}$	$(40.1 \pm 1.3)\%$	$(40.1 \pm 1.1)\%$

Table 4.4: *Average values of b-hadron production rates obtained from direct measurements and using time integrated mixing as of the ICHEP 2002 conference [87].*

where $g_{\mathrm{B}_s^0}$ and $g_{\mathrm{B}_d^0}$ are the fractions of B_d^0 and B_s^0 mesons in a semileptonic sample. Assuming that the semileptonic width is the same for all B hadrons, we obtain

$$g_{B_i} = f_{B_i} R_i \quad \text{where} \quad R_i = \frac{\tau_i}{\tau_B}. \tag{48}$$

This results in

$$\begin{aligned} f_{\mathrm{B}_s^0} &= \frac{1}{R_s} \frac{(1+r)\,\bar{\chi} - (1 - f_{b-\mathrm{baryon}}\, R_{b-\mathrm{baryon}})\,\chi_d}{(1+r)\,\chi_s - \chi_d} \\ f_{\mathrm{B}_d^0} &= \frac{1}{R_d} \frac{\bar{\chi} - (1 - f_{b-\mathrm{baryon}}\, R_{b-\mathrm{baryon}})\,\chi_s}{\chi_d - (1+r)\,\chi_s} \end{aligned} \tag{49}$$

where $r = R_u/R_d = \tau(B^+)/\tau(\mathrm{B}_d^0)$. We assume $f_{\mathrm{B}_d^0} = f_{\mathrm{B}^+}$, $f_{\mathrm{B}^+} + f_{\mathrm{B}_d^0} + f_{\mathrm{B}_s^0} + f_{b-\mathrm{baryon}} = 1$ and $\chi_s = 0.5$.

From the previous expressions, the values of $f_{\mathrm{B}_s^0}$ and $f_{\mathrm{B}_d^0}$ are determined and combined with those obtained from direct measurements (for more details see Ref. [85]). The results are shown in Table 4.4. It is clear that $f_{\mathrm{B}_s^0}$ is essentially determined from the time integrated mixing measurement. The error on $f_{\mathrm{B}_s^0}$ is dominated by the uncertainty on the integrated oscillation parameter $\bar{\chi}$, which is not expected to improve substantially in the near future. Different uncertainties contribute to the error on $f_{\mathrm{B}_d^0}$. The most important one is the poor knowledge of the b-baryon production rates. It has to be noted that $f_{\mathrm{B}_d^0}$ is essentially determined by the DELPHI direct measurement [86].

3.2. Flavour tagging techniques

In general, a measurement of the time dependence of B^0–$\overline{\mathrm{B}}^0$ oscillations requires the knowledge of:

- the proper decay time t of the B^0 meson (see Sec. 3.3.),
- the flavour of the B or $\overline{\mathrm{B}}$ meson at both production and decay in order to determine whether the B^0 meson has oscillated.

Events are classified on the basis of the sign of the production and decay tagging variables as mixed or unmixed. To accomplish this, it is necessary to determine the b quark content (b or $\bar{b}$) of the B meson at production and at decay time. The figure of merit to compare different flavour tags is the so-called effective tagging efficiency $\varepsilon(1 - 2\,p_W)^2$, where the efficiency ε represents the fraction of events for which a flavour tag exists and p_W is the mistag probability indicating the fraction of events with a wrong flavour tag. The mistag probability is related to the dilution $\mathcal{D}$, another quantity used to express the power of a flavour tag:

$$\mathcal{D} = 1 - 2\,p_W. \tag{50}$$

The dilution $\mathcal{D}$ is defined as the number of correctly tagged events N_R minus the number of incorrectly identified events N_W divided by the sum:

$$\mathcal{D} = \frac{N_R - N_W}{N_R + N_W}. \tag{51}$$

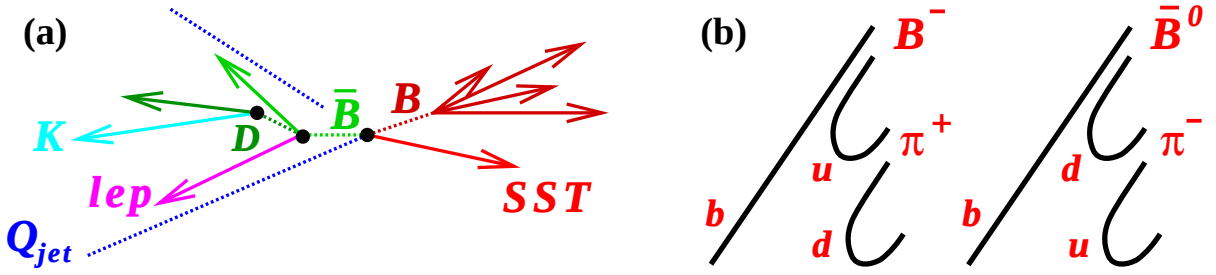

Fig. 4.7: *(a) Schematic sketch of a typical* $\mathrm{B\overline{B}}$ *event. (b) A simplified picture of* b *quark fragmentation into* B *mesons.*

Fig. 4.7(a) is a sketch of a $\mathrm{B\overline{B}}$ event showing the B and $\overline{\mathrm{B}}$ mesons originating from the primary production vertex and decaying at a secondary vertex indicating possible flavour tags on the decay vertex side (SST) and opposite side (lep, K, Q_{jet}).

3.2.1. Decay flavour tagging

Several techniques are used to determine the b quark flavour at decay time. The B flavour can be identified by the charge of a lepton from a semileptonic B decay. In a prompt $b \to \ell^-$ decay, the charge of the lepton reflects the b flavour. However, other processes can also give a lepton in the final state such as cascade decays $b \to c \to \ell^+$ resulting in a wrong sign tag, right sign cascade decays $b \to W^- \to \bar{c} \to \ell^-$, semileptonic τ decays $b \to W^- \to \tau^- \to \ell^-$ or $b \to J/\psi X \to \ell^\pm$ decays giving both sign leptons. These processes resulting in wrong sign leptons can be suppressed by using the lepton momentum or transverse momentum with respect to the b jet axis.

The b quark flavour can also be inferred from the charge of a reconstructed charm meson ($\mathrm{D}^{(*)-}$ from B^0_d or D^-_s from B^0_s) or that of a kaon assumed to come from a $b \to c \to s$ transition. In fully inclusive analyses, the b flavour can be obtained from the jet charge (see Eq. (52)), the charge of a reconstructed dipole or from multitags as further detailed in Sec. 3.4.

3.2.2. Production flavour tagging

Methods to tag the production b quark flavour differ somewhat between high energy colliders (LEP, SLC, Tevatron) and the B factories. At high-energy colliders, the production flavour tags can be divided into two groups, those that tag the initial charge of the b quark contained in the B candidate itself (same side tag) and those that tag the initial charge of the other quark ($\bar{b}$) produced in the same event (opposite side tag).

Same side tagging methods exploit correlations of the B flavour with the charge of particles produced in association with the B meson. Such correlations are expected to arise from b quark hadronization and from B^{**} decays. It has been suggested [88] that the electric charge of particles produced near a B meson can be used to determine its initial flavour. This can be understood in a simplified picture of b quark fragmentation as shown in Fig. 4.7(b). For example, if a b quark combines with a $\bar{u}$ quark to form a B^- meson, the remaining u quark may combine with a $\bar{d}$ quark to form a π^+. Similarly, if a b quark hadronizes to form a $\overline{\mathrm{B}}^0$ meson, the associated pion would be a π^-. A similar charge correlation is expected for a charged kaon produced in association with a B^0_s meson. Decays of the orbitally excited ($L = 1$) B^{**} mesons, $\mathrm{B}^{**0} \to \mathrm{B}^{(*)+}\pi^-$ or $\mathrm{B}^{**+} \to \mathrm{B}^{(*)0}\pi^+$, also produce pions with the same charge correlation. This tagging method has been successfully used for example at CDF [89,90].

There are several methods of opposite side flavour tagging as illustrated in Fig. 4.7(a). The methods using a lepton from the semileptonic decay of a B hadron, a kaon or the presence of a charmed particle from the other $\bar{B}$ hadron in the event, were already discussed above.

The technique based on the jet charge exploits the fact that the momentum weighted sum of the particle charges of a b jet is related to the b-quark charge. In the most basic form, the jet charge can be defined as:

$$Q_{\rm jet} = \frac{\sum_i q_i \cdot (\vec{p}_i \cdot \hat{a})}{\sum_i \vec{p}_i \cdot \hat{a}}, \tag{52}$$

where q_i and $\vec{p}_i$ are the charge and momentum of track i in the jet and $\hat{a}$ is a unit vector defining the jet direction. On average, the sign of the jet charge is the same as the sign of the b quark charge that produced the jet. More sophisticated weights (e.g. $(\vec{p}_i \cdot \hat{a})^\kappa$) or track impact parameter information are often introduced to improve the b flavour separation. The jet charge can also be used as a same side tag, if tracks from primary vertex can be efficiently distinguished with respect to those from secondary decay vertices.

Other tagging methods include the charge dipole method that aims of reconstructing the b hadron decay chain topologically. This method has been utilized at SLD taking advantage of the superb decay length resolution of the SLD CCD pixel vertex detector to separate tracks from the B decay point from tertiary tracks emitted at the charm decay vertex [91]. A charge dipole is defined as the distance between secondary and tertiary vertices signed by the charge difference between them (see also Sec. 3.4.).

Another interesting production flavour tagging method is available at SLD. It exploits the large polarized forward-backward asymmetry in $Z \to b\bar{b}$ decays [92–95]. This b flavour production tag makes use of the large electron beam polarization $P_e \sim 73\%$ at the SLC collider. A left- or right-handed incident electron tags the quark produced in the forward hemisphere as a b or $\bar{b}$ quark with a mistag rate p_W of 28% at nearly 100% efficiency [91].

At asymmetric e^+e^- B factories, $\mathrm{B}^0_d - \overline{\mathrm{B}}^0_d$ pairs are produced through the $\Upsilon(4S)$ resonance with a boost $\beta\gamma = 0.425$ and 0.55 at KEKB and PEP II, respectively. The two neutral B_d mesons produced from the $\Upsilon(4S)$ decay evolve in time in a coherent P-wave state where they keep opposite flavours until one of the B_d mesons decays. From this point in time onwards, the other B meson follows a time evolution according to the expression $\Gamma \mathrm{e}^{-\Gamma|\Delta t|}\,(1 \pm \cos \Delta M\, \Delta t)$ where Δt is the proper time difference between the two B decays. Hence, the production flavour tag of one of the B mesons can be taken as the decay flavour tag of the other. The main flavour tagging methods currently used at BaBar and Belle include $b \to \ell^-$ lepton tagging and $b \to c \to s$ kaon tagging.

It is common to combine different production tags in an oscillation analysis to achieve mistag probabilities of $p_W \sim 26\%$ at LEP [96–101] or even 22% for SLD [102]. An equivalent figure for CDF in Run I of the Tevatron is $p_W \sim 40\%$ [103]. Effective mistag probabilities of $p_W \sim 24\%$ are achieved by the BaBar and Belle experiments [104,105]. It is interesting to mention that the effect of B^0_d and B^0_s mixing substantially decreases the tagging power of opposite side tagging methods at high-energy colliders while mixing of the other B meson (i.e. the coherent mixing occurring before the first B decay) does not contribute to a mistag probability at the $\Upsilon(4S)$.

3.3. Analytical description of oscillation analyses

A physics function of the form $\Gamma \mathrm{e}^{-\Gamma t}\,(1 \pm \cos \Delta M\, t)$ is used to describe the signal in B oscillation analyses. At high energy colliders such as LEP, SLC or the Tevatron, the B meson decay proper time t can be obtained from a measurement of the distance L_B between the B production vertex and the B decay vertex. The proper time t is related to the decay distance L_B and to the boost $\beta\gamma$ by

$$c\,t = \frac{L_B}{\beta\gamma} = L_B\,\frac{M_B}{p_B}. \tag{53}$$

At asymmetric e^+e^- B factories, the proper time difference Δt between the two B candidate decays is the relevant measure. It is computed as:

$$\Delta t = \Delta z / \beta\gamma c, \tag{54}$$

where Δz is the spatial separation between the two B decay vertices along the boost direction.

The uncertainty on the decay time σ_t can be expressed in units of the B lifetime τ_B as

$$\frac{\sigma_t}{\tau_B} = \sqrt{\left(\frac{\sigma(L_B)}{L_B^0}\right)^2 + \left(\frac{t}{\tau_B}\frac{\sigma(p_B)}{p}\right)^2} \qquad \text{where} \quad L_B^0 = c\tau_B \cdot p_B/M_B. \tag{55}$$

The proper time resolution σ_t depends on the uncertainty $\sigma(L_B)$ to infer the decay length from the primary to the B decay vertex and on the B momentum resolution $\sigma(p_B)$. Note that the latter uncertainty scales with t/τ_B, while the vertexing resolution is independent of the proper time and only adds a constant error.

The dependence of B oscillations on the proper time resolution and other detector effects is illustrated in Fig. 4.8. Rather than plotting the mixed and unmixed probabilities $\mathcal{P}_{\text{unmix/mix}}(t) = 1/2\ \Gamma\mathrm{e}^{-\Gamma t}\,(1\pm\cos\Delta M\, t)$ as introduced in Eq. (27) and Eq. (26), it is customary in B oscillation analyses to either determine a mixing asymmetry $\mathcal{A}_{\text{mix}}$ or to calculate the fraction of mixed events $\mathcal{F}_{\text{mix}}$

$$\mathcal{A}_{\text{mix}} = \frac{\mathcal{P}_{\text{unmix}} - \mathcal{P}_{\text{mix}}}{\mathcal{P}_{\text{unmix}} + \mathcal{P}_{\text{mix}}} = \cos\Delta M\, t, \qquad \mathcal{F}_{\text{mix}} = \frac{\mathcal{P}_{\text{mix}}}{\mathcal{P}_{\text{unmix}} + \mathcal{P}_{\text{mix}}} = (1 - \cos\Delta M\, t)/2. \tag{56}$$

As an example, Fig. 4.8(a) shows the oscillation pattern of $\mathcal{A}_{\text{mix}}$ for $\Delta M = 5$ ps^{-1} assuming an ideal case with perfect tagging, ideal proper time resolution and no background. The reduction of the amplitude due to a finite decay length resolution is shown in Fig. 4.8(b). Figure 4.8(c) indicates what happens when the resolution of the (silicon) vertex detector is not sufficient to resolve the oscillations: $\mathcal{A}_{\text{mix}}$ is completely smeared out and oscillations are no longer visible. The effect of a finite momentum resolution is displayed in Fig. 4.8(d). Since the uncertainty on the proper time coming from the momentum resolution is linear in proper time t, as seen in Eq. (55), the rapid oscillation damps in time while the first few "wiggles" can still be seen completely. The oscillation amplitude is reduced if a mistag probability is introduced, as can be seen in Fig. 4.8(e). Finally, in a real measurement, background will also be present which additionally reduces the relative importance of the oscillation amplitude. The effect of background on the mixing amplitude, in addition to a finite decay length and momentum resolution, as well as a non-zero mistag probability, is shown in Fig. 4.8(f). Note, however, that this "realistic" distribution is based on half a million signal events. Imagine the corresponding error bars for a measurement with a few hundred signal events and an oscillation frequency of $\Delta M = 20$ ps^{-1}.

In a B^0 mixing measurement, a value for ΔM is usually extracted from the data using a maximum likelihood method. In the following, we illustrate some of the essential steps for a B_d^0 analysis determining ΔM_d in more detail. We use the example of an analysis where like-sign (unlike-sign) events describe mixed (unmixed) events as would be the case, for example, in a dilepton analysis. The total probability to observe a like-sign tagged event at the reconstructed proper time t_{rec} is:

$$\mathcal{P}^{like}(t_{rec}) = f_{b\bar{b}} \sum_{q=d,s} f_{\mathrm{B_q}} p_W^{\mathrm{B_q}} \mathcal{P}^{\text{mix}}_{rec.\mathrm{B_q}}(t_{rec}) + f_b \sum_{q=u,d,s,baryons} f_{\mathrm{B_q}}(1 - p_W^{\mathrm{B_q}}) \mathcal{P}^{\text{unmix}}_{rec.\mathrm{B_q}}(t_{rec}) \quad + \\ f_{bkg.}(1 - p_W^{bkg.}) \mathcal{P}_{bkg.}(t_{rec}) \tag{57}$$

and correspondingly for an unlike-sign tagged event:

$$\mathcal{P}^{unlike}(t_{rec}) = f_{b\bar{b}} \sum_{q=d,s} f_{\mathrm{B_q}}(1 - p_W^{\mathrm{B_q}}) \mathcal{P}^{\text{mix}}_{rec.\mathrm{B_q}}(t_{rec}) + f_b \sum_{q=u,d,s,baryons} f_{\mathrm{B_q}} p_W^{\mathrm{B_q}} \mathcal{P}^{\text{unmix}}_{\mathrm{B_q}}(t_{rec}) \quad + \\ f_{bkg.} p_W^{bkg.} \mathcal{P}_{bkg.}(t_{rec}). \tag{58}$$

where $f_{b\bar{b}}$ is the fraction of $b\bar{b}$ events and p_W^i are the mistag probabilities. The probability $\mathcal{P}^{\text{mix}}_{rec.\mathrm{B_q}}(t_{rec})$ to observe the mixed B_d^0 or B_s^0 mesons at proper time t_{rec} is the result of a convolution of the oscillation probability function as given in Eq. (26) and Eq. (27) with the detector resolution function $\mathcal{R}$ and

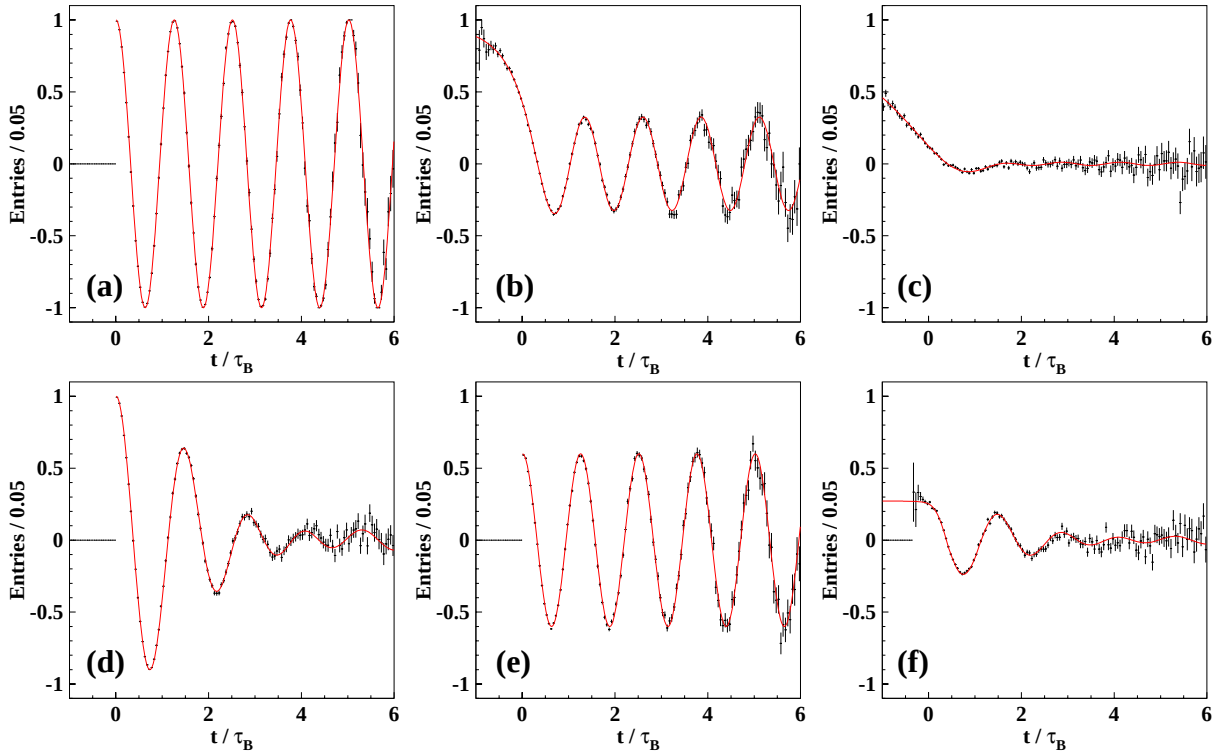

Fig. 4.8: *Illustration of various detector and analysis effects on the mixing amplitude $\mathcal{A}_{\rm mix}$: (a) perfect resolution, (b) good decay length resolution, (c) poor decay length resolution, (d) finite momentum resolution, (e) mistag probability and (f) decay length and momentum resolution plus mistag including background.*

weighted with an acceptance function $Acc(t)$

$$\mathcal{P}^{(\rm un)\rm mix}_{rec.\rm B_q}(t_{rec}) = \int_0^\infty Acc(t)\mathcal{R}(t_{rec}-t,t)\mathcal{P}^{(\rm un)\rm mix}_{\rm B_q}(t)dt. \tag{59}$$

To extract the value ΔM of the oscillation frequency, the following likelihood function is minimized :

$$\mathcal{L} = -\sum_{like-sign}\ln(\mathcal{P}^{like}(t_{rec})) - \sum_{unlike-sign}\ln(\mathcal{P}^{unlike}(t_{rec})). \tag{60}$$

In order to fully exploit the available statistics, more sophisticated mixing analyses make use of those variables on an event-by-event basis, or often divide the event sample into classes with e.g. different tagging capabilities.

3.3.1. The amplitude method

For ΔM_s measurements, the amplitude method [106] is used to set limits on ΔM_s and to combine results from different analyses. For the mixed and unmixed $\rm B^0_s$ events an amplitude A is introduced in the expressions describing the mixed and unmixed probabilities:

$$\mathcal{P}^{\rm unmix}_{\rm B^0_s} = \frac{1}{2}\Gamma_{\rm B_s}{\rm e}^{-\Gamma_{\rm B_s}t}[1+A\cos\Delta M_s t] \tag{61}$$

and similarly:

$$\mathcal{P}^{\rm mix}_{\rm B^0_s} = \frac{1}{2}\Gamma_{\rm B_s}{\rm e}^{-\Gamma_{\rm B_s}t}[1-A\cos\Delta M_s t] \tag{62}$$

The amplitude method works as follows. A B_s^0 oscillation amplitude A and its error σ_A are extracted as a function of a fixed test value of ΔM_s using a likelihood method in analogy to Eq. (60) based on the physics functions defined in Eq. (61) and Eq. (62). To a very good approximation, the statistical uncertainty on A is Gaussian and the experimental sensitivity is :

$$\mathcal{S} = \frac{1}{\sigma_A} \sim \sqrt{N/2}\, f_{\rm sig}\, (1-2p_w)\, \mathrm{e}^{-(\Delta M\, \sigma_t)^2/2} \tag{63}$$

where N and $f_{\rm sig}$ are the number of candidate events and the fraction of signal in the selected sample, p_W is the mistag probability to incorrectly tag a decay as mixed or unmixed characterizing the effective flavour tagging efficiency as discussed in Sec. 3.2., and σ_t is the resolution on proper time or proper time difference in the case of the B factories. The sensitivity $\mathcal{S}$ decreases rapidly as ΔM increases. This dependence is controlled by σ_t.

If ΔM_s equals its true value $\Delta M_s^{\rm true}$, the amplitude method expects $A = 1$ within the total uncertainty σ_A. If ΔM_s is tested far below its true value, a measurement consistent with $A = 0$ is expected. A value of ΔM_s can be excluded at 95% C.L. if $A + 1.645\,\sigma_A \leq 1$. If the true B_s^0 oscillation frequency $\Delta M_s^{\rm true}$ is very large, far above the experimental sensitivity, $A = 0$ is expected to be measured and all values of ΔM_s such that $1.645\,\sigma_A(\Delta M_s) < 1$ are expected to be excluded at 95% C.L. Because of proper time resolution, the quantity $\sigma_A(\Delta M_s)$ is an increasing function of ΔM_s. It is therefore expected that individual values of ΔM_s can be excluded up to $\Delta M_s^{\rm sens}$, where $\Delta M_s^{\rm sens}$ is called the sensitivity of the analysis defined by $1.645\,\sigma_A(\Delta M_s^{\rm sens}) = 1$. The results from different analyses and experiments can be combined by simple averaging different amplitude spectra.

3.4. Description of oscillation analyses

Many different analysis methods have been devised to study B_d^0 and B_s^0 mixing. These range from fully inclusive to fully exclusive analyses and, thus, they differ significantly in terms of selection efficiency, sample purity and mistag rates. Moreover, they make use of various production and decay tags. The methods also differ in the techniques used to reconstruct the B decay length and to estimate the B momentum, and therefore have different proper time resolutions. In the following, analysis methods developed to measure ΔM_d are discussed first and those used in the search for B_s^0 oscillations are presented afterwards.

3.4.1. B_d^0–$\overline{\mathrm{B}}_d^0$ *oscillation analyses*

Exclusive methods

The most straightforward and cleanest method relies on the exclusive reconstruction of the B_d^0 decay chain. However, because of its low efficiency, it has only recently become accessible with the advent of e^+e^- asymmetric B factories. Using samples of $\sim$30M $\mathrm{B\overline{B}}$ events, BaBar [107] and Belle [108] reconstruct the decays $\mathrm{B}_d^0 \to \mathrm{D}^{(*)-}\pi^+$, $\mathrm{D}^{(*)-}\rho^+$, $\mathrm{D}^{(*)-}a_1^+$, $J/\psi\mathrm{K}^{*0}$ (BaBar), and $\mathrm{B}_d^0 \to \mathrm{D}^{(*)-}\pi^+$, $\mathrm{D}^{*-}\rho^+$ (Belle), where charmed mesons are fully reconstructed in several D^{*-} and $\overline{\mathrm{D}}^0$ decay modes. Very clean signals are obtained, see Fig. 4.9, and the decay flavour is unambiguously determined by the charge of the $\mathrm{D}^{(*)}$ meson (or the charged kaon in case of the $J/\psi\mathrm{K}^{*0}$ decay).

The average separation of the two B decay points is $\Delta z = 255\ (200)\ \mu\mathrm{m}$ with $\sigma_z \simeq 180\ (140)\mu\mathrm{m}$ for Babar (Belle), which corresponds to a resolution on Δt (Eq. 54) of about $1.1\,\mathrm{ps}$. For a measurement of the B_d^0 oscillation frequency it is therefore critical to have good control over the resolution. Table 4.5 summarizes the number of events, signal mode purity and production flavour tag information for these as well as all other analyses presented below.

Decay modes	Analysis	Events/Signal	$f_{\rm mode}$	Production flavour tag
$\mathrm{B}^0_d \to \mathrm{D}^{(*)-}h^{+a}$	BaBar exclusive [107]	7380/6347	86%	Multiple tags
$J/\psi \mathrm{K}^{*0}$	Belle exclusive [108]	8325/6660	80%	Multiple tags
$\mathrm{B}^0_d \to \mathrm{D}^{*-}\pi^+$	Belle semi-incl. [116]	4899/3433	70%	Lepton
$\mathrm{B}^0_d \to \mathrm{D}^{(*)-}X$	ALEPH semi-excl. [113]	4059/2395	38?%	Lepton+jet charge
	CDF semi-excl. [103]	874/358	27%	Lepton
	DELPHI semi-excl. [114]	10030/4212	27?%	Jet charge
	OPAL semi-excl. [112]	347/253	48%	Lepton
$\mathrm{B}^0_d \to \mathrm{D}^{(*)-}\ell^+\nu$	BaBar semi-excl. [109]	17506/14182	74%	Multiple tags
	Belle semi-excl. [110]	16397/15118	80%	Multiple tags
	CDF semi-excl. [111]	888/530		Lepton
	CDF semi-excl. [89]	/6266		Same-side tag
	OPAL semi-excl. [112]	1200/926	65%	Jet charge
	DELPHI semi-incl. [114]	5958/4135	59%	Jet charge
	OPAL semi-incl. [115]	/7000	36%	Multiple tags
$\mathrm{B}^0_d \to X\ell^+\nu$	BaBar semi-incl. [120]	99k/	37%	Lepton
	Belle semi-incl. [121]	281k/		Lepton
	ALEPH semi-incl. [113]	5957/		Lepton
	CDF semi-incl. [117]	5968/	39%	Lepton ($\mu\mu$)
	CDF semi-incl. [103]	10180/		Lepton ($e\mu$)
	DELPHI semi-incl. [114]	4778/	33%	Lepton
	L3 semi-incl. [119]	1490/		Lepton
	L3 semi-incl. [119]	2596/	34%	Lepton (impact parameter)
	OPAL semi-incl. [100]	5357/		Lepton
	ALEPH semi-incl. [113]	62k/		Jet charge
	CDF semi-incl. [118]	13k/		Lepton+jet charge
	DELPHI semi-incl. [114]	60k/	29%	Jet charge
	OPAL semi-incl. [101]	95k/	30%	Jet charge
	L3 semi-incl. [119]	8707/		Jet charge
	SLD semi-incl. [93]	581/	51%	Polarization+jet charge
	SLD semi-incl. [92]	2609/	31%	Polarization+jet charge
$\mathrm{B}^0_d \to$ all	ALEPH inclusive [123]	423k/	35%	Jet charge
	DELPHI inclusive [122]	770k/	40%	Multiple tags
	SLD inclusive [94]	3291/	60%	Polarization+jet charge; Charge dipole decay tag
	SLD inclusive [94]	5694/	60%	Polarization+jet charge; Kaon decay tag 1993–95
	SLD inclusive [95]	7844/	60%	Multiple tags; Kaon decay tag 1996–98

[a] h^+ stands for π^+, ρ^+, a_1^+.

Table 4.5: *Summary of* B^0_d *mixing analyses showing the signal decay modes, analysis method, total number of selected events and estimated signal, fraction of signal decay mode in the selected sample (*$f_{\rm mode}$*), and production flavour tag.*

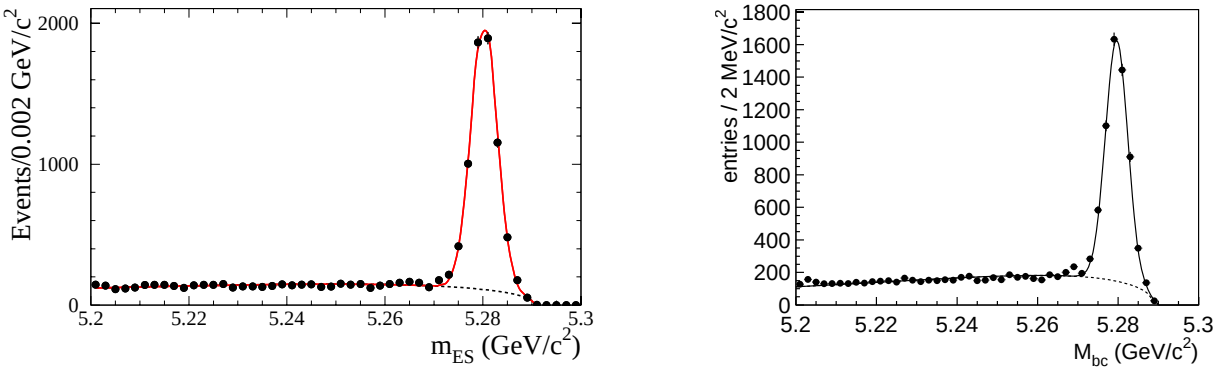

Fig. 4.9: *Distributions of beam-energy substituted mass for exclusively reconstructed* B^0_d *decays in the BaBar (left) and Belle (right) analyses.*

Semi-exclusive methods

Several analyses have combined an identified lepton with a fully reconstructed charmed hadron. Generally, the presence of a $\mathrm{D}^{(*)-}$, with charge opposite that of the lepton, tags the decay of a B^0_d meson. This simple picture is complicated by decays of the type $\mathrm{B}^+ \to \overline{\mathrm{D}}^{**0}\ell^+\nu$, where the $\overline{\mathrm{D}}^{**0}$ decays into a $\mathrm{D}^{(*)-}$ meson.

Measurements have been performed at B factories by BaBar [109] and Belle [110] and at high energy colliders by CDF [111,89] and OPAL [112]. B^0_d mesons are partially reconstructed in the mode $\mathrm{B}^0_d \to \mathrm{D}^{(*)-}\ell^+\nu$, where the D^{*-} or D^- meson is fully reconstructed. The selection relies on the kinematical properties of B^0_d and $\mathrm{D}^{(*)-}$ decays. In particular, the low Q value of the decay $\mathrm{D}^{*-} \to \overline{\mathrm{D}}^0\pi^-$ is exploited to identify D^{*-} mesons efficiently and cleanly. Fig. 4.10 shows the mass difference $\Delta M = M(\mathrm{D}^{*-}) - \mathrm{M}(\overline{\mathrm{D}}^0)$ in the BaBar and OPAL analyses. Signal purities range from ∼45% to ∼90% for the different experiments, depending mostly on the $\overline{\mathrm{D}}^0$ decay mode.

In order to increase the selection efficiency, analyses by ALEPH [113], CDF [103], DELPHI [114], and OPAL [112] select $\mathrm{B}^0_d \to \mathrm{D}^{(*)-}X$ decays, where the $\mathrm{D}^{(*)-}$ meson is also fully reconstructed. Despite the more inclusive nature of this method, the identification of a $\mathrm{D}^{(*)-}$ decay guarantees that the B^0_d purity remains high. However, $b\overline{b}$ tagging is generally needed to suppress the significant number of D^{*-} produced in $c\overline{c}$ events.

Semi-inclusive methods

One of the semi-inclusive methods selects $\mathrm{B}^0_d \to \mathrm{D}^{*-}\ell^+\nu$ decays without attempting to fully reconstruct the $\overline{\mathrm{D}}^0$ meson but only the lepton and the slow π^- from the $\mathrm{D}^{*-} \to \overline{\mathrm{D}}^0\pi^-$ decay. This partial reconstruction method yields much larger data samples than obtained with the exclusive reconstruction but suffers from higher background. It has been applied by DELPHI [114] and OPAL [115]. The combinatorial background can be studied with same-sign lepton-pion pairs and ΔM side bands. The $\mathrm{B}^+ \to \overline{\mathrm{D}}^{**0}\ell^+\nu$ component needs to be estimated from the simulation.

A similar technique is used by Belle [116] to reconstruct $\mathrm{B}^0_d \to \mathrm{D}^{*-}\pi^+$ decays. In this analysis, only the fast π^+ and the slow π^- are reconstructed. This information is sufficient to compute the $\overline{\mathrm{D}}^0$ missing mass, assuming that the B^0_d meson is at rest in the $\Upsilon(4\mathrm{S})$ rest frame and using energy and momentum conservation. The event is required to contain a high-momentum lepton to tag the other B meson flavour and to suppress the large non-$\mathrm{B}\overline{\mathrm{B}}$ background. This method is only possible at the $\Upsilon(4\mathrm{S})$ where sufficient kinematical constraints are available.

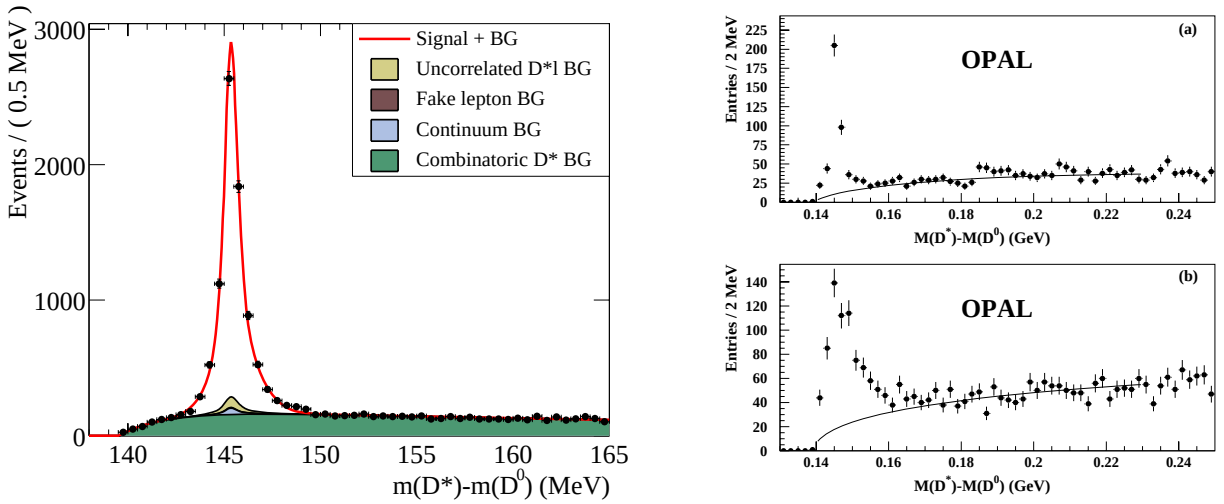

Fig. 4.10: *Distributions of the mass difference between* D^{*-} *and* $\overline{\mathrm{D}}^0$ *candidates for BaBar (left) and OPAL (right). The BaBar distribution is shown for* $\mathrm{D}^{*-}e^+\nu$ *candidates. The distributions on the right correspond to the modes* $\overline{\mathrm{D}}^0 \to \mathrm{K}^+\pi^-$ *(top) and* $\overline{\mathrm{D}}^0 \to \mathrm{K}^+\pi^-\pi^0$ *(bottom), where the* π^0 *is not reconstructed.*

The most widely used method relies on the inclusive reconstruction of semileptonic decays. At high energy colliders, it has been employed by ALEPH [113], CDF [117,118,103], DELPHI [114], L3 [119], OPAL [100,101], and SLD [93,92]. This method is efficient since the decay rate for $\mathrm{B}^0_d \to X\ell^+\nu$ is approximately 20% (using electrons and muons) and the decay flavour tag is excellent. A high-p and high-p_T lepton is selected to suppress the contribution from cascade leptons (from $b \to c \to \ell^+$ transitions) and the accompanying charmed hadron (denoted "D" in the following) is reconstructed inclusively using charged tracks in the jet containing the lepton. The position of the B decay vertex and the B momentum are obtained using algorithms that aim to classify tracks as coming from either primary or secondary vertices. The B decay vertex is then obtained by intersecting the trajectories of the lepton and that of a D candidate.

The analyses are combined with a variety of different production flavour tags and are thus referred to as "dilepton", "lepton-jet charge" and "Multiple tags" analyses (see Table 4.5).

Dilepton analyses have also been performed by both BaBar [120] and Belle [121]. Here, there is no attempt to reconstruct the D decay and the time difference is extracted directly from the separation Δz between the intersections of the two leptons with the beam axis. Momentum and angular cuts are applied to reduce the wrong-sign background from cascade leptons. In the BaBar analysis, the main background consists of $\mathrm{B}^+\mathrm{B}^-$ events and is determined to be $\sim$55% and the main source of mistag originates from events containing one direct lepton and one cascade lepton, amounting to 13% of the total sample.

Inclusive methods

A few analyses rely on fully inclusive techniques to select large samples of B^0_d decays. These techniques aim to capture most decays by using topological vertexing. As for the semi-inclusive methods, the selection algorithms generally do not provide any enhancement in the B^0_d purity. The primary issue here is the decay flavour tag.

SLD uses two different decay tags: the charge of a kaon coming from the B decay chain [94,95] or the charge dipole of the secondary vertex [94]. These analyses require the net charge of all tracks associated with the decay to be zero to enhance the B^0_d fraction from $\sim$40% to $\sim$60%. The kaon decay tag is more efficient than the lepton decay tag but has a worse mistag rate of $\sim$20%. The charge dipole technique takes advantage of the $\mathrm{B}^0_d \to \mathrm{D}^- X^+$ dipole structure and the fact that the B^0_d and D^- vertices

are separated along the B_d^0 line of flight due to the finite charm lifetime. For the B_d^0 analyses the charge dipole is defined as the difference between the weighted mean location of the positive tracks and of the negative tracks along the axis joining the primary and secondary vertices. The track weights account for the uncertainty in determining the location of each track. A positive (negative) charge dipole tags the decay flavour of the $\overline{B}_d^0$ (B_d^0) meson.

At LEP, DELPHI [122] also developed a fully inclusive method based on the charge dipole tag. The vertex algorithm uses topological and kinematical information to separate primary and secondary tracks. A secondary lepton is found in a subset of the vertices and provides the decay flavour tag (these leptons are referred to as "soft" leptons since decays with high p and p_T are used in other DELPHI analyses). For the remainder of the sample, the B decay products are boosted back into the B meson rest frame and a charge dipole is formed between the forward and backward hemispheres (as defined by the thrust axis). Given that the forward (backward) hemisphere contains mostly tracks from the D (B) decay vertex, one expects a ± 2 charge difference between the two hemispheres. The ALEPH inclusive analysis [123] reconstructs topological vertices in both event hemispheres as in the inclusive semileptonic analysis. The flavour tagging is performed by computing the product of the jet charges in the two hemispheres of each event. This product thus combines production and decay flavour tags and is sensitive to whether mixing occurred or not.

Table 4.5 summarizes the different B_d^0 mixing analyses. It should be noted that this Table provides only an approximate representation of the performance of each analysis. The reader is referred to the specific papers for more detailed comparisons.

3.4.2. B_s^0 *and* $\overline{B}_s^0$ *oscillation analyses*

The study of time dependent B_s^0 oscillations has been performed with a wide range of analysis techniques at high energy colliders. The study of B_s^0 oscillations is more challenging than that of B_d^0 oscillations due to two main differences. Only about 10% of b quarks hadronize into B_s^0 mesons, as compared to about 40% into B_d^0 mesons. The B_s^0 oscillation frequency is expected to be at least a factor of 20 larger than that for B_d^0 oscillations. To address this, sophisticated analyses have been developed with an emphasis on lowering the mistag rate, increasing the B_s^0 purity and, especially, improving the proper time resolution, all of which affect the sensitivity to B_s^0 oscillations.

Exclusive methods

Fully exclusive analyses have been performed by ALEPH [124] and DELPHI [98] via the (all charged particles) modes $B_s^0 \to D_s^- \pi^+$, $D_s^- a_1^+$, $\overline{D}^0 K^- \pi^+$, $\overline{D}^0 K^- a_1^+$ (last two for DELPHI only), where the D_s^- and $\overline{D}^0$ are fully reconstructed in several decay modes. The decays $B_s^0 \to D_s^{*-} \pi^+$, $D_s^{*-} a_1^+$ and $D_s^{(*)-} \rho^+$ are also reconstructed by adding one or more photons to the above final states (ALEPH only) or by considering the events falling into the "satellite" mass region below the B_s^0 mass peak.

The number of selected signal decays is small (see Table 4.6) but the method provides excellent proper time resolution for two reasons. As there is no missing particle in the decay (at least for events in the main peak), the B_s^0 momentum is known with good precision and therefore the contribution of the momentum uncertainty to the proper time resolution is small. As a result, unlike all other methods, σ_t does not grow significantly when increasing the proper time t. In addition, the reconstructed channels are two-body or quasi two-body decays, with an opening angle of their decay products which is on average larger than that in multi-body final states; this results in a better accuracy on the B decay length. Despite the limited statistics, this method contributes to the study of B_s^0 oscillations at the highest values of ΔM_s. As detailed in Sec. 3.7., this is the preferred method for future studies of B_s^0 oscillations at hadron colliders.

Semi-exclusive methods

Many analyses have been developed with semi-exclusive methods. B_s^0 decays are partially reconstructed in the modes $\mathrm{B}_s^0 \to \mathrm{D}_s^- \ell^+ \nu_\ell X$ and $\mathrm{B}_s^0 \to \mathrm{D}_s^- h^+ X$, where h represents any charged hadron (or system of several hadrons) and the D_s^- meson decay is either fully or partially reconstructed in the modes $\mathrm{D}_s^- \to \phi\pi^-$, $\mathrm{K}^{*0}\mathrm{K}^-$, $K_s^0\mathrm{K}^-$, $\phi\rho^-$, $\mathrm{K}^{*0}\mathrm{K}^{*-}$, $\phi\pi^-\pi^+\pi^-$, $\phi\ell^-\overline{\nu}$, $\phi h^- X$. Partial reconstruction in $\mathrm{D}_s^- h^+$ modes has the benefit of larger statistics but the $\mathrm{D}_s^- \ell^+ \nu_\ell X$ channel has the advantage of a considerably higher B_s^0 purity, lower mistag rate and higher proper time resolution.

Analyses in the mode $\mathrm{B}_s^0 \to \mathrm{D}_s^- \ell^+ \nu$ have been performed by ALEPH [124], CDF [125], DELPHI [126] and OPAL [127]. Selection of D_s^- decays proceeds as described above. CDF only uses a partial reconstruction of the mode $\mathrm{D}_s^- \to \phi\pi^- X$. Some background suppression (especially from $\mathrm{B} \to \mathrm{D}_s \mathrm{D} X$) is achieved by requiring that the lepton and the D_s^- comes from the same vertex.

The hadronic channel $\mathrm{B}_s^0 \to \mathrm{D}_s^- h^+ X$ has been used by DELPHI [98] and SLD [128]. Fully reconstructed D_s^- decays are selected only in the modes $\mathrm{D}_s^- \to \phi\pi^-$ and $\mathrm{K}^{*0}\mathrm{K}^-$ because of their lower background level. D_s^- candidates are then combined with one or more secondary tracks to form B_s^0 decay candidates. Among B_s^0 decays contributing to the D_s^- signal, approximately 10% have the wrong decay flavour tag due to the process $W^+ \to \mathrm{D}_s^+$ ($b \to c\bar{c}s$ transition). This source of mistag is essentially absent in the semileptonic analyses. Despite lower statistics, the SLD analysis contributes to the B_s^0 oscillation sensitivity at large ΔM_s thanks to its excellent decay length resolution (see Table 4.6).

Semi-inclusive methods

The semi-inclusive lepton method, based on the process $\mathrm{B}_s^0 \to X \ell^+ \nu_\ell$, is the most sensitive method at LEP and has been used by ALEPH [124], DELPHI [126], OPAL [129] and SLD [130]. The principle of the method (see the discussion above in the case of B_d^0 mixing) is to reconstruct the D_s^- inclusively by relying on topological vertexing and kinematical information. Fairly loose criteria are applied to select large event samples, see Table 4.6.

For this method, it is important to reduce the contribution from cascade decays and to increase the B_s^0 purity of the sample (B_s^0 mesons represent about 10% of all b-hadrons produced, see Table 4.4). To enrich the sample in direct B_s^0 semileptonic decays, the following quantities are used: momentum and transverse momentum of the lepton, impact parameters of all tracks in the opposite hemisphere relative to the main event vertex, kaons at primary or secondary vertices in the same hemisphere, and charge of the secondary vertex. Those variables are usually combined in a global discriminant variable. The result of this procedure is to increase the B_s^0 purity by about 30%; the corresponding mistag rate at decay is $\sim$10% or less. The above information, as well as the proper time resolution, is then used on an event-by-event basis. As an example, Fig. 4.11 shows the neural network output distributions sensitive to the $b \to \ell^-$ fraction and the B_s^0 purities in the ALEPH data. The decay length resolution is somewhat worse than in the case of semi-exclusive analyses due to missing or mis-assigned tracks.

Inclusive methods

Fully inclusive methods are sensitive to most B decay modes and, thus, have high efficiency. Such techniques have been developed by DELPHI [122] and SLD [131]. The analyses rely on inclusive topological vertexing to select B decay products and to reconstruct the B decay vertex. The DELPHI analysis is the same as the one described earlier for B_d^0 mixing. A very large data sample is obtained but the mistag rates are high (see Table 4.6). SLD is able to exploit the excellent 3D spatial resolution of its CCD-pixel vertex detector to cleanly separate the charged decay products from secondary (originating directly from the B decay) and tertiary (originating from cascade D decays) vertices. The decay flavour is determined from the charge dipole δQ defined as the distance between secondary and tertiary vertices

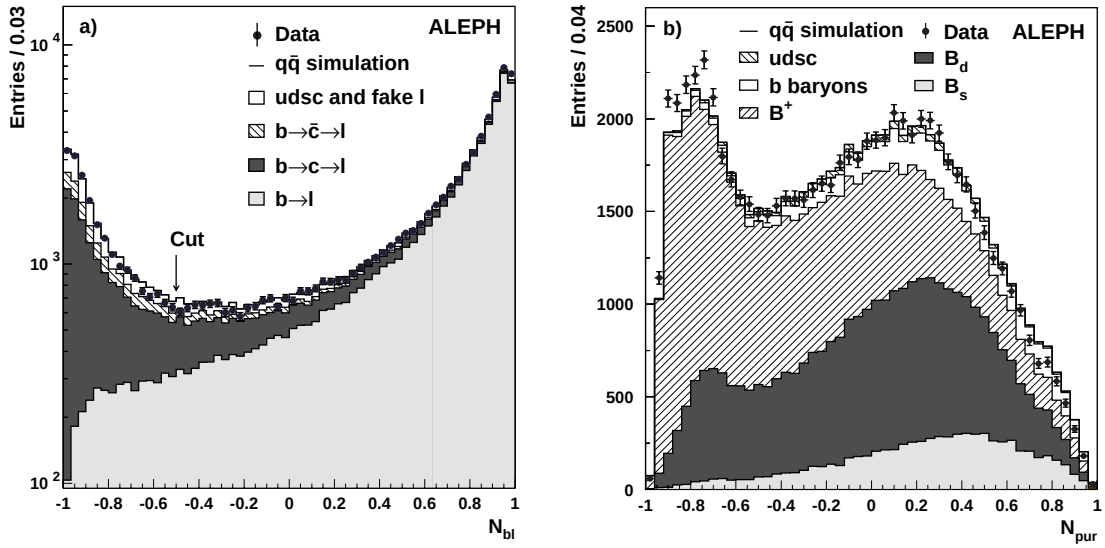

Fig. 4.11: *Distribution of (a) the $b \to \ell^-$ tagging variable and (b) the B_s^0 purity variable for data (points) and Monte Carlo simulation (histograms) in the ALEPH inclusive lepton analysis.*

signed by the charge difference between them. Positive (negative) values of δQ tag $\overline{\mathrm{B}^0}$ (B^0) decays as shown in Fig. 4.12.

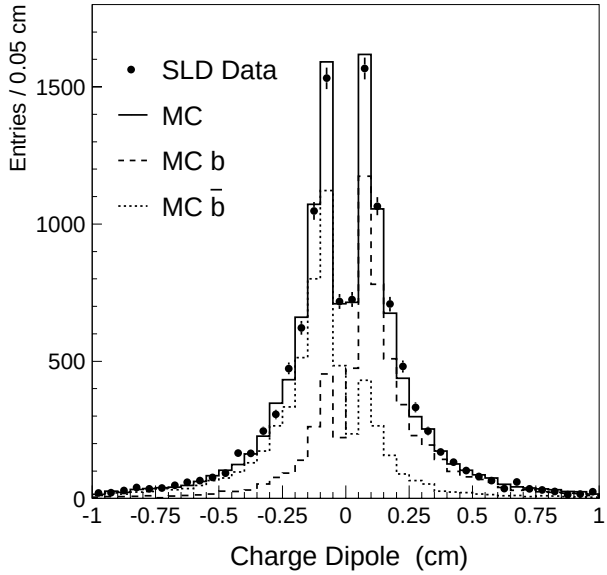

Fig. 4.12: *Distribution of the charge dipole for SLD data (points) and Monte Carlo (solid histogram). Also shown are the contributions from hadrons containing a b quark (dashed histogram) or a $\bar{b}$ quark (dotted histogram).*

Table 4.6 summarizes the different B_s^0 mixing analyses. It should be noted that the Table presents only the average performance of the analyses and that most analyses substantially increase their sensitivity by relying on event-by-event information.

Decay modes	Analysis	Events/Signal	$f_{\rm mode}$	p_W	σ_L	σ_p/p
$\mathrm{B}_s^0 \to \mathrm{D}_s^{(*)-} h^{+a}$	ALEPH [124]	80/29	36%	0	180 μm	0.005 (peak)
	exclusive					0.03 (satellite)
$\mathrm{B}_s^0 \to \mathrm{D}_s^{(*)-} h^{+a}$	DELPHI [98]	44/23	52%	0	117 μm (58%)	[b]
$\bar{\mathrm{D}}^0 \mathrm{K}^- h'^+$	exclusive				216 μm (42%)	
$\mathrm{B}_s^0 \to \mathrm{D}_s^- X$	DELPHI [98]	3079/1266	50%	10%	260 μm (77%)	0.10 (77%) [c]
	semi-excl.				304 μm (13%)	0.26 (23%)
					650 μm (10%)	
	SLD [128]	361/174	55%	10%	50 μm (60%)	0.08 (60%)
	semi-excl.				151 μm (40%)	0.18 (40%)
$\mathrm{B}_s^0 \to \mathrm{D}_s^- \ell^+ \nu$	ALEPH [124]	333/156	47%		240 μm	0.11
	semi-excl.					
	CDF [125]	/1068	61%			
	semi-excl.					
	DELPHI [126]	/436	53%		200 μm (82%)	0.07 (82%)
	semi-excl.				740 μm (16%)	0.16 (16%)
	OPAL [127]	244/116	48%		500 μm	0.10
	semi-excl.					
$\mathrm{B}_s^0 \to X \ell^+ \nu$	ALEPH [124]	74k/	10%	13% [d]	251 μm (75%)	0.064 (60%)
	semi-incl.				718 μm (25%)	0.020 (40%)
	DELPHI [126]	68k/	10%	8-18%		
	semi-incl.					
	OPAL [129]	53k/	8%	12% [d]		
	semi-incl.					
	SLD [130]	2k/	16%	4%	55 μm (60%)	0.06 (60%)
	semi-incl.				217 μm (40%)	0.18 (40%)
$\mathrm{B}_s^0 \to$ all	DELPHI [122]	770k/	10%	43% [e]	400 μm	0.15
	inclusive			33% [f]		
	SLD [131]	11k/	16%	22%	78 μm (60%)	0.07 (60%)
	inclusive				304 μm (40%)	0.21 (40%)

[a] h^+ stands for π^+, ρ^+, a_1^+ and h'^+ stands for π^+, a_1^+.

[b] For the best data subset (B_s^0 peak and 1994-95 data).

[c] Evaluated at $t = 1$ ps for the best subset of data.

[d] Fraction of non-$(b \to \ell^-)$ decays.

[e] For 615k vertices with charge dipole tag.

[f] For 155k vertices with soft lepton tag.

Table 4.6: *Summary of* B_s^0 *mixing analyses showing the signal decay modes, analysis method, total number of selected events and estimated signal, fraction of signal mode in the selected sample* $f_{\rm mode}$*, decay flavour mistag rate* p_W *for* B_s^0 *decays, decay length and momentum resolutions. For semi-exclusive analyses, the number of signal events corresponds to the number of* D_s^- *signal decays (not the number of signal events in the selected decay mode) and* $f_{\rm mode}$ *represents the fraction of* B_s^0 *in the* D_s^- *signal.*

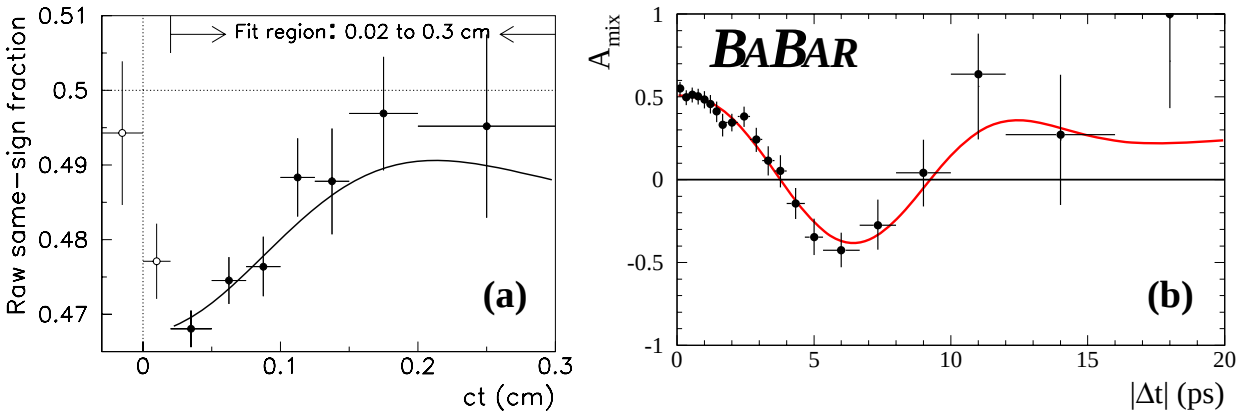

Fig. 4.13: *Examples of ΔM_d results from (a) CDF (Ref. [118]) and (b) BaBar (Ref. [120]). See text for details.*

3.5. B^0_d oscillation results. Measurement of the ΔM_d frequency

As detailed in Sec. 3.4., many methods and channels have been used to study B^0_d–$\overline{B}^0_d$ oscillations. These analyses have been performed by the ALEPH [113,123], BaBar [107,109,120], Belle [108,110,116,121], CDF [89,90,103,111,117,118], DELPHI [98,114,122], L3 [119], OPAL [112,115,100,101] and SLD [92–95] collaborations.

In the following, we will discuss the results of a few representative measurements of ΔM_d. Fig. 4.13(a) showss the fraction of mixed events as a function of proper decay length for a semi-inclusive analysis at CDF using a lepton sample with an inclusively reconstructed vertex combined, on the opposite side, with a lepton and jet charge tag to infer the production flavour [118]. Although this analysis is based on about 240,000 events, the total height of the oscillation amplitude is small (~ 0.05) due to an effective tagging efficiency of $\varepsilon(1-2p_W)^2 \sim 1\%$ for each tag yielding a value of $\Delta M_d = (0.500 \pm 0.052 \pm 0.043)$ ps^{-1}. In this analysis, a large mistag rate p_W resulting in $(1-2p_W)$ being small is compensated by the number of events N being large (see Eq. (63)). This result can be compared to a measurement from BaBar [120] based on about ~ 6300 neutral B mesons fully reconstructed in multihadronic modes (mainly $B^0_d \to \bar{D}^{(*)}X$). An opposite lepton and kaon tag with low mistag fractions of $p_W \sim 8\%$ and $\sim 16\%$, respectively, are the reason for an oscillation amplitude of ~ 0.5 in the mixed asymmetry as shown in Fig. 4.13(b). Note the statistical error on the ΔM_d value obtained by BaBar for this analysis: $\Delta M_d = (0.516 \pm 0.016 \pm 0.010)$ ps^{-1}. From this example we can see the trade-off between a poor tagging power in high statistics B samples produced for example in a hadronic $p\bar{p}$ environment at the Tevatron and lower statistics analyses with superior tagging and low mistag probabilities in an e^+e^- environment for example at the B factories. In addition, compared to inclusive methods, analyses with fully reconstructed B mesons have a higher sample purity.

Fig. 4.14 shows the result of two other ΔM_d analyses. One of the most precise single measurements performed at the Z^0 resonance is an inclusive D^* analysis by OPAL [115] using $B^0 \to D^{*-}\ell^+\nu$ decays. High statistics $D^{*-} \to \bar{D}^0\pi^-$ decays were reconstructed using the slow π^- from the D^{*-} decay while inferring the $\overline{D}^0$ with an inclusive technique. Same-sign lepton-pion pairs serve to constrain the combinatorial background in the opposite sign lepton-pion pair signature. A clear oscillation signal is observed in the fraction of mixed events as can be seen in Fig. 4.14(a). A value of $\Delta M_d = (0.497 \pm 0.024 \pm 0.025)$ ps^{-1} is extracted. Another example of a precise ΔM_d analysis at the Z^0 pole by DELPHI is shown in Fig. 4.14(b). A sample of 770,000 events with an inclusively reconstructed vertex has been selected. Tags based on several separating variables such as the jet charge, dipole charge and the transverse momentum of the (soft) lepton have been combined into a probability to determine the fraction of like-sign events as displayed in Fig. 4.14(b). DELPHI obtains a value of $\Delta M_d = (0.531 \pm 0.025 \pm 0.007)$ ps^{-1}.

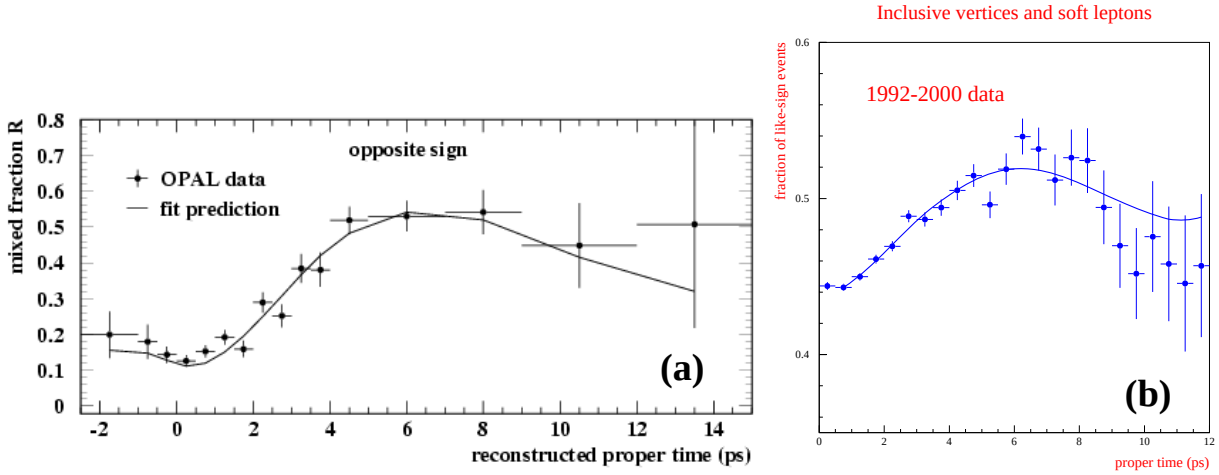

Fig. 4.14: *Examples of ΔM_d results from (a) OPAL (Ref. [115]) and (b) DELPHI (Ref. [122]). See text for details.*

In order to combine all individual ΔM_d results to obtain a world average value, possible statistical correlations between individual analyses have to be taken into account and also the systematic errors which are often not negligible have to be combined properly. The main sources of systematic uncertainties are determinations of sample compositions, mistag probabilities, b hadron production fractions and contributions from b hadron lifetimes. Before being combined, the measurements are adjusted on the basis of a common set of input values. Details of the averaging procedure are described in Ref. [85].

A compilation of all ΔM_d measurements available as of the 2002 ICHEP conference, can be found in Fig. 4.15. The individual results from each experiment are combined and averaged using the procedure described above. There exist also time-integrated measurements of B^0_d mixing from the ARGUS [132,133] and CLEO [134,135] collaborations which can be converted into a value for ΔM_d assuming the width difference $\Delta\Gamma_d$ in the B^0_d system to be zero and no CP violation in B^0_d mixing. The quoted world average, at the bottom of Fig. 4.15, also includes χ_d measurements by ARGUS and CLEO. The ΔM_d averages per experiment are displayed in Fig. 4.16.

The different results from the combination procedure are [87]:

$$\Delta M_d = \begin{cases} (0.491 \pm 0.041)\ \mathrm{ps}^{-1} & \text{Argus-CLEO (from } \chi_d\text{)} \\ (0.498 \pm 0.013)\ \mathrm{ps}^{-1} & \text{LEP-SLD-CDF} \\ (0.503 \pm 0.007)\ \mathrm{ps}^{-1} & \text{Belle-BaBar} \\ (0.503 \pm 0.006)\ \mathrm{ps}^{-1} & \text{world average} \end{cases} . \tag{64}$$

At the end of the LEP-CDF-SLD era, ΔM_d has been determined with a relative precision of about 2.6%. The LEP-CDF-SLD results are in excellent agreement with the Belle-BaBar measurements. After the inclusion of the results from B factories, the precision on ΔM_d is improved by a factor of two. The world average B^0_d mixing frequency is now dominated by the results of B factories.

3.6. Results on B^0_s oscillations. Limits on the ΔM_s frequency

B^0_s–$\overline{\mathrm{B}}^0_s$ oscillations have also been the subject of many studies by ALEPH [96,97,124], CDF [125], DELPHI [98,99,126,136], OPAL [127,129] and SLD [102,128,130,131]. No oscillation signal has been observed to date. To set lower limits on the oscillation frequency ΔM_s, all B^0_s mixing analyses use the amplitude method [106] described in Sec. 3.3.1.

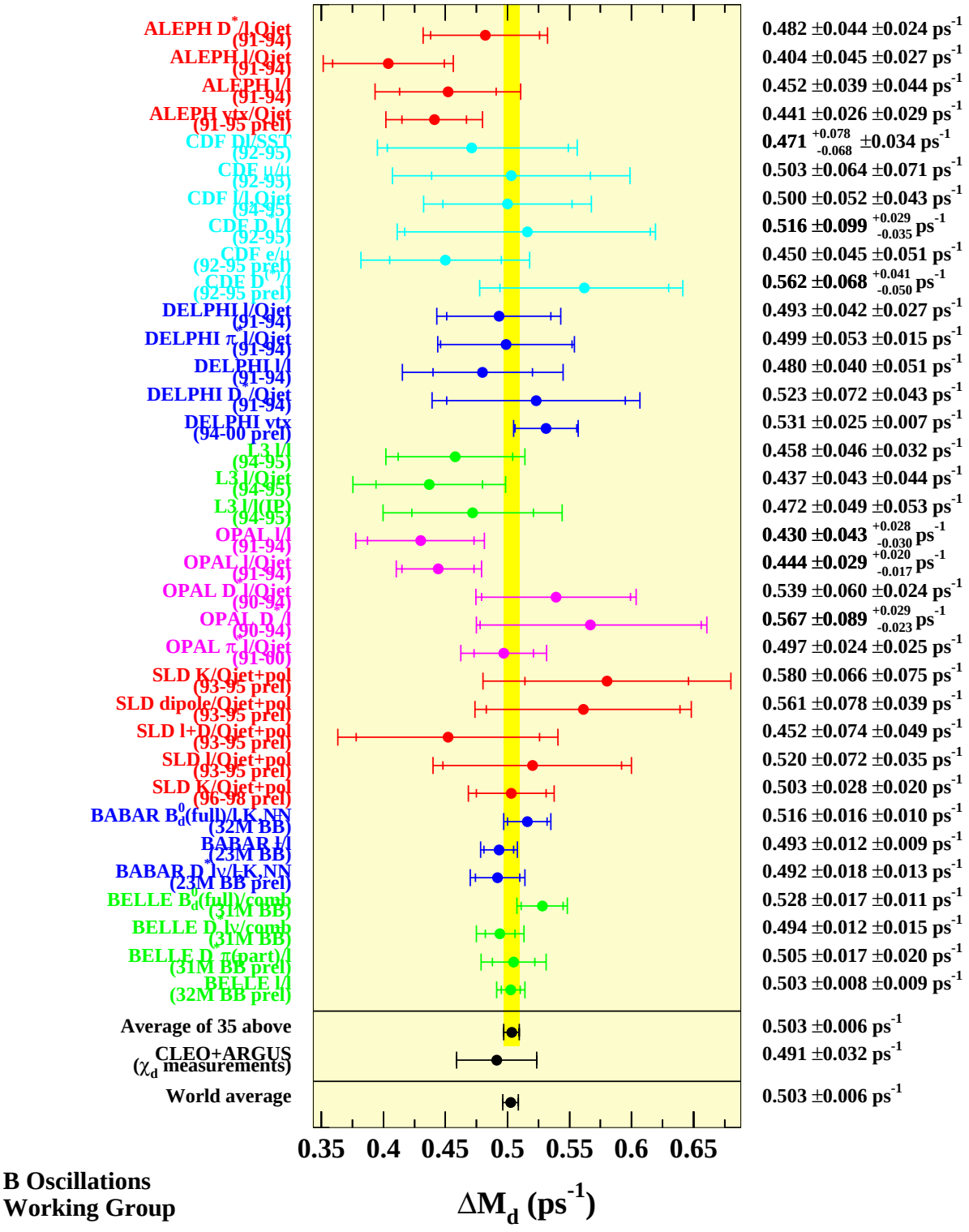

Fig. 4.15: *Individual and combined measurements of* ΔM_d *at* B *factories, LEP, SLD and CDF as of the ICHEP 2002 conference [87]. The quoted world average, at the bottom, also includes* χ_d *measurements performed by ARGUS and CLEO.*

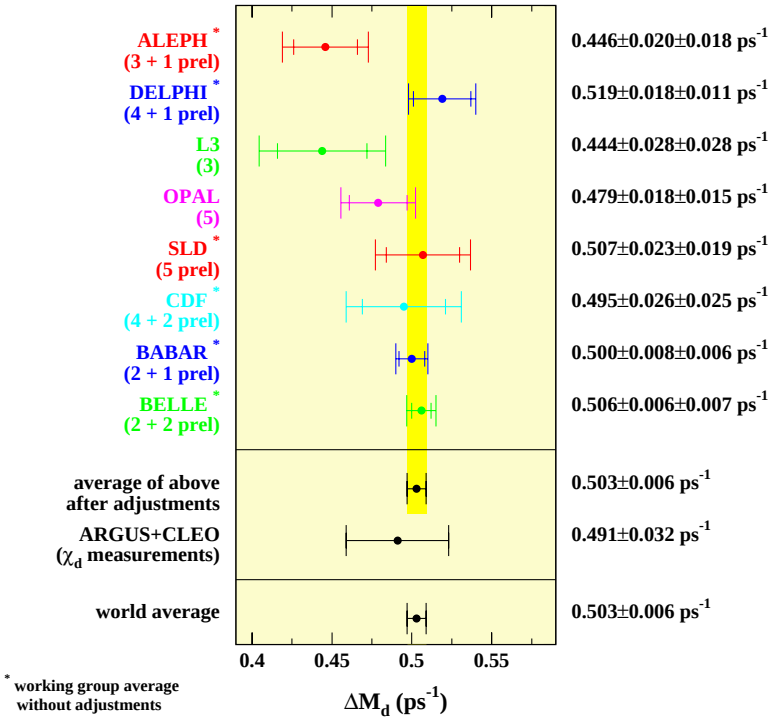

Fig. 4.16: *Combined measurements of ΔM_d averaged by experiment as of the ICHEP 2002 conference [87]. The quoted world average, at the bottom, also includes χ_d measurements by ARGUS and CLEO.*

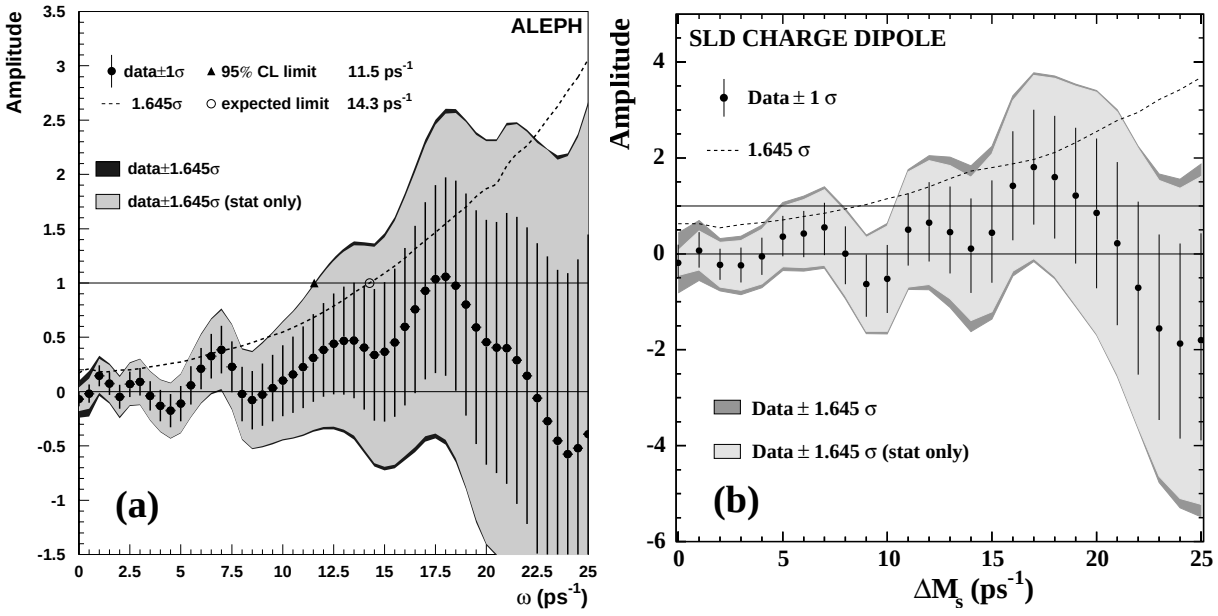

Fig. 4.17: *Examples of measured B_s^0 oscillation amplitudes as a function of the mixing frequency ΔM_s from (a) ALEPH (Ref. [124]) and (b) SLD (Ref. [131]). See text for details.*

Two examples of measured B_s^0 oscillation amplitudes as a function of the mixing frequency ΔM_s are shown in Fig. 4.17. The ALEPH collaboration recently presented an improved search for B_s^0 oscillations combining three analyses based on different final states [124]. First, fully reconstructed decays of B_s^0 mesons yield a small sample of B_s^0 candidates with excellent decay length and momentum resolution. Semileptonic decays with a reconstructed D_s^+ meson provide a second sample with larger statistics, high B_s^0 purity but with a poorer momentum and decay length resolution due to the partial decay reconstruction. Finally, semileptonic B hadron decays are inclusively selected and yield the data sample with the highest sensitivity to B_s^0 oscillations since the higher statistics compensates for the low average B_s^0 purity and the poorer proper time resolution. Fig. 4.17(a) shows the fitted amplitude spectrum as a function of ΔM_s for the third sample. From this inclusive semileptonic sample alone, ALEPH excludes all frequencies below 11.5 ps^{-1}, while the combined 95% C.L. limit from all three analyses yields $\Delta M_s > 10.9$ ps^{-1}.

Fig. 4.17(b) shows the amplitude spectrum from an analysis by SLD [131]. This analysis determined the B flavour at production time by exploiting the large forward-backward asymmetry of polarized $Z^0 \to b\bar{b}$ decays and uses additional information from the hemisphere opposite to that of the reconstructed B decay such as the jet charge, the lepton and kaon tags. The B flavour at decay is tagged by a charge dipole method as explained in Sec. 3.4.2. Although this analysis is based on only 11,000 decays, it reaches a sensitivity of 8.8 ps^{-1} because of the slower rise of the uncertainty on the amplitude due to the excellent proper time resolution.

No B_s^0 oscillation signal has been seen so far. The most sensitive analyses are the ones based on the inclusive lepton samples at LEP. Because of better proper time resolution, smaller data samples of inclusive decays analyzed at SLD as well as measurements using only a few fully reconstructed B_s^0 decays at LEP, turn out to be very useful to explore the high ΔM_s region. This point is illustrated in Fig. 4.18(a) showing the ΔM_s sensitivities for the different B_s^0 oscillation analysis methods. The uncertainty on the amplitude A (actually $1.645\,\sigma_A$) is plotted as a function of ΔM_s combining the existing results of the various B_s^0 analyses methods from different experiments. The combination of all fully inclusive methods crosses the dashed line corresponding to the condition $1.645\,\sigma_A = 1$ used to define the 95% C.L. sensitivity at about 9.5 ps^{-1}. This represents the combined sensitivity of all inclusive methods from the various experiments. Due to the combination of high statistics and adequate vertexing resolution, the inclusive lepton methods give currently the most sensitive results. The D_s^+-lepton samples also reach a high sensitivity while the exclusive methods that attempt to fully reconstruct hadronic B_s^0 decays have a lower sensitivity because of the small number of B_s^0 candidates that have been exclusively reconstructed to date. However, the slow growth of the amplitude error for the exclusive method can be inferred from Fig. 4.18(a). Note, the visible scattering of points for the exclusive method which results from the small number of events contributing in these analyses.

All available results on ΔM_s oscillations can be combined into a world average exclusion limit using the amplitude method. All data on the measurements of B_s^0 oscillation amplitudes versus ΔM_s, as provided by the experiments, are averaged to yield a combined amplitude A as a function of ΔM_s as shown in Fig. 4.18(b). The individual results have been adjusted to common physics inputs and all known correlations have been accounted for. The sensitivities of the inclusive analyses which depend on the assumed fraction, $f_{B_s^0}$, of B_s^0 mesons have been re-scaled to a common average of $f_{B_s^0} = 0.093 \pm 0.011$ (see Table 4.4). Figure 4.18(b) includes all results as of the ICHEP 2002 conference. The measurements are dominated by statistical uncertainties. Neighbouring points are statistically correlated. The combined result is [87]:

$$\boxed{\begin{array}{c}\Delta M_s > 14.4\ \mathrm{ps}^{-1}\ \text{at}\ 95\%\mathrm{C.L.} \\ \text{with a sensitivity of}\ \Delta\mathrm{M_s} = 19.2\ \mathrm{ps}^{-1}\end{array}} \tag{65}$$

Values between $14.4\,\mathrm{ps}^{-1}$ and $\sim 22\,\mathrm{ps}^{-1}$ cannot be excluded because the data appear to be compatible with a signal in this region. The amplitude plot presents a deviation from $A = 0$ at about $\Delta M_s \sim 17.5$ ps^{-1} for which a significance of $\sim 2.2\ \sigma$ can be derived. This means that there is not enough

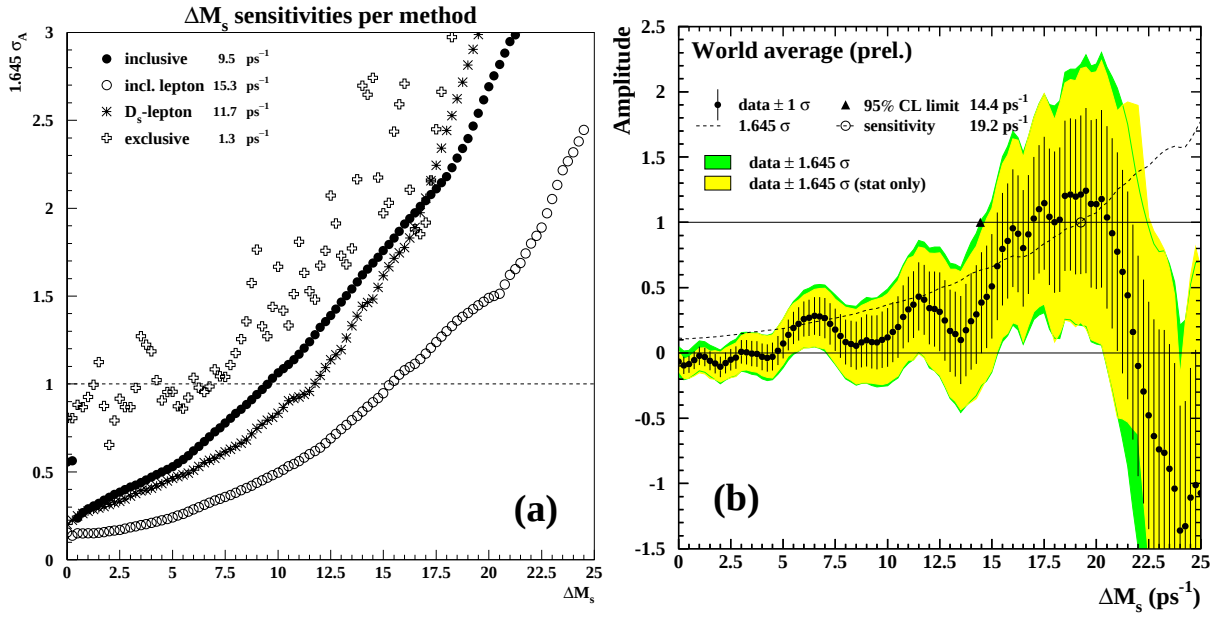

Fig. 4.18: *(a) Uncertainty ($1.645\,\sigma_A$) on the amplitude A as a function of ΔM_s for the various B_s^0 oscillation analyses. The dashed line corresponds to the condition $1.645\,\sigma_A = 1$ used to define the 95% C.L. sensitivity. (b) Combined measurements of the B_s^0 oscillation amplitude as a function of ΔM_s, including all results as of the ICHEP 2002 conference. Neighbouring points are statistically correlated.*

sensitivity for the observation of a B_s^0–$\overline{\mathrm{B}}_s^0$ signal at this frequency.

The different measurements of the B_s^0 oscillation amplitude as of the ICHEP 2002 conference are shown in Fig. 4.19, where the amplitudes for the various analyses are given at $\Delta M_s = 15$ ps^{-1} along with the relevant statistic and systematic errors. The exclusion sensitivities are also indicated Fig. 4.19 shows which analyses contribute most in the high ΔM_s region. Note that the individual measurements are quoted as in the original publications, but the averages include the effects of adjustments to a common set of input parameters.

Although all B_s^0 mixing results are presently limited by statistics, a discussion of systematic uncertainties in these analyses is relevant for a future measurement of B_s^0 oscillations. Critical analysis parameters (σ_L, σ_p and p_W) are extracted from detailed Monte Carlo simulation and are subject to modelling uncertainties. A first level of control is typically achieved with detailed comparisons between data and MC. In addition, measurements from calibration samples are performed to cross-check the parameters directly from the data but not all critical parameters can be tested in this manner. Of particular importance to the sensitivity at large ΔM_s values is the proper time resolution and, in particular, the decay length resolution. The latter has been tested with a variety of techniques: fit to the decay length distribution of τ decays, fit for the primary vertex in Z^0 decays to light-flavour quarks, study of tracks with negative impact parameter. These studies find that the decay length resolution is typically understood at the 10% level or better.

3.7. Future prospects for ΔM_d and ΔM_s determination

The current world average B_d^0 oscillation frequency constitutes a measurement at about 1% precision. It is dominated by the results of the B factories which will further improve the precision on Δm_d. The uncertainty on the B_d^0 lifetime starts to become a main contributor to the systematic error on future measurements of ΔM_d. A simultaneous fit of the B lifetime and ΔM_d will improve this situation as

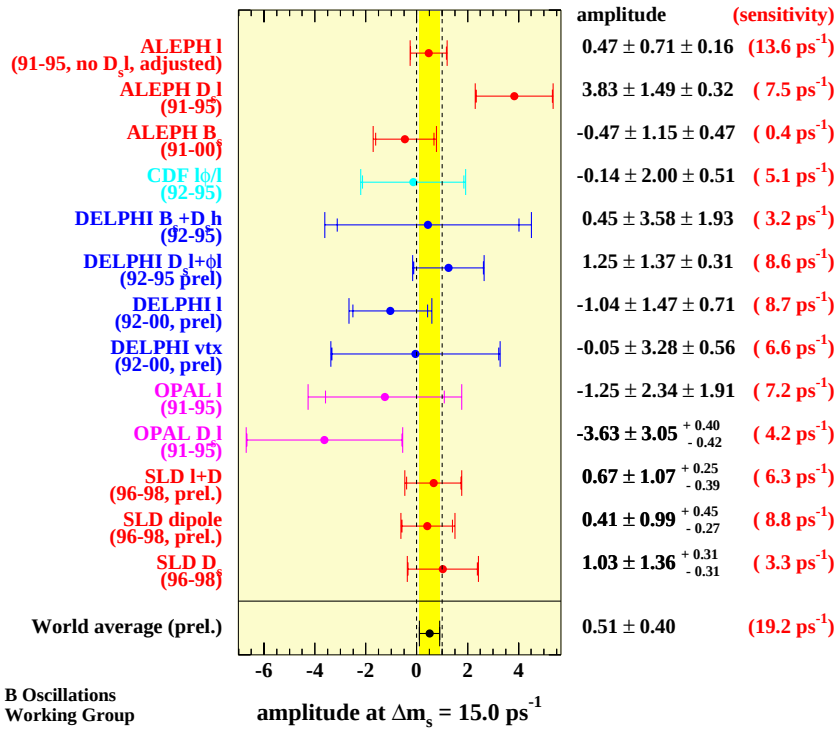

Fig. 4.19: *Measurements of the B_s^0 oscillation amplitude as of the ICHEP 2002 conference. The amplitudes are given at $\Delta M_s = 15\ ps^{-1}$ along with the relevant statistical and systematic errors. The exclusion sensitivities are indicated on the right, within in parentheses, The shaded area indicates the $\pm$ 1 σ region on the average, and the dashed lines correspond to the values 0 and 1.*

demonstrated in Ref. [107]. For a data sample of 300 fb^{-1}, the BaBar and Belle experiments expect to improve the B_d^0 oscillation frequency by a factor two, down to a precision of about 0.4%

The future interest in B mixing clearly lies in a measurement of B_s^0 oscillations. Some of the still preliminary analyses from LEP and SLD are in the process of being finalized for publication while no new measurements or improved limits are to be expected. Since no B_s^0 mesons are produced at the B factories and running at the $\Upsilon(5S)$ resonance as a source of B_s^0 mesons is not foreseen in the near future, the hopes of the heavy flavour community focus on the Tevatron Collider experiments CDF and DØ to measure B_s^0 oscillations. For such a measurement it is important that the resolution of the vertexing device is good enough to resolve the expected (rapid) oscillations while a small boost correction will prevent the measured oscillations to damp out with proper time. The path to measure B_s^0 oscillations is therefore to use fully reconstructed B_s^0 mesons rather than higher statistics samples of partially reconstructed B_s^0 candidates from e.g. semileptonic decays.

A measurement of ΔM_s will be the next crucial test of the Standard Model probing whether the obtained result will fit to the current constraints on the CKM triangle which are all in beautiful agreement (see results in Chapter 5). It is noteworthy to mention that physics with B_s^0 mesons is unique to the Tevatron until the start of the LHC in 2007.

3.7.1. CDF and DØ detector upgrades in Run II at Tevatron

The Fermilab accelerator complex has undergone a major upgrade since the end of Run I in 1996. The centre-of-mass energy has been increased to 1.96 TeV and the Main Injector, a new 150 GeV proton storage ring, has replaced the Main Ring as injector of protons and anti-protons into the Tevatron. The Main Injector also provides higher proton intensity onto the anti-proton production target, with the goal

to allow for more than an order of magnitude higher luminosities. Run II officially started in March 2001. The design luminosity during the first phase of Run II (Run IIa) is 5-8$\cdot 10^{31}$ cm^{-2}s^{-1} for a final integrated luminosity of ~ 2 fb^{-1} by the end of Run IIa.

Since 1996, the CDF and DØ detectors have also undergone major upgrades [138,139] to allow operation at high luminosities and bunch spacing of up to 132 ns. Details of the DØ detector upgrade can be found elsewhere [139]. The main upgrade for DØ is the installation of a tracking system contained in a 2T superconducting solenoid surrounded by a scintillator preshower detector. The tracking upgrade includes a silicon microstrip tracker which consists of six barrel segments with disks in between plus three more disks located at each end of the tracker. In addition, there are two large disks placed at the end of the silicon tracker to increase the pseudorapidity coverage. The silicon system is enclosed within a central fiber tracker providing momentum resolution at the level of $\sigma(p_T)/p_T = 0.02$-0.05 for low-p_T tracks with high tracking efficiency for charged particles with pseudo-rapidity $\eta < 2.5$. Vertex reconstruction is expected with a resolution of 15-30 μm in the $r\phi$-plane and about 80 μm in the rz-plane. A major upgrade of the muon system together with central and forward scintillators will allow DØ to trigger and reconstruct muon tracks. The B physics triggers at DØ allow to trigger on muons and electrons while a new Level 1 tracking trigger and a Level 2 silicon trigger are under construction.

The CDF detector improvements for Run II [138] were motivated by the shorter accelerator bunch spacing of up to 132 ns and the increase in luminosity by an order of magnitude. All front-end and trigger electronics has been significantly redesigned and replaced. A DAQ upgrade allows the operation of a pipelined trigger system. CDF's tracking devices were completely replaced. They consist of a new Central Outer Tracker (COT) with 30,200 sense wires arranged in 96 layers combined into four axial and four stereo superlayers. It also provides dE/dx information for particle identification. The Run II silicon vertex detector, covering a total radial area from 1.5-28 cm, consists of seven double sided layers and one single sided layer mounted on the beampipe. The silicon vertex detector covers the full Tevatron luminous region which has a RMS spread of about 30 cm along the beamline and allows for standalone silicon tracking up to a pseudo-rapidity $|\eta|$ of 2. The forward calorimeters have been replaced by a new scintillator tile based plug calorimeter which gives good electron identification up to $|\eta| = 2$. The upgrades to the muon system almost double the central muon coverage and extent it up to $|\eta| \sim 1.5$.

3.7.2. Prospects for B_s^0 mixing at CDF

The most important improvements for B physics at CDF are a Silicon Vertex Trigger (SVT) and a Time-of-Flight (ToF) system with a resolution of about 100 ps. The later employs 216 three-meter-long scintillator bars located between the outer radius of the COT and the superconducting solenoid. More details about the CDF II Time-of-Flight detector and its performance can be found in Ref. [140,141]. The ToF system will be most beneficiary for the identification of kaons with a 2 σ-separation between π and K for $p < 1.6$ GeV/c. This will enable CDF to make use of opposite side kaon tagging and allows to identify same side fragmentation kaons accompanying B_s^0 mesons [140,141].

In Run I, all B physics triggers at CDF were based on leptons including single and dilepton triggers. A newly implemented Silicon Vertex Trigger gives CDF access to purely hadronic B decays and makes CDF's B physics program fully competitive with the one at the e^+e^- B factories. The hadronic track trigger is the first of its kind operating successfully at a hadron collider. It works as follows: with a fast track trigger at Level 1, CDF finds track pairs in the COT with $p_T > 1.5$ GeV/c. At Level 2, these tracks are linked into the silicon vertex detector and cuts on the track impact parameter (e.g. $d > 100\ \mu$m) are applied. The original motivation for CDF's hadronic track trigger was to select the two tracks from the rare decay $\mathrm{B}^0 \to \pi\pi$ but it will play a major role in collecting hadronic B_s^0 decays for the measurement of B_s^0 oscillations. Since the beginning of Run II, much work has gone into commissioning the CDF detector. The Silicon Vertex Trigger was fully operational at the beginning of 2002. A detailed discussion of the SVT and its initial performance can be found elsewhere [142,143].

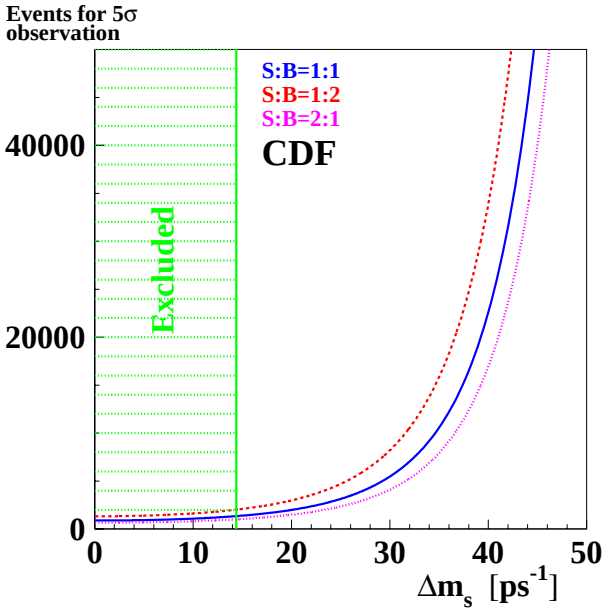

Fig. 4.20: *Expected event yield of fully reconstructed* B_s^0 *decays at CDF necessary for a* $5\,\sigma$*-observation of* B_s^0 *oscillations as a function of* ΔM_s *for different signal-to-background ratios.*

The CDF detector upgrades described above play an important role in CDF's prospects for measuring B_s^0 mixing. The inner layer of silicon mounted on the beampipe improves the time resolution for measuring the B_s^0 decay length from originally $\sigma_t = 0.060$ ps to 0.045 ps. This will be important if ΔM_s is unexpectedly large. The Time-of-Flight system will enhance the effectiveness of B flavour tagging, especially through same side tagging with kaons and opposite side kaon tagging, to a total expected $\varepsilon\mathcal{D}^2 \sim 11.3\%$ [144,141].

Fig. 4.20 shows the expected event yield of fully reconstructed B_s^0 decays necessary for a $5\,\sigma$ observation of B_s^0 oscillations as a function of the mixing frequency ΔM_s for different signal-to-background ratios. If the B_s^0 mixing frequency is around the current Standard Model expectation of $\Delta M_s \sim 18\mathrm{ps}^{-1}$ (see discussion in Sec. 6. of Chapter 5), Fig. 4.20 indicates that CDF would only need a few thousand fully reconstructed B_s^0 mesons to discover B_s^0 flavour oscillations. Originally, CDF estimated to fully reconstruct a signal of about 75,000 $\overline{\mathrm{B}}_s^0 \to \mathrm{D}_s^+\pi^-$ and $\overline{\mathrm{B}}_s^0 \to \mathrm{D}_s^+\pi^-\pi^+\pi^-$ events from the two-track hadronic trigger in 2 fb^{-1} [144]. This assumes all detector components and triggers work as expected. Although with the beginning of 2002, the CDF detector is in stable running conditions operating with reliable physics triggers, including the hadronic two-track trigger, there appear to be indications that the projected event yield might be overestimated. Given this and the small amount of data delivered by the Tevatron and recorded by CDF to date (about 100 pb^{-1} by the end of 2002) it will take some time until CDF can present first results on B_s^0 mixing [145].

3.7.3. Prospects for B_s^0 *mixing at DØ*

The major difference for a search of B_s^0 oscillations at DØ is the collection of B_s^0 candidate events. DØ currently does not operate a hadronic track trigger. However, it will be able to collect B_s^0 candidate events using lepton triggers. Various B_s^0 decay modes such as $\overline{\mathrm{B}}_s^0 \to \mathrm{D}_s^+\pi^-$, $\overline{\mathrm{B}}_s^0 \to \mathrm{D}_s^+\pi^-\pi^+\pi^-$ and $\overline{\mathrm{B}}_s^0 \to \mathrm{D}_s^+\ell^-\nu$ are under investigation by the DØ collaboration. The fully hadronic decay modes can be collected by single lepton triggers where the trigger lepton serves as an opposite side lepton tag and the B_s^0 meson is reconstructed on the other side. In this case the event yield is suppressed leaving DØ with a few thousand events of this type in a data sample of 2 fb^{-1}. If the B_s^0 oscillation frequency is

small enough, semileptonic B_s^0 decays can be used utilizing DØ's lepton trigger data. But due to the escaping neutrino, the boost resolution is reduced limiting the ΔM_s reach. DØ expects to collect about 40,000 events in the semileptonic channel in 2 fb^{-1}. Monte Carlo studies indicate that DØ will be able to measure B_s^0 oscillations in this mode up to a mixing frequency of $\Delta M_s \sim 20$ ps^{-1}.

4. Use of the amplitude spectrum for CKM fits

In this Section we discuss how to include ΔM_s information in CKM fits starting from the amplitude spectrum given by the LEP Oscillation Working Group [87].

The 95% C.L. limit and the sensitivity (see definition in Eq. (63)), are useful to summarize the results of the analysis. However to include ΔM_s in a CKM fit and to determine probability regions for the Unitarity Triangle parameters, continuous information about the degree of exclusion of a given value of ΔM_s is needed. We describe how to include this information in both Bayesian and frequentist approaches. The requirements for an optimal method are:

- the method should be independent 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.

For the discussion in this Section we use the World Average computed by the LEP Oscillation Working Group [87] and presented a the CKM-Workshop, corresponding to a 95% C.L. lower limit at 15.0 ps^{-1} and to a sensitivity at 18.0 ps^{-1}.

In Sec. 4.1. we review and analyse how to include ΔM_s information for the CKM fits. Sec. 4.2. describes the newly-proposed frequentist method for including ΔM_s information in CKM fits.

4.1. Review of the available methods. The likelihood ratio method

Modified χ^2 method

The first CKM fits [146–148] used the χ^2 of the complete amplitude spectrum w.r.t. 1:

$$\chi^2 = \left(\frac{1-A}{\sigma_A}\right)^2 \tag{66}$$

The main drawback of this method is that the sign of the deviation of the amplitude with respect to the value $A = 1$ is not used. A signal might manifest itself by giving an amplitude value simultaneously compatible with $A = 1$ and incompatible with $A = 0$; in fact, with this method, values of $A > 1$ (but still compatible with $A = 1$) are disfavoured w.r.t. $A = 1$, 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 using ΔM_s in CKM fits, by taking $A = 1$ whenever it was in fact higher.

A modified χ^2 has been introduced in [149] to solve the second problem:

$$\chi^2 = 2 \cdot \left[\mathrm{Erfc}^{-1}\left(\frac{1}{2}\mathrm{Erfc}\left(\frac{1-A}{\sqrt{2}\sigma_A}\right)\right)\right]^2 \tag{67}$$

Relation between the log-likelihood and the Amplitude

The log-likelihood values can be easily deduced from A and σ_A using the expressions given in [106]:

$$\Delta \log \mathcal{L}^{\infty}(\Delta M_s) \;=\; \frac{1}{2}\left[\left(\frac{A-1}{\sigma_A}\right)^2 - \left(\frac{A}{\sigma_A}\right)^2\right] = \left(\frac{1}{2} - A\right)\frac{1}{\sigma_A^2}\,, \tag{68}$$

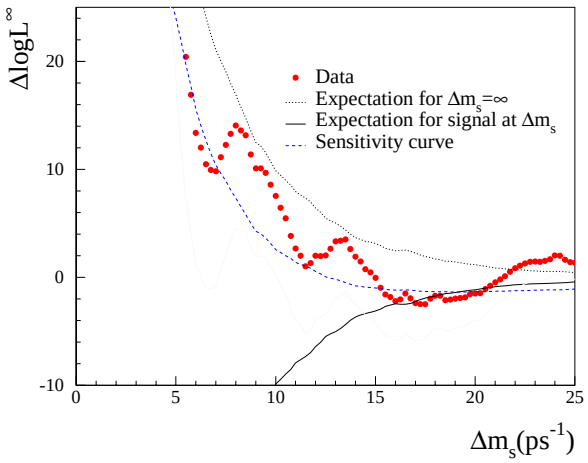

Fig. 4.21: *World average amplitude analysis:* $\Delta \log \mathcal{L}^{\infty}(\Delta M_s)$.

$$\Delta \log \mathcal{L}^{\infty}(\Delta M_s)_{\rm mix} = -\frac{1}{2}\frac{1}{\sigma_A^2}\,, \tag{69}$$

$$\Delta \log \mathcal{L}^{\infty}(\Delta M_s)_{\rm nomix} = \frac{1}{2}\frac{1}{\sigma_A^2}\,. \tag{70}$$

The last two equations give the average log-likelihood value for ΔM_s corresponding to the true oscillation frequency (mixing case) and for ΔM_s being far from the oscillation frequency ($|\Delta M_s - \Delta M_s^{\rm true}| \gg \Gamma/2$, no-mixing case). Γ is here the full width at half maximum of the amplitude distribution in case of a signal; typically $\Gamma \simeq 1/\tau_{\rm B_s^0}$. Fig. 4.21 shows the variation of $\Delta\mathcal{L}^{\infty}(\Delta M_s)$ corresponding to the amplitude spectrum of Fig. 4.18(b).

Likelihood ratio method R

Instead of the χ^2 or the modified χ^2 methods, the log-likelihood function $\Delta \log \mathcal{L}^{\infty}(\Delta M_s)$ can be used: this is the log-likelihood referenced to its value obtained for $\Delta M_s = \infty$ [150,151]. The log-likelihood values can easily be deduced from A and σ_A, in the Gaussian approximation, by using the expressions given in Eqs.(68), (69), (70). The Likelihood Ratio R, defined as,

$$\mathrm{R}(\Delta M_s) = e^{-\Delta \log \mathcal{L}^{\infty}(\Delta M_s)} = \frac{\mathcal{L}(\Delta M_s)}{\mathcal{L}(\Delta M_s = \infty)}\,, \tag{71}$$

has been adopted in [151] to incorporate the ΔM_s constraint.

Comparison between 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 Fig. 4.22-(a) and (b). The constraints obtained using the Likelihood Ratio method (R) and the Modified χ^2 method (χ^2) are shown in Fig. 4.22-(c). In this comparison the Modified χ^2 has been converted to a likelihood using $\mathcal{L} \propto \exp(-\chi^2/2)$. It is clear that the two methods (R and χ^2) give very different constraints. In particular the Modified χ^2 method, with the present World Average, corresponds to a looser constraint for CKM fits (and in particular for the determination of the $\bar{\rho}$ and γ parameters).

The toy Monte Carlo

In order to test and compare the statistical properties of the two methods it is necessary to generate several experiments having similar characteristics as the data used for the World Average. We will call

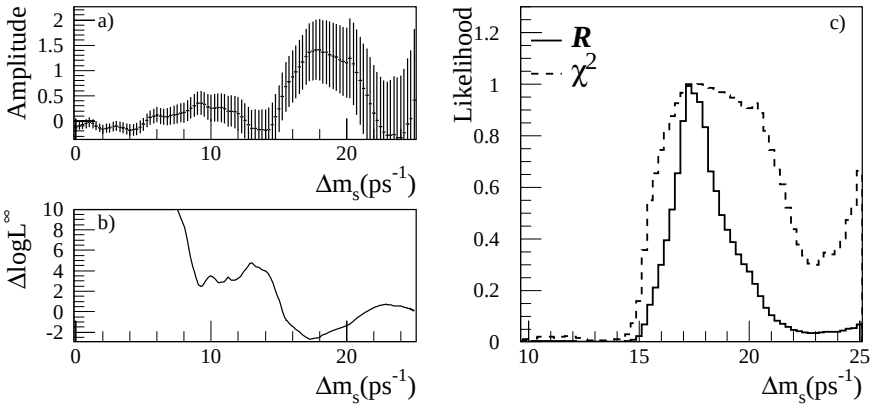

Fig. 4.22: *World Average amplitude analysis: (a) amplitude spectrum, (b) $\Delta \log \mathcal{L}^{\infty}(\Delta M_s)$, (c) comparison between the Likelihood Ratio method (R) and the Modified χ^2 method (χ^2). The information in (b) and in the solid histogram in (c) is identical.*

equivalent those experiments having the same dependence of σ_A as a function of ΔM_s.

The dependence of σ_A 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 [152]. The error on the amplitude can be written as:

$$\sigma_A^{-1} = \sqrt{N}\,\eta_{\mathrm{B}_s^0}\,(2\epsilon_d - 1)\,(2\epsilon_p - 1)\,W(\sigma_L, \sigma_P, \Delta M_s) \tag{72}$$

where N is the total number of events, $\eta_{\mathrm{B}_s^0}$ the B_s^0 purity of the sample, $\epsilon_{d(p)}$ the tagging purity at the decay (production) time, σ_L the uncertainty on the B_s^0 flight length and σ_p the relative uncertainty in the B_s^0 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 multiplies the W function are obtained by adjusting the simulated error distribution to the one measured with real events. Figure 4.23 shows the agreement between the toy-MC calculation and the real data up to $\Delta M_s = 25\,\mathrm{ps}^{-1}$ (the upper value of ΔM_s at which amplitudes are given). An additional problem is that, in principle, one would like to define the likelihood within the interval $[0, \infty]$ whereas the amplitude spectrum is measured only up to a certain value. For the present World Average the value is $25\,\mathrm{ps}^{-1}$. A procedure has to be introduced to continue σ_A and A.

The continuation for σ_A is shown in Fig.4.23. The continuation of A is more delicate. In particular it is more sensitive to the real amplitude spectrum. Nevertheless if $\Delta M_s^{\mathrm{sens}} << \Delta M_s^{\mathrm{last}}$, the significance S ($S = A/\sigma_A$) is approximately constant. It is then a good approximation to continue using:

$$A(\Delta M_s) = \frac{A(\Delta M_s^{\mathrm{last}})}{\sigma_A(\Delta M_s^{\mathrm{last}})}\sigma_A(\Delta M_s). \tag{73}$$

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 remains stable.

Comparison of the methods in case of an oscillation signal

In this Section we compare the two methods in the presence of a clear ΔM_s oscillation signal. We perform several ΔM_s toy-MC analyses with the same σ_A versus ΔM_s behaviour as the World Average analysis. For this study we have generated a ΔM_s signal at $17\,\mathrm{ps}^{-1}$. This value corresponds to the value

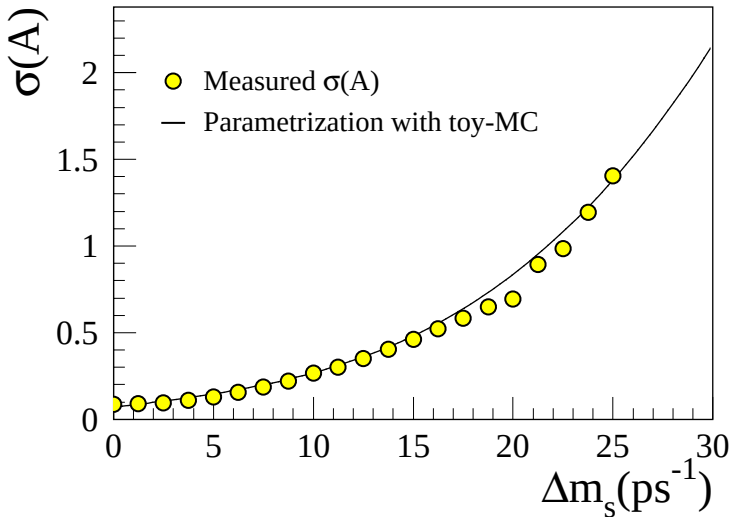

Fig. 4.23: *Comparison between the error distribution computed with the toy-MC (solid line) and the measured amplitude errors (circles).*

where there is the bump in the World Average amplitude spectrum. The statistics of the virtual experiments is much larger than the registered data, at present, so that clear oscillation signals are expected.

The results in Fig. 4.24 show that only the Likelihood Ratio method is able to see the signal at the correct ΔM_s value. The same exercise has been repeated for different generated values of ΔM_s, always giving the same result.

Test of the coverage of the two methods applied to CKM fits

In the absence of a clear B_s^0 oscillation signal, the Likelihood Ratio method results in a ΔM_s range which extends to infinity at any C.L. A criticism was made in [149] that it is then dangerous to use this information in a CKM fit. The best way to answer this objection is to test the coverage of the probability regions (68%, 95% and 99%) computed by the fit by performing a Monte Carlo simulation.

To do this we have prepared a simplified CKM fit where we measure the quantity R_t (see Chapter 1), using only the ΔM_d and the $\Delta M_d/\Delta M_s$ constraints. The set of constraints on the quantity R_t is:

$$\Delta M_d = a^2 R_t^2 \quad (74)$$

$$\Delta M_d/\Delta M_s = b^2 R_t^2 \qquad (\text{or } \Delta M_s = a^2/b^2) \quad (75)$$

where a and b are Gaussian distributed parameters with errors $\sigma_a = 20\%$ and $\sigma_b = 10\%$, thus taking into account the theoretical uncertainties.

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

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

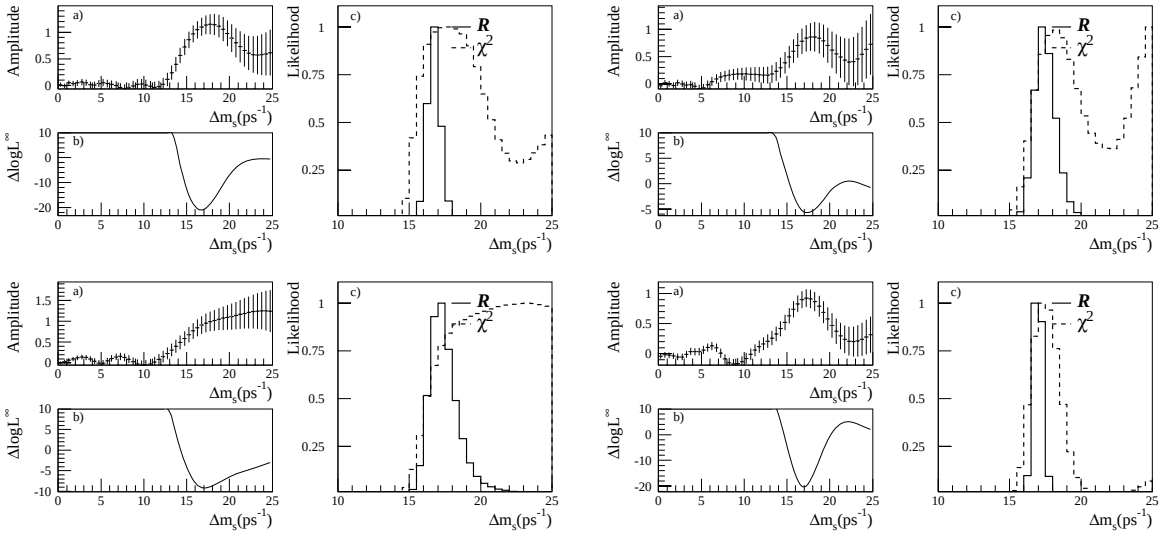

Fig. 4.24: *Toy-MC analyses with ΔM_s generated at $17\,\mathrm{ps}^{-1}$ corresponding to four virtual experiments. Each experiment is summarized in three plots: (a) amplitude spectrum, (b) $\Delta \log \mathcal{L}^{\infty}(\Delta M_s)$, (c) comparison between the Likelihood Ratio method (R) and the Modified χ^2 method (χ^2).*

	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 4.7: *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.*

For each experiment the best-fit value for R_t was determined and it was counted how many times it fell 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. The measured frequencies for the three probability regions using the Likelihood Ratio or the Modified χ^2 method are given in Table 4.7 and 4.8 respectively.

For the Likelihood Ratio method the measured frequencies correspond to the confidence level intervals and the coverage is close to correct. This is not the case for the Modified χ^2 method where the confidence levels are significantly underestimated for the true value of ΔM_s. The effect stems from the fact that the χ^2 defined in Eq. 67 reaches its minimum systematically above the true value of ΔM_s.

Some conclusions

In this first part we have studied the problem of including in CKM fits the ΔM_s World Average amplitude spectrum. We have tested two different methods and compared the results in case of an oscillation signal. MC simulations also were performed for a CKM fit to test the coverage of the two methods. The conclusion is that the Likelihood Ratio method, proposed in [150,151], is optimal because it gives probability intervals with correct coverage and, in case of a signal, it also gives the correct value of ΔM_s.

	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 4.8: *As for Table 4.7, but for the Modified χ^2 method.*

4.2. Use of the amplitude spectrum in a frequentist approach

The aim of this Section is to describe the frequentist method for incorporating experimental constraints derived from the amplitude spectrum as a function of the B_s^0 oscillation frequency (ΔM_s) into a global CKM fit. In other words, we address the questions: what is the pdf of a likelihood measurement of ΔM_s, and what is the confidence level (CL) as a function of ΔM_s to be associated with an observation obtained with a given level of sensitivity?

Infinite statistics

We assume that the x_s measurement is performed using the log-likelihood. The measured value of x_s ($x_\mathrm{s}^\mathrm{mes}$) is defined to be the one maximizing $\mathcal{L}(x_\mathrm{s})$: the outcome of one experiment $x_\mathrm{s}^\mathrm{mes}$ is a random number. For infinite statistics, the $x_\mathrm{s}^\mathrm{mes} = \Delta M_s \tau_b$ random number follows a (leading-order: lo) Gaussian probability density function:

$$\Phi_\mathrm{lo}^{x_\mathrm{s}}(x_\mathrm{s}^\mathrm{mes}) = \frac{1}{\sqrt{2\pi}\Sigma(x_\mathrm{s})} \exp\left(-\frac{1}{2}\left(\frac{x_\mathrm{s}^\mathrm{mes} - x_\mathrm{s}}{\Sigma(x_\mathrm{s})}\right)^2\right) \tag{76}$$

where the standard deviation $\Sigma(x_\mathrm{s})$ is given by the second derivative of the expected $\mathcal{L}$, through the integral A

$$(\sqrt{N}\Sigma(x_\mathrm{s}))^{-2} = \int_{-\infty}^{+\infty}\left(\frac{(\dot{\mathrm{P}}_-)^2}{\mathrm{P}_-} + \frac{(\dot{\mathrm{P}}_+)^2}{\mathrm{P}_+}\right)\mathrm{d}t_\mathrm{mes} \equiv \mathrm{A}(x_\mathrm{s}) \tag{77}$$

$$\dot{\mathrm{P}}_\pm = \frac{\partial \mathrm{P}_\pm}{\partial x_\mathrm{s}} = \mp f_\mathrm{s}\frac{1}{2} d\, t \sin(x_\mathrm{s} t) e^{-t} \otimes \mathrm{G_t} \tag{78}$$

N is the total number of mixed and unmixed events and the integrals are performed using the true value of x_s, not the measured one. It follows from Eq. (76) that one may set a confidence level $\mathrm{CL_{lo}}(x_\mathrm{s}^\mathrm{hyp})$ on a given hypothetical value $x_\mathrm{s}^\mathrm{hyp}$ using the χ^2 law:

$$\mathrm{CL_{lo}}(x_\mathrm{s}^\mathrm{hyp}) = \int_{<} \Phi_\mathrm{lo}^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^{\mathrm{mes}\prime})\mathrm{d}x_\mathrm{s}^{\mathrm{mes}\prime} = \mathrm{Prob}(\chi^2, 1) \tag{79}$$

$$\chi \equiv \chi^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^\mathrm{mes}) = \frac{x_\mathrm{s}^\mathrm{mes} - x_\mathrm{s}^\mathrm{hyp}}{\Sigma(x_\mathrm{s}^\mathrm{hyp})} \tag{80}$$

where the integral is performed over the $x_\mathrm{s}^{\mathrm{mes}\prime}$ domain where $\Phi_\mathrm{lo}^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^{\mathrm{mes}\prime}) < \Phi_\mathrm{lo}^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^\mathrm{mes})$, that is to say where $|\chi^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^{\mathrm{mes}\prime})| > |\chi^{x_\mathrm{s}^\mathrm{hyp}}(x_\mathrm{s}^\mathrm{mes})|$.

If the log-likelihood is parabolic near its maximum, as is the case for infinite statistics, then, in the vicinity of $x_\mathrm{s}^\mathrm{mes}$, $\Sigma(x_\mathrm{s}^\mathrm{hyp}) \simeq \mathrm{cst} = \Sigma(x_\mathrm{s}^\mathrm{mes})$, and one can evaluate Σ as the second derivative of the experimental log-likelihood, taken at the measured value $x_\mathrm{s}^\mathrm{mes}$. In effect:

$$\frac{\partial^2 \mathcal{L}}{\partial {x_\mathrm{s}}^2}_{|x_\mathrm{s}=x_\mathrm{s}^\mathrm{mes}} = \sum_-\left(\frac{\ddot{\mathrm{P}}_-\mathrm{P}_- - (\dot{\mathrm{P}}_-)^2}{\mathrm{P}_-^2}\right)^2 + \sum_+\left(\frac{\ddot{\mathrm{P}}_+\mathrm{P}_+ - (\dot{\mathrm{P}}_+)^2}{\mathrm{P}_+^2}\right)^2 \tag{81}$$

$$\stackrel{N\to\infty}{=} \quad -N\mathrm{A}(x_\mathrm{s}) = -\Sigma^{-2} \tag{82}$$

where $\ddot{\mathrm{P}}_\pm$ denotes the second derivative with respect to x_s:

$$\ddot{\mathrm{P}}_\pm = \frac{\partial^2 \mathrm{P}_\pm}{\partial {x_\mathrm{s}}^2} = \mp f_\mathrm{s} \frac{1}{2} d\, t^2 \cos(x_\mathrm{s} t) e^{-t} \otimes \mathrm{G_t} \tag{83}$$

which does not appear in the final expression thanks to the normalization of the probability density function, and assuming that $x_\mathrm{s}^\mathrm{mes} = x_\mathrm{s}$ (which is true for infinite statistics).

Equivalently, one can evaluate Σ by locating the value of $x_\mathrm{s}^\mathrm{hyp}$ which yields a drop of $-1/2$ in the log-likelihood, for the experiment at hand, or one can compute the χ^2 directly using the approximation

$$\chi^2(x_\mathrm{s}^\mathrm{hyp}) = \left(\frac{x_\mathrm{s}^\mathrm{mes} - x_\mathrm{s}^\mathrm{hyp}}{\Sigma(x_\mathrm{s}^\mathrm{hyp})} \right)^2 \simeq 2(\mathcal{L}(x_\mathrm{s}^\mathrm{mes}) - \mathcal{L}(x_\mathrm{s}^\mathrm{hyp})) \equiv \tilde{\chi}^2(x_\mathrm{s}^\mathrm{hyp}) \tag{84}$$

Finite statistics

For large enough $x_\mathrm{s}^\mathrm{hyp}$, the approximation $\Sigma(x_\mathrm{s}^\mathrm{hyp}) \simeq \Sigma(x_\mathrm{s}^\mathrm{mes})$ breaks down since the sensitivity of the experiment vanishes: $\Sigma(x_\mathrm{s}^\mathrm{hyp}) \to \infty$ for $x_\mathrm{s}^\mathrm{hyp} \to \infty$. It follows that the likelihood is not parabolic for large enough $x_\mathrm{s}^\mathrm{hyp}$, however large the statistics.

The vanishing sensitivity makes χ^2, as defined by Eq. (80), a poor test statistic to probe for large x_s values. Furthermore, it is not a straightforward task to infer the correct $\mathrm{CL}(x_\mathrm{s}^\mathrm{hyp})$ from the χ^2 value: Eq. (79) does not apply (i.e., it is not a true χ^2) because Eq. (76) is a poor approximation**.

In the realistic case of finite statistics, the next-to-leading order statistical analysis of a likelihood measurement [153] is used here to obtain the key-formula expressing the probability density function of the random number $x_\mathrm{s}^\mathrm{mes}$ beyond the Gaussian approximation:

$$\Phi_\mathrm{nlo}^{x_\mathrm{s}}(x_\mathrm{s}^\mathrm{mes}) \;=\; \Phi_\mathrm{lo}^{x_\mathrm{s}}(x_\mathrm{s}^\mathrm{mes})\, e^{-{\mathrm{a}_3}^{x_\mathrm{s}} \chi^3} (1 + {\mathrm{a}_0}^{x_\mathrm{s}} \chi) \tag{85}$$

$${\mathrm{a}_0}^{x_\mathrm{s}} \;=\; \frac{2\mathrm{B} - \mathbb{C}}{2\mathrm{A}} \frac{1}{\sqrt{N\mathrm{A}}} = -\dot{\Sigma} \tag{86}$$

$${\mathrm{a}_3}^{x_\mathrm{s}} \;=\; \frac{3\mathrm{B} - \mathbb{C}}{6\mathrm{A}} \frac{1}{\sqrt{N\mathrm{A}}} \tag{87}$$

where $\mathrm{A}(x_\mathrm{s})$ is the integral defined in Eq. (77), $B(x_\mathrm{s})$ and $\mathbb{C}(x_\mathrm{s})$ being two new integrals:

$$B(x_\mathrm{s}) \;=\; \int_{-\infty}^{+\infty} \left(\frac{\dot{\mathrm{P}}_- \ddot{\mathrm{P}}_-}{\mathrm{P}_-} + \frac{\dot{\mathrm{P}}_+ \ddot{\mathrm{P}}_+}{\mathrm{P}_+} \right) \mathrm{d}t_\mathrm{mes} \tag{88}$$

**The redefinition of the χ^2 using the right-hand side of Eq. (84) provides a test statistic more appropriate for large values of $x_\mathrm{s}^\mathrm{hyp}$. Although Eq. (79) does not apply, $\tilde{\chi}^2$ is capable of ruling out $x_\mathrm{s}^\mathrm{hyp}$ values lying beyond the sensitivity reach (if $\mathcal{L}(x_\mathrm{s}^\mathrm{mes})$ is large enough) provided one computes the CL using:

$$\mathrm{CL}(x_\mathrm{s}^\mathrm{hyp}) = \int_{\tilde{\chi}^2(x_\mathrm{s}^\mathrm{hyp})}^{\infty} \Psi^{x_\mathrm{s}^\mathrm{hyp}}(\tilde{\chi}^{2\prime})\, \mathrm{d}\tilde{\chi}^{2\prime}$$

where $\Psi^{x_\mathrm{s}^\mathrm{hyp}}$ is the probability density function of the $\tilde{\chi}^2$ test statistic, for $x_\mathrm{s} = x_\mathrm{s}^\mathrm{hyp}$, obtained using a toy Monte Carlo. The rejection of $x_\mathrm{s}^\mathrm{hyp}$ values beyond the sensitivity reach is not a paradox: it uses the fact that large values are unlikely to yield an indication of a clear signal, especially at low values of x_s. Such a treatment, as well as others (e.g., the minimum value of the likelihood could be used to define another test statistics) are satisfactory. We prefer here to use x_s^mes, and only this quantity, because an analytical expression for its probability density function is available (Eq. 85) and thus the computation of the CL can be carried out in practice. This is nothing but the standard choice made when dealing with better defined measurements.

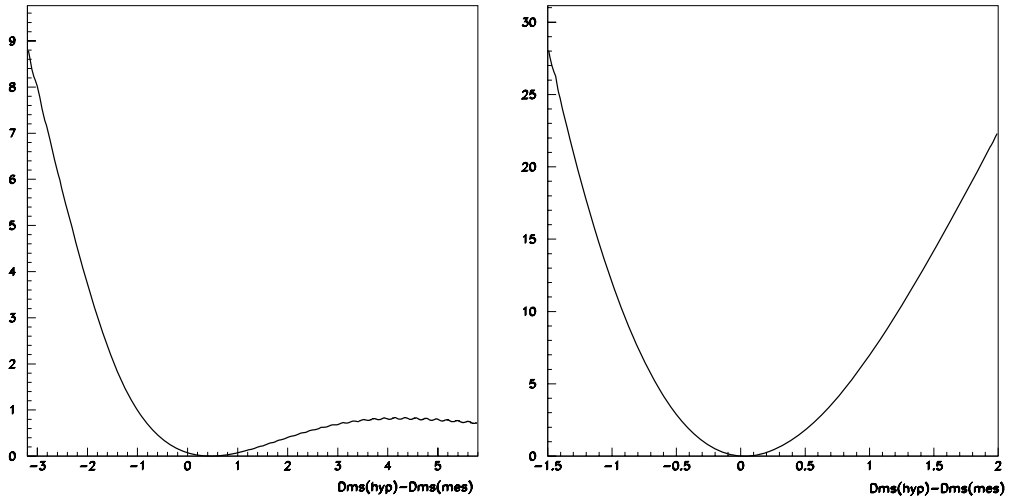

Fig. 4.25: *Left Plot: The equivalent χ^2 (cf. Eq. (79)) expressing the confidence levels computed using the next-to-leading order expression Eq. (85) in the actual situation where the maximum value of the likelihood is reached for $\Delta M_s(\mathrm{mes}) = 17.2\,\mathrm{ps}^{-1}$. The horizontal axis is the difference $\Delta M_s(\mathrm{hypothetical}) - \Delta M_s(\mathrm{mes})$. The minimum value of the equivalent χ^2 is not reached for $\Delta M_s(\mathrm{hypothetical}) = \Delta M_s(\mathrm{mes})$ because the maximum of the next-to-leading order probability density function is slightly shifted below the true ΔM_s value. The left hand side of the plot is nearly parabolic and resembles closely the one that a simplistic interpretation of the likelihood curve provides. The right hand side of the plot states that there is almost no constraint on high values of ΔM_s. One is far from dealing with a measurement in the usual (Gaussian) sense. Right Plot: The equivalent χ^2 in the would-be situation where the maximum value of the likelihood is reached for $\Delta M_s(\mathrm{mes}) = 10\,\mathrm{ps}^{-1}$. Although the equivalent χ^2 is not truly parabolic, the Gaussian limit is almost reached: one is close to dealing with a measurement in the usual (Gaussian) sense.*

$$\mathbb{C}(x_\mathrm{s}) \;=\; \int_{-\infty}^{+\infty} \left(\frac{(\dot{\mathrm{P}}_-)^3}{\mathrm{P}_-^2} + \frac{(\dot{\mathrm{P}}_+)^3}{\mathrm{P}_+^2} \right) \mathrm{d}t_\mathrm{mes} \tag{89}$$

The integral $\mathbb{C}$ tends to be small because, on the one hand the two contributions have opposite signs, and on the other hand the denominator is of order two: it follows that $\mathrm{a}_3 \simeq \mathrm{a}_0/2$. The right hand side of Eq. (86) links the next-to-leading order correction terms a_0 and a_3 to the dependence on x_s of Σ. When Σ depends significantly on x_s, not only is the standard treatment of Sec. 4.2. invalid, but the well-known formula Eq. (79) itself becomes incorrect, even if one uses the correct $\Sigma(x_\mathrm{s})$.

The expression Eq. (85) is identical to Eq. (76) for small χ values. Although it extends the range of validity to larger χ values, it cannot be trusted too far away from the origin, where higher order corrections start to play a role. In particular, Φ_nlo becomes negative (hence meaningless) for $\chi > -\mathrm{a}_0{}^{-1}$ (a_0 is negative since it is equal to minus the derivative of Σ with respect to x_s). Since Φ is sizable only when $\chi \sim \mathcal{O}(1)$ the next-to-leading order terms, when relevant, are of the form $N^{-\frac{1}{2}} \times (\mathrm{ratio\ of\ integrals})$: they are negligible for large enough N and for small enough ratio of integrals. The double-sided CL is computed as in Eq. (79), replacing Φ_lo by the next-to-leading order approximation. Using the right hand side of Eq. (79) to translate the confidence level thus obtained into a more familiar equivalent[††] χ^2, one obtains the results shown in Fig. (4.25) in two cases: first for the actual situation using the parametrization of the world average likelihood as described in Sec. 4.1. where the maximum of the likelihood is

[††] In the CKMfitter package, it is this equivalent χ^2 which is added to the overall χ^2.

reached at the boundary of the experimental sensitivity; second for a hypothetical situation where the maximum of the likelihood would be reached well within the sensitivity region.

In conclusion, we have presented a frequentist analysis of the B_s oscillation. Its domain of validity extends to the level of sensitivity reached by LEP and SLD. The treatment presented here provides, in a frequentist approach, a practical means to incorporate into a CKM fit the information on ΔM_s contained in the data, both present and future.

References

[1] A.J. Buras, M. Jamin and P.H. Weisz, Nucl. Phys. B **347** (1990) 491.

[2] A.J. Buras, W. Slominski and H. Steger, Nucl. Phys. B **245** (1984) 369.

[3] S. Herrlich and U. Nierste, Nucl. Phys. B **419** (1994) 292, [hep-ph/9310311].

[4] S. Herrlich and U. Nierste, Phys. Rev. D **52** (1995) 6505, [hep-ph/9507262].

[5] S. Herrlich and U. Nierste, Nucl. Phys. B **476** (1996) 27, [hep-ph/9604330].

[6] M. Jamin and U. Nierste (2003). Recent update for this document.

[7] E. de Rafael, in Proceedings of TASI 1994 (1995), hep-ph/9502254.

[8] J. Urban, F. Krauss, U. Jentschura and G. Soff, Nucl. Phys. B **523** (1998) 40, [hep-ph/9710245].

[9] M. Ciuchini *et al.*, JHEP **07** (2001) 013, [hep-ph/0012308].

[10] C. Bernard *et al.*, Nucl. Phys. Proc. Suppl. **106** (2002) 412, [hep-lat/0110072].

[11] MILC Collaboration, C. Bernard *et al.*, hep-lat/0209163.

[12] C. Bernard *et al.*, in Proceedings of Lattice 2002 (2002) hep-lat/0209086.

[13] M.J. Booth, Phys. Rev. D **51** (1995) 2338, [hep-ph/9411433].

[14] A.S. Kronfeld and S.M. Ryan, Phys. Lett. B **543** (2002) 59, [hep-ph/0206058].

[15] JLQCD Collaboration, N. Yamada *et al.*, Nucl. Phys. Proc. Suppl. **106** (2002) 397, [hep-lat/0110087].

[16] N. Yamada, in Proceedings of 20th Int. Symp. on Lattice Field Theory, Lattice 2002, Boston, June 2002 (2002) hep-lat/0210035.

[17] MILC Collaboration, C. Bernard *et al.*, Phys. Rev. D **66** (2002) 094501, [hep-lat/0206016].

[18] D. Becirevic, S. Fajfer, S. Prelovsek and J. Zupan, hep-ph/0211271.

[19] S.R. Sharpe and Y. Zhang, Phys. Rev. D **53** (1996) 5125, [hep-lat/9510037].

[20] A.V. Manohar and M.B. Wise, Heavy Quark Physics, vol. 10 of Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, 2000).

[21] CLEO Collaboration, S. Ahmed *et al.*, Phys. Rev. Lett. **87** (2001) 251801, [hep-ex/0108013].

[22] UKQCD Collaboration, G.M. de Divitiis *et al.*, JHEP **10** (1998) 010, [hep-lat/9807032].

[23] A. Abada *et al.*, hep-lat/0209092.

[24] JLQCD Collaboration, S. Hashimoto *et al.*, hep-lat/0209091.

[25] L. Lellouch, in Proceedings of ICHEP 2002, 31st Int. Conf. on High Energy Physics, Amsterdam, July 2002, hep-ph/0211359.

[26] D. Becirevic *et al.*, Nucl. Phys. B **618** (2001) 241, [hep-lat/0002025].

[27] UKQCD Collaboration, L. Lellouch and C.J.D. Lin, Phys. Rev. D **64** (2001) 094501, [hep-ph/0011086].

[28] UKQCD Collaboration, K.C. Bowler *et al.*, Nucl. Phys. B **619** (2001) 507, [hep-lat/0007020].

[29] A.X. El-Khadra *et al.*, Phys. Rev. D **58** (1998) 014506, [hep-ph/9711426].

[30] CP-PACS Collaboration, A. Ali Khan *et al.*, Phys. Rev. D **64** (2001) 034505, [hep-lat/0010009].

[31] A. Ali Khan *et al.*, Phys. Lett. B **427** (1998) 132, [hep-lat/9801038].

[32] S. Collins *et al.*, Phys. Rev. D **60** (1999) 074504, [hep-lat/9901001].

[33] JLQCD Collaboration, K.I. Ishikawa *et al.*, Phys. Rev. D **61** (2000) 074501, [hep-lat/9905036].

[34] CP-PACS Collaboration, A. Ali Khan *et al.*, Phys. Rev. D **64** (2001) 054504, [hep-lat/0103020].

[35] C.W. Bernard, T. Blum and A. Soni, Phys. Rev. D **58** (1998) 014501, [hep-lat/9801039].

[36] M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Nucl. Phys. B **147** (1979) 385.

[37] M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Nucl. Phys. B **147** (1979) 448.

[38] A.A. Penin and M. Steinhauser, Phys. Rev. D **65** (2002) 054006, [hep-ph/0108110].

[39] M. Jamin and B.O. Lange, Phys. Rev. D **65** (2002) 056005, [hep-ph/0108135].

[40] S. Narison, Phys. Lett. B **520** (2001) 115, [hep-ph/0108242].

[41] P. Colangelo and A. Khodjamirian, hep-ph/0010175.

[42] K.G. Chetyrkin and M. Steinhauser, Phys. Lett. B **502** (2001) 104, [hep-ph/0012002].

[43] K.G. Chetyrkin and M. Steinhauser, Eur. Phys. J. C **21** (2001) 319, [hep-ph/0108017].

[44] M. Jamin, Phys. Lett. B **538** (2002) 71, [hep-ph/0201174].

[45] A.X. El-Khadra and M. Luke, hep-ph/0208114.

[46] A.A. Ovchinnikov and A.A. Pivovarov, Phys. Lett. B **207** (1988) 333.

[47] L.J. Reinders and S. Yazaki, Phys. Lett. B **212** (1988) 245.

[48] S. Narison and A.A. Pivovarov, Phys. Lett. B **327** (1994) 341, [hep-ph/9403225].

[49] K. Hagiwara, S. Narison and D. Nomura, Phys. Lett. B **540** (2002) 233, [hep-ph/0205092].

[50] V. Chernyak, Nucl. Phys. B **457** (1995) 96 [hep-ph/9503208].

[51] N. Cabibbo, G. Martinelli and R. Petronzio, Nucl. Phys. B **244** (1984) 381.

[52] R.C. Brower, G. Maturana, M. Belen Gavela and R. Gupta, Phys. Rev. Lett. **53** (1984) 1318.

[53] C.W. Bernard, in Proceedings of Gauge Theory On A Lattice, Argonne (1984), p. 85.

[54] C.W. Bernard *et al.*, Phys. Rev. Lett. **55** (1985) 2770.

[55] M.B. Gavela *et al.*, Nucl. Phys. B **306** (1988) 677.

[56] Staggered Collaboration, G. Kilcup, R. Gupta and S.R. Sharpe, Phys. Rev. D **57** (1998) 1654, [hep-lat/9707006].

[57] JLQCD Collaboration, S. Aoki *et al.*, Phys. Rev. Lett. **80** (1998) 5271, [hep-lat/9710073].

[58] SPQCDR Collaboration, D. Becirevic *et al.*, [hep-lat/0209136].

[59] CP-PACS Collaboration, A. Ali Khan *et al.*, Phys. Rev. D **64** (2001) 114506, [hep-lat/0105020].

[60] RBC Collaboration, T. Blum *et al.*, hep-lat/0110075.

[61] MILC Collaboration, T. DeGrand, hep-lat/0208054.

[62] GGHLR Collaboration, N. Garron *et al.*, hep-lat/0212015.

[63] T. Blum and A. Soni, Phys. Rev. Lett. **79** (1997) 3595, [hep-lat/9706023].

[64] M. Guagnelli *et al.*, hep-lat/0110097.

[65] N. Ishizuka *et al.*, Phys. Rev. Lett. **71** (1993) 24.

[66] G. Kilcup, Phys. Rev. Lett. **71** (1993) 1677.

[67] S. Sharpe, Nucl. Phys. Proc. Suppl. **53** (1997) 181, [hep-lat/9609029].

[68] J. Bijnens, H. Sonoda and M.B. Wise, Phys. Rev. Lett. **53** (1984) 2367.

[69] S.R. Sharpe, Phys. Rev. D **46** (1992) 3146, [hep-lat/9205020].

[70] S.R. Sharpe, hep-ph/9412243.

[71] M.F.L. Golterman and K.C. Leung, Phys. Rev. D **57** (1998) 5703, [hep-lat/9711033].

[72] J. Prades *et al.*, Z. Phys. C **51** (1991) 287.

[73] M. Jamin and A. Pich, Nucl. Phys. B **425** (1994) 15, [hep-ph/9402363].

[74] S. Narison, Phys. Lett. B **351** (1995) 369, [hep-ph/9409428].

[75] L.J. Reinders and S. Yazaki, Nucl. Phys. B **288** (1987) 789.

[76] N. Bilic, C.A. Dominguez and B. Guberina, Z. Phys. C **39** (1988) 351.

[77] W. Bardeen, A. Buras and J.M. Gérard, Phys. Lett. **211** (1988) 343.

[78] T. Hambye, G.O. Köhler and P.H. Soldan, Eur. Phys. J. C **10** (1999) 271, [hep-ph/9902334].

[79] S. Peris and E. de Rafael, Phys. Lett. B **490** (2000) 213, [hep-ph/0006146].

[80] J. Bijnens and J. Prades, J. High Energy Phys. **01** (2000) 002, [hep-ph/9909244].

[81] S. Bertolini, J.O. Eeg, M. Fabbrichesi and E.I. Lashin, Nucl. Phys. B **514** (1998) 63, [hep-ph/9705244].

[82] A.J. Buras, M. Jamin and P.H. Weisz, Nucl. Phys. B **347** (1990) 491.

[83] J.F. Donoghue, E. Golowich and B.R. Holstein, Phys. Lett. B **119** (1982) 412.

[84] LEP/SLD Electroweak Working Group, CERN-EP/2001-021.

[85] LEP/SLD/CDF Collaborations, CERN-EP/2001-050.

[86] DELPHI Collaborations, EPS-HEP99, contributed paper 5.515.

[87] The LEP B Oscillation Working Group, *http://www.cern.ch/LEPBOSC/*.

[88] M. Gronau, A. Nippe and J.L. Rosner, Phys. Rev. D **47** (1993) 1988, [hep-ph/9211311].

[89] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **80** (1998) 2057, [hep-ex/9712004].

[90] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **59** (1999) 032001, [hep-ex/9806026].

[91] P.C. Rowson, D. Su and S. Willocq, Ann. Rev. Nucl. Part. Sci. **51** (2001) 345, [hep-ph/0110168].

[92] SLD Collaboration (1996). SLAC-PUB-7228, contribution to ICHEP96 Warsaw.

[93] SLD Collaboration, K. Abe *et al.*, SLAC-PUB-7229, contribution to ICHEP96 Warsaw.

[94] SLD Collaboration, SLAC-PUB-7230, contribution to ICHEP96 Warsaw.

[95] J.L. Wittlin (2001). Ph.D. thesis, SLAC-R-582.

[96] ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. C **4** (1998) 367.

[97] ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. C **7** (1999) 553, [hep-ex/9811018].

[98] DELPHI Collaboration, P. Abreu *et al.*, Eur. Phys. J. C **18** (2000) 229, [hep-ex/0105077].

[99] DELPHI Collaboration, P. Abreu *et al.*, Eur. Phys. J. C **16** (2000) 555, [hep-ex/0107077].

[100] OPAL Collaboration, K. Ackerstaff *et al.*, Z. Phys. C **76** (1997) 417, [hep-ex/9707010].

[101] OPAL Collaboration, K. Ackerstaff *et al.*, Z. Phys. C **76** (1997) 401, [hep-ex/9707009].

[102] J. Thom, SLAC-R-585 (2002).

[103] CDF Collaboration, M. Paulini, Int. J. Mod. Phys. A **14** (1999) 2791, [hep-ex/9903002].

[104] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87** (2001) 091801, [hep-ex/0107013].

[105] Belle Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **87** (2001) 091802, [hep-ex/0107061].

[106] H.G. Moser and A. Roussarie, Nucl. Instrum. Meth. A **384** (1997) 491.

[107] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **88** (2002) 221802, [hep-ex/0112044].

[108] Belle Collaboration, T. Tomura *et al.*, Phys. Lett. B **542** (2002) 207, [hep-ex/0207022].

[109] BABAR Collaboration, B. Aubert *et al.*, [hep-ex/0212017].

[110] Belle Collaboration, K. Hara *et al.* Phys. Rev. Lett. **89** (2002) 251803, [hep-ex/0207045].

[111] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **60** (1999) 112004, [hep-ex/9907053].

[112] OPAL Collaboration, G. Alexander *et al.*, Z. Phys. C **72** (1996) 377.

[113] ALEPH Collaboration, D. Buskulic *et al.*, Z. Phys. C **75** (1997) 397.

[114] DELPHI Collaboration, P. Abreu *et al.*, Z. Phys. C **76** (1997) 579.

[115] OPAL Collaboration, G. Abbiendi *et al.*, Phys. Lett. B **493** (2000) 266, [hep-ex/0010013].

[116] Belle Collaboration, Y. Zheng *et al.*, to appear in Phys. ReV. D, [hep-ex/0211065].

[117] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **60** (1999) 051101.

[118] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **60** (1999) 072003, [hep-ex/9903011].

[119] L3 Collaboration, M. Acciarri *et al.*, Eur. Phys. J. C **5** (1998) 195.

[120] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **88** (2002) 221803, [hep-ex/0112045].

[121] Belle Collaboration, N. Hastings *et al.*,to appear in Phys. Rev. D, [hep-ex/0212033].

[122] DELPHI Collaboration, J. Abdallah *et al.*, CERN-EP-2002-078.

[123] ALEPH Collaboration, D. Buskulic *et al.*, ALEPH-97/027, contribution to EPS-HEP97 Jerusalem.

[124] ALEPH Collaboration, A. Heister *et al.*, CERN-EP-2002-016.

[125] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **82** (1999) 3576.

[126] DELPHI Collaboration, J. Abdallah *et al.* DELPHI 2002-073, contribution to ICHEP 2002 Amsterdam.

[127] OPAL Collaboration, G. Abbiendi *et al.*, Eur. Phys. J. C **19** (2001) 241, [hep-ex/0011052].

[128] SLD Collaboration, K. Abe *et al.*, Phys. Rev. D **66** (2002) 032009, [hep-ex/0207048].

[129] OPAL Collaboration, G. Abbiendi *et al.*, Eur. Phys. J. C **11** (1999) 587, [hep-ex/9907061].

[130] SLD Collaboration, K. Abe *et al.*, SLAC-PUB-8568, [hep-ex/0012043].

[131] SLD Collaboration, K. Abe *et al.*, Phys. Rev. D **67** (2003) 012006, [hep-ex/0209002].

[132] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. C **55** (1992) 357.

[133] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **324** (1994) 249.

[134] CLEO Collaboration, J. Bartelt *et al.*, Phys. Rev. Lett. **71** (1993) 1680.

[135] CLEO Collaboration, B.H. Behrens *et al.*, Phys. Lett. B **490** (2000) 36, [hep-ex/0005013].

[136] DELPHI Collaboration, W. Adam *et al.*, Phys. Lett. B **414** (1997) 382.

[137] Particle Data Group Collaboration, K. Hagiwara *et al.*, Phys. Rev. D **66** (2002) 010001.

[138] CDF-II Collaboration, R. Blair *et al.*, FERMILAB-PUB-96-390-E.

[139] D0 Collaboration, J. Ellison, prepared for 15th International Workshop on High-Energy Physics and Quantum Field Theory (QFTHEP 2000), Tver, Russia, 14-20 Sep 2000.

[140] S. Cabrera *et al.*, Nucl. Instrum. Meth. A **494** (2002) 416.

[141] C. Grozis *et al.*, hep-ex/0209027.

[142] CDF Collaboration, A. Cerri, presented at 31st International Conference on High Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 24-31 Jul 2002.

[143] CDF Collaboration, D. Lucchesi, presented at Beauty 2002: 8th International Conference on B Physics at Hadron machines, Santiago de Compostela, Spain, 17-21 Jun 2002.

[144] K. Anikeev *et al.*, hep-ph/0201071.

[145] CDF Collaboration, M. Paulini, hep-ex/0302016.

[146] P. Paganini, F. Parodi, P. Roudeau and A. Stocchi, Phys. Scripta **58** (1998) 556, [hep-ph/9711261].

[147] Y. Grossman, Y. Nir, S. Plaszczynski and M.H. Schune, Nucl. Phys. B **511** (1998) 69, [hep-ph/9709288].

[148] F. Parodi, P. Roudeau and A. Stocchi, Nuovo Cim. A **112** (1999) 833, [hep-ex/9903063].

[149] A. Hoecker, H. Lacker, S. Laplace and F. Le Diberder, Eur. Phys. J. C **21** (2001) 225, [hep-ph/0104062].

[150] P. Checchia, E. Piotto and F. Simonetto (1999), hep-ph/9907300.

[151] M. Ciuchini *et al.*, JHEP **07** (2001) 013, [hep-ph/0012308].

[152] G. Boix and D. Abbaneo, JHEP **08** (1999) 004, [hep-ex/9909033].

[153] R.N. Cahn, Private communication.