Chapter 5

FIT OF THE UNITARITY TRIANGLE PARAMETERS

Conveners : A.J. Buras, F. Parodi.
Contributors : M. Ciuchini, G. Dubois-Felsmann, G. Eigen, P. Faccioli, E. Franco, A. Hocker, D. Hitlin, H. Lacker, S. Laplace, F. LeDiberder, V. Lubicz, G. Martinelli, F. Porter, P. Roudeau, L. Silvestrini, A. Stocchi, M. Villa

1. Introduction

In this Chapter we will discuss the determination of the Unitarity Triangle (UT) using as input the values of $|V_{us}|$, $|V_{cb}|$, and $|V_{ub}|$ from Chapters 2 and 3 and the constraints from ε_K and $\Delta M_{d,s}$ with the values of the non-perturbative parameters $\hat{B}_K$, $F_{B_d}\sqrt{\hat{B}_{B_d}}$, $F_{B_s}\sqrt{\hat{B}_{B_s}}$ and ξ determined in Chapter 4. We will also include in this analysis the most recent results for the CP asymmetry in $\mathrm{B}_d \to J/\psi \mathrm{K}_S$ that allows to determine the angle β of the UT essentially without any theoretical uncertainty. The list of the common quantities which have been used for the analyses performed in this Chapter are summarised in Table 5.1.

A very important issue in constraining the apex $(\bar{\varrho}, \bar{\eta})$ of the UT is the treatment of the experimental and especially the theoretical uncertainties. In the literature five different approaches can be found: Gaussian approach [1], Bayesian approach [2], frequentist approach [3], 95% C.L. scan method [4] and the simple (naive) scanning within one standard deviation. Moreover the fact that different authors often use different input parameters makes the comparison of various analyses very difficult.

This situation is clearly unsatisfactory as different analyses give generally different allowed ranges for $(\bar{\varrho}, \bar{\eta})$. While all these analyses find presently the SM consistent with all the available data, the situation may change in the future when the experimental and theoretical uncertainties will be reduced and additional decays relevant for the determination of the UT will be available.

It is then conceivable that some approaches will find the SM consistent with the data whereas other will claim an indication for new physics contributions. This clearly would be rather unfortunate. However, even in the absence of new physics contributions, the increasing accuracy of the data and the expected reduction of theoretical uncertainties calls for an extensive comparison of the different methods to gain the best understanding on the UT.

Another important issue is the sensitivity of the UT analysis to theoretical uncertainties. Some theoretical parameters have more impact on this analysis than others and it is important to identify those for which the reduction of errors through improved non-perturbative calculations can contribute to the quality of the determination of the UT most efficiently.

Parameter	Value	Gaussian σ	Theory uncertainty
λ	0.2240(0.2210)	0.0036 (*0.0020*)	-
$\|V_{cb}\|$ $(\times 10^{-3})$ (excl.)	42.1	2.1	-
$\|V_{cb}\|$ $(\times 10^{-3})$ (incl.)	41.4 (*40.4*)	0.7	0.6(*0.8*)
$\|V_{ub}\|$ $(\times 10^{-4})$ (excl.)	33.0(*32.5*)	2.4(*2.9*)	4.6(*5.5*)
$\|V_{ub}\|$ $(\times 10^{-4})$ (incl.)	40.9	4.6	3.6
ΔM_d (ps^{-1})	0.503 (*0.494*)	0.006 (*0.007*)	-
ΔM_s (ps^{-1})	$>$ 14.4 (*14.9*) at 95% C.L.	sensitivity 19.2 (*19.3*)	
m_t (GeV)	167	5	-
m_c (GeV)	1.3	-	0.1
$F_{B_d}\sqrt{\hat{B}_{B_d}}$(MeV)	223 (*230*)	33 (*30*)	12 (*15*)
$\xi = \frac{F_{B_s}\sqrt{\hat{B}_{B_s}}}{F_{B_d}\sqrt{\hat{B}_{B_d}}}$	1.24(1.18)	0.04 (*0.03*)	0.06 (*0.04*)
$\hat{B}_K$	0.86	0.06	0.14
sin 2β	0.734 (*0.762*)	0.054 (*0.064*)	-

Table 5.1: *Latest values of the relevant quantities entering into the expressions of ϵ_K, ΔM_d and ΔM_s. In the third and fourth columns the Gaussian and the flat part of the uncertainty are given, respectively. The values within parentheses are the ones available at the time of the Workshop and used when comparing different fitting procedures. In case of asymmetric theoretical errors, like for $|V_{ub}|$ exclusive, the central values have been shifted to make them symmetric.*

The goals of this Chapter are:

- to describe in some detail two of the most developed methods: the Bayesian approach and the frequentist approach,
- to compare the resulting allowed regions for $(\bar{\varrho}, \bar{\eta})$ obtained from the Bayesian and frequentist approaches for the same input parameters,
- to identify those non-perturbative parameters for which the reduction of the uncertainty is most urgent.

This Chapter is organized as follows. In Section 2. we express the constraints from $|V_{ub}/V_{cb}|$, ε_K and $\Delta M_{d,s}$ in terms of Wolfenstein parameters [5] including the generalization of [6]. The Bayesian method and the frequentist methods are discussed in Sections 3.1. and 3.2., respectively. The discussion in the frequentist case includes the *R*fit and the scanning methods. In Section 4. the impact of the uncertainties of theoretical parameters on the determination of the UT is discussed in detail using both the Bayesian approach and the scanning method. Finally in Section 5. we compare the Bayesian and *R*fit methods and draw conclusions. In Section 6. we show some important results obtained in testing the consistency of the CKM picture of the Standard Model.

2. Constraints on the Unitarity Triangle parameters

Five measurements restrict at present the range of $(\bar{\varrho}, \bar{\eta})$ within the SM:

- The $|V_{ub}|$ constraint:

The length of the side AC of the UT (see Fig. 5.1) is determined from

$$R_b = \sqrt{\bar{\varrho}^2 + \bar{\eta}^2} = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|. \tag{1}$$

The constraint in the $(\bar{\varrho}, \bar{\eta})$ plane resulting from (1) is represented by a circle of radius R_b that is centered at $(\bar{\varrho}, \bar{\eta}) = (0, 0)$ (for the visualisation of this and following constraints see Fig. 5.1).

- The ε_K–constraint:

$$\bar{\eta}\left[(1-\bar{\varrho})A^2\eta_2 S(x_t) + P_c(\varepsilon)\right]A^2\hat{B}_K = 0.187\,\left(\frac{0.224}{\lambda}\right)^{10}\,, \tag{2}$$

that follows from the experimental value for ε_K and the formula (Eq. (12) of Chapter 4). Here

$$P_c(\varepsilon) = \left[\eta_3 S_0(x_c, x_t) - \eta_1 x_c\right]\frac{1}{\lambda^4}, \qquad x_t = \frac{m_{\rm t}^2}{M_{\rm W}^2}. \tag{3}$$

$P_c(\varepsilon) = 0.29 \pm 0.07$ [7] summarizes the contributions of box diagrams with two charm quark exchanges and the mixed charm-top exchanges. We observe a very strong dependence of the r.h.s. in (2) on the parameter $\lambda = |V_{us}|$. However, this dependence is cancelled to a large extent by the λ dependence of $P_c(\varepsilon)$ and of $A = |V_{cb}|/\lambda^2$ that enter the l.h.s of (2). The main uncertainties in the constraint (2) reside then in $\hat{B}_K$ and to some extent in the factor A^4 or equivalently $|V_{cb}|^4$ which multiplies the dominant term. The status of $\hat{B}_K$ has been reviewed in Chapter 4. Eq. (2) specifies an hyperbola in the $(\bar{\varrho}, \bar{\eta})$ plane. This hyperbola intersects the circle found from the $|V_{ub}|$ constraint in two points which correspond to two solutions for the angle γ.

- The ΔM_d–constraint:
The length R_t of the side AB of the UT (see Fig. 5.1) can be determined from the observed $\mathrm{B}_d^0 - \overline{\mathrm{B}}_d^0$ mixing, parametrized by ΔM_d and given in Eq. 22 (in Chapter 4), with the result

$$R_t = \sqrt{(1-\bar{\varrho})^2 + \bar{\eta}^2} = \frac{1}{\lambda}\frac{|V_{td}|}{|V_{cb}|} = 0.85 \cdot \left[\frac{|V_{td}|}{7.8\cdot 10^{-3}}\right]\left[\frac{0.041}{|V_{cb}|}\right] \tag{4}$$

where

$$|V_{td}| = 7.8\cdot 10^{-3}\left[\frac{230\mathrm{MeV}}{\sqrt{F_{B_d}\hat{B}_{B_d}}}\right]\sqrt{\frac{\Delta M_d}{0.50/\mathrm{ps}}}\sqrt{\frac{0.55}{\eta_B}}\sqrt{\frac{2.34}{S_0(x_t)}}\,. \tag{5}$$

Since $m_{\rm t}$, ΔM_d and η_B are already rather precisely known, the main uncertainty in the determination of R_t and $|V_{td}|$ from $\mathrm{B}_d^0 - \overline{\mathrm{B}}_d^0$ mixing comes from $F_{B_d}\sqrt{\hat{B}_{B_d}}$. Its theoretical status has been reviewed in Chapter 4. R_t suffers from additional uncertainty in $|V_{cb}|$. The constraint in the $(\bar{\varrho}, \bar{\eta})$ plane resulting from (4) is represented by a circle of radius R_t that is centered at $(\bar{\varrho}, \bar{\eta}) = (1, 0)$.

- The $\Delta M_d/\Delta M_s$–constraint:
The measurement of $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ mixing parametrized by ΔM_s together with ΔM_d allows to determine R_t in a different manner:

$$R_t = \frac{1}{\lambda}\xi\sqrt{\frac{M_{\mathrm{B}_s}}{M_{\mathrm{B}_d}}}\sqrt{\frac{\Delta M_s}{\Delta M_d}}(1 - \frac{\lambda^2}{2} + \bar{\varrho}\lambda^2), \qquad \xi = \frac{F_{B_s}\sqrt{\hat{B}_s}}{F_{B_d}\sqrt{\hat{B}_d}}\,. \tag{6}$$

This constraint follows from Eq. (22) (in Chapter 4) with the factor $(1 - \lambda^2/2 + \bar{\varrho}\lambda^2)$ representing the departure of $|V_{ts}/V_{cb}|$ from unity. For $0 \le \bar{\varrho} \le 0.5$ this factor deviates from unity by less than 2%. Neglecting this correction gives ($\lambda = 0.224$)

$$R_t = 0.86\sqrt{\frac{\Delta M_d}{0.50/\mathrm{ps}}}\sqrt{\frac{18.4/\mathrm{ps}}{\Delta M_s}}\left[\frac{\xi}{1.18}\right]. \tag{7}$$

The advantage of determining R_t by means of the ratio $\Delta M_d/\Delta M_s$ with respect to the ΔM_d constraint are smaller hadronic uncertainties in ξ than in $F_{B_d}\sqrt{\hat{B}_d}$ and the absence of m_t and $|V_{cb}|$ dependence. The present status of ξ has been reviewed in Chapter 4.

- The $a(\psi K_S)$–constraint:
  The mixing induced CP asymmetry $a_{\psi K_S}$ in $\mathrm{B} \to \psi \mathrm{K_S}$ allows to determine the angle β of the UT essentially without any hadronic uncertainties through

$$(\sin 2\beta)_{\psi K_S} = 0.734 \pm 0.054 \; . \tag{8}$$

  The value given in (8) is the world average from [8] and is dominated by the results of the BaBar [9] and Belle [10] Collaborations.

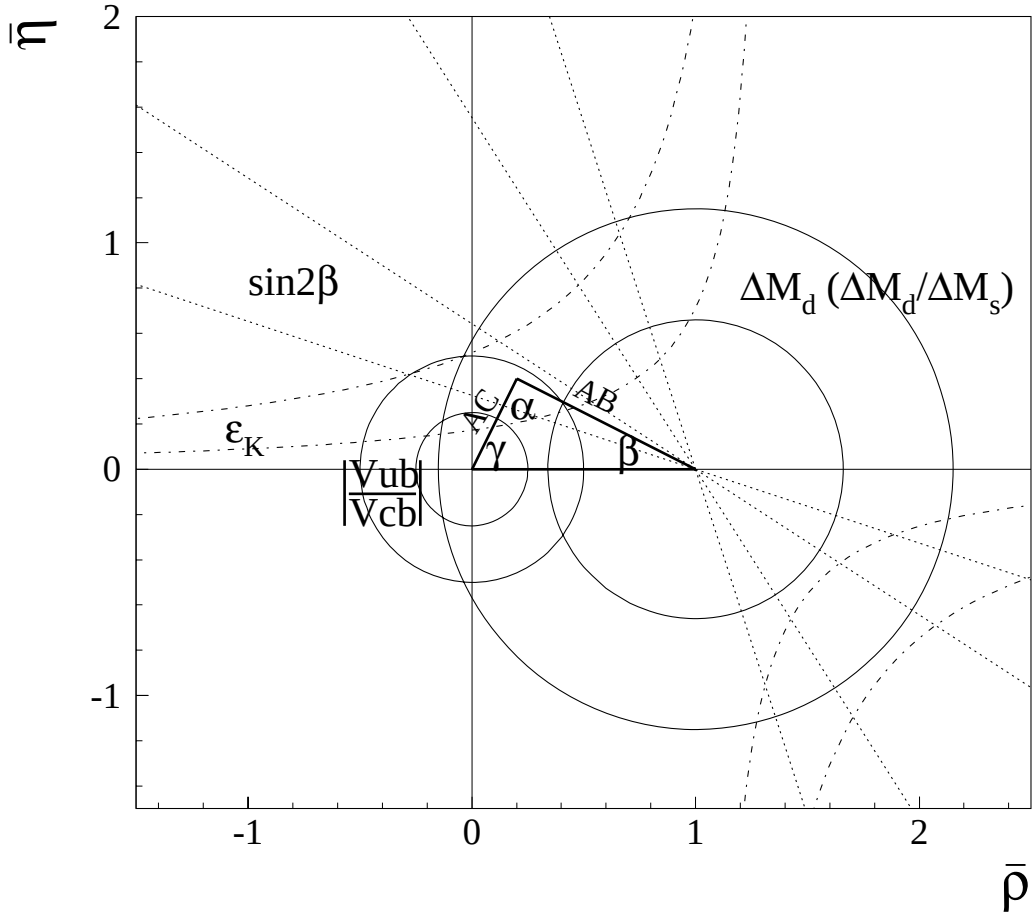


Fig. 5.1: *Constraints which are contributing to the Unitarity Triangle parameter determination.*

3. Statistical methods for CKM fits

In this Section we describe the basic ingredients for the different statistical approaches. The plots and the results presented here have to be taken as illustrations of the methods. Quantitative results and comparisons are given in the next Sections.

3.1. Bayesian methods

In this Section we describe the basic ingredients of the Bayesian approach and discuss the role of the systematic and theoretical uncertainties in deriving probability intervals for the relevant parameters.

Each of Eqs. (1, 2, 4, 6, 8) relates a constraint c_j to the parameters $\bar{\rho}$ and $\bar{\eta}$, via the set of additional parameters $\mathbf{x}$, where $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ stand for all experimentally determined or theoretically calculated quantities on which the various c_j depend (m_t, ξ ...)

$$c_j = c_j(\bar{\rho}, \bar{\eta}; \mathbf{x}). \tag{9}$$

In an ideal case of exact knowledge of c_j and $\mathbf{x}$, each of the constraints provides a curve in the $(\bar{\rho}, \bar{\eta})$ plane. In a realistic case, the analysis suffers from several uncertainties on the quantities c_j and $\mathbf{x}$. This means that, instead of a single curve (9) in the $(\bar{\rho}, \bar{\eta})$ plane, we have a family of curves which depends on the distribution of the set $\{c_j, \mathbf{x}\}$. As a result, the points in the $(\bar{\rho}, \bar{\eta})$ plane get different weights (even if they were taken to be equally probable *a priori*) and the *confidence* on the values of $\bar{\rho}$ and $\bar{\eta}$ clusters in a region of the plane.

The above arguments can be formalized by using the so called Bayesian approach (see [11] for an introduction). In this approach, the uncertainty is described in terms of a probability density function (pdf) which quantifies the confidence on the values of a given quantity. Applying Bayes Theorem in the case of a single constraint we obtain

$$\begin{aligned} f(\bar{\rho}, \bar{\eta}, c_j, \mathbf{x} \,|\, \hat{c}_j) &\propto f(\hat{c}_j \,|\, c_j, \bar{\rho}, \bar{\eta}, \mathbf{x}) \cdot f(c_j, \bar{\rho}, \bar{\eta}, \mathbf{x}) && (10)\\ &\propto f(\hat{c}_j \,|\, c_j) \cdot f(c_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \cdot f(\mathbf{x}, \bar{\rho}, \bar{\eta}) && (11)\\ &\propto f(\hat{c}_j \,|\, c_j) \cdot \delta(c_j - c_j(\bar{\rho}, \bar{\eta}, \mathbf{x})) \cdot f(\mathbf{x}) \cdot f_\circ(\bar{\rho}, \bar{\eta}) \,, && (12) \end{aligned}$$

where $\hat{c}_j$ is the experimental best estimate of c_j and $f_\circ(\bar{\rho}, \bar{\eta})$ denotes the *prior* distribution.

The various steps follow from probability rules, by assuming the independence of the different quantities and by noting that $\hat{c}_j$ depends on $(\bar{\rho}, \bar{\eta}, \mathbf{x})$ only via c_j. This is true since c_j is unambiguously determined, within the Standard Model, from the values of $\bar{\rho}$, $\bar{\eta}$ and $\mathbf{x}$.

The extension of the formalism to several constraints is straightforward. We can rewrite Eq. (10) as

$$f(\bar{\rho}, \bar{\eta}, \mathbf{x} \,|\, \hat{c}_1, ..., \hat{c}_{\mathrm{M}}) \;\propto\; \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \times \prod_{i=1,\mathrm{N}} f_i(x_i) \times f_\circ(\bar{\rho}, \bar{\eta}) \,. \tag{13}$$

M and N run over the constraints and the parameters respectively. In the derivation of (13), we have used the independence of the different quantities. By integrating Eq. (13) over $\mathbf{x}$ we obtain

$$f(\bar{\rho}, \bar{\eta} \,|\, \hat{\mathbf{c}}, \mathbf{f}) \propto \mathcal{L}(\hat{\mathbf{c}} \,|\, \bar{\rho}, \bar{\eta}, \mathbf{f}) \times f_\circ(\bar{\rho}, \bar{\eta}) \,, \tag{14}$$

where $\hat{\mathbf{c}}$ stands for the set of measured constraints, and

$$\mathcal{L}(\hat{\mathbf{c}} \,|\, \bar{\rho}, \bar{\eta}, \mathbf{f}) = \int \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \prod_{i=1,\mathrm{N}} f_i(x_i) \,\mathrm{d}x_i \tag{15}$$

is the effective overall likelihood which takes into account all possible values of x_i, properly weighted. We have written explicitly that the overall likelihood depends on the best knowledge of all x_i, described by $f(\mathbf{x})$. Whereas *a priori* all values for $\bar{\rho}$ and $\bar{\eta}$ are considered equally likely ($f_\circ(\bar{\rho}, \bar{\eta})$=cst), *a posteriori* the probability clusters around the point which maximizes the likelihood.

In conclusion, the final (unnormalized) pdf obtained starting from a uniform pdf for $\bar{\rho}$ and $\bar{\eta}$ is

$$f(\bar{\rho}, \bar{\eta}) \propto \int \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \prod_{i=1,\mathrm{N}} f_i(x_i) \,\mathrm{d}x_i \,. \tag{16}$$

The integration can be performed by Monte Carlo methods and the normalization is trivial. Starting from the pdf for $\bar{\rho}$ and $\bar{\eta}$, probability regions $P(w)$ are defined by the conditions:

$$\begin{array}{l} (\bar{\rho}, \bar{\eta}) \in P(w) \text{ if } f(\bar{\rho}, \bar{\eta}) \;>\; z_w \\ \int_{P(w)} f(\bar{\rho}, \bar{\eta}) d\bar{\rho} d\bar{\eta} = w \end{array}$$

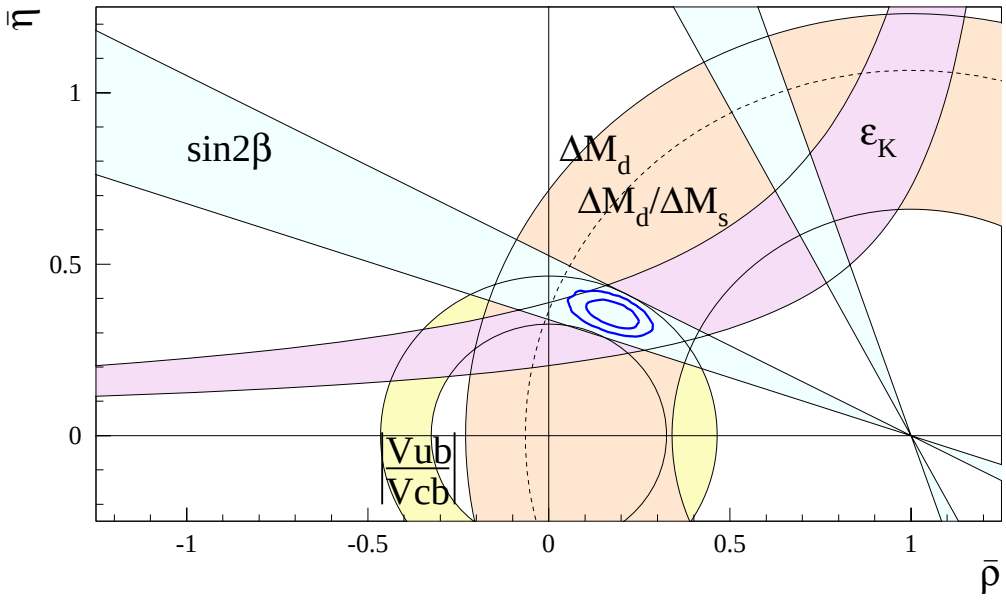


Fig. 5.2: *The contours at 68%, 95% probability regions in $\bar{\rho}$ and $\bar{\eta}$ as obtained using the Bayesian method, superimposed to the experimental constraints.*

An example of the typical output of this fit approach is shown in Fig. 5.2 where the probability regions at 68% and 95% are shown together with the experimental constraints.

It is important to note that this method does not make any distinction on whether the individual likelihood associated with some constraint is different from zero only in a narrow region (and we usually refer to this case as "measurement"), or if it goes to zero only on one side (e.g. when $c_j \to \infty$ or 0). In the latter case, the data only provide an upper/lower bound to the value of the constraint. This is precisely what happens, at present, with ΔM_s. Therefore, the experimental information about this constraint enters naturally in the analysis.

One of the feature of the Bayesian approach is that there is no conceptual distinction between the uncertainty due to random fluctuations, which might have occurred in the measuring process, the uncertainty about the parameters of the theory, and the uncertainty about systematics of not-exactly-known value, which can be both of experimental or theoretical origin (in the Bayesian jargon there are often indicated as influence parameters).

We can simply extend the notation to include in $\mathbf{x}$ these influence parameters responsible for the systematic uncertainty, and use Eq. (15) in an extended way. Irrespectively of the assumptions made on the pdf of $\mathbf{x}$, the overall likelihoods $f(\hat{c}_j)$ are approximately Gaussian because of a mechanism similar to the central limit theorem (i.e. just a matter of combinatorics). This makes the results largely stable against variations within choices of the distributions used to describe the uncertainties due to theory or systematics. For this reason we simplify the problem, by reducing the choice to only two possibilities. We choose a Gaussian model when the uncertainty is dominated by statistical effects, or there are many comparable contributions to the systematic errors, so that the central limit theorem applies ($\mathcal{G}(x - x_0)$). We choose a uniform pdf if the parameter is believed to be (almost) certainly in a given interval, and the points inside this interval are considered as equally probable. The second model is applied to some theoretical uncertainties. $\mathcal{U}(x) = 1/2\sigma_{\rm theo}$ for $x \in [x_0 - \sigma_{\rm theo}, x_0 + \sigma_{\rm theo}]$ and $\mathcal{U}(x) = 0$ elsewhere. The combined pdf $\mathcal{P}$ is then obtained by convoluting the Gaussian pdf $\mathcal{G}$ with the uniform pdf $\mathcal{U}$: $\mathcal{P} = \mathcal{G} \otimes \mathcal{U}$. When several determinations of the same quantity are available, the final p.d.f, in the Bayesian approach, is obtained by the product of the single pdfs.

An important point is how to evaluate the compatibility among individual constraints. In the CKM fits based on χ^2 minimization, a conventional evaluation of compatibility stems automatically from the value of the χ^2 at its minimum.

The compatibility between constraints in the Bayesian method is evaluated by comparing partial pdfs obtained when removing each constraint at a time. The integral over the overlap between the pdf with and without a given constraint quantifies the compatibility. In case of poor overlap the difference Δ_j between the two pdfs can be determined, for each constraint c_j, by substituting

$$c_j \to c_j(1 + \Delta_j). \tag{17}$$

Further investigation (based on physics and not on statistics) will be necessary to tell if the difference Δ_j comes from an incorrect evaluation of the input parameters or from new physics.

3.2. Frequentist methods

As said in the introduction theoretical quantities play an important role in the formulae relating the measured quantities to the UT parameters. These quantities are often inferred from theoretical calculations with uncertainties which can be associated to approximations. Uncertainties due to approximations are often estimated from more or less educated guesswork. For example, we recall that i) The quenched approximation in Lattice QCD calculations; ii) Model calculations of form factors where model parameters are varied within some range; iii) Higher order terms neglected in a power series for which the error is estimated from the "natural size" of the expansion parameter or a scale dependence in a perturbative series where the error is estimated by varying the scale within some *reasonable* range. This has driven the developments of statistical approaches based on a frequentist understanding of systematic theoretical uncertainties, which cannot be treated as statistically distributed quantities.

In this framework two approaches are presented: the *R*fit method and the Scanning method. In both methods, the "theoretically allowed" values for some theoretical parameters are "scanned", *i.e.* no statistical weight is assigned to these parameters as long as their values are inside a "theoretically allowed" range.

The *R*fit method starts by choosing a point in a parameter subspace of interest, *e.g.* a point in the $\bar{\rho}$-$\bar{\eta}$ plane, and ask for the best set of theoretical parameters for this given point. This set is determined by minimizing a χ^2 function with respect to all model parameters, except $\bar{\rho}$ and $\bar{\eta}$. The theoretical parameters are free to vary inside their theoretically allowed range without obtaining any statistical weight. In this way, upper limits of confidence levels in the parameter subspace of interest can be determined.

The basic idea of the Scanning method is to choose a possible set of values for the theoretical parameters and to ask whether this particular model leads to an acceptable fit for the given data set. If so, a confidence contour is drawn in a parameter subspace of interest, *e.g.* the $\bar{\rho}$-$\bar{\eta}$ plane, representing the constraints obtained for this particular set of model parameters. This procedure is repeated for a large number of possible theoretical models by scanning allowed ranges of the non-perturbative parameters. The single confidence level contours cannot be compared from a statistical point of view. This method has been extended to facilitate an analysis of the relative influence of experimental and theoretical uncertainties in determining the consistency of the measurements with theory.

3.2.1. The Rfit approach

The CKM analysis using the *R*fit method * is performed in three steps: 1. Testing the overall consistency between data and the theoretical framework, here the SM. 2. If data and the SM are found to be in reasonable agreement, confidence levels (CL) in parameter subspaces are determined. 3. Testing extensions of the SM.

*The *R*fit method is implemented in the software package *CKMfitter* [12]. More details can be found in [3].

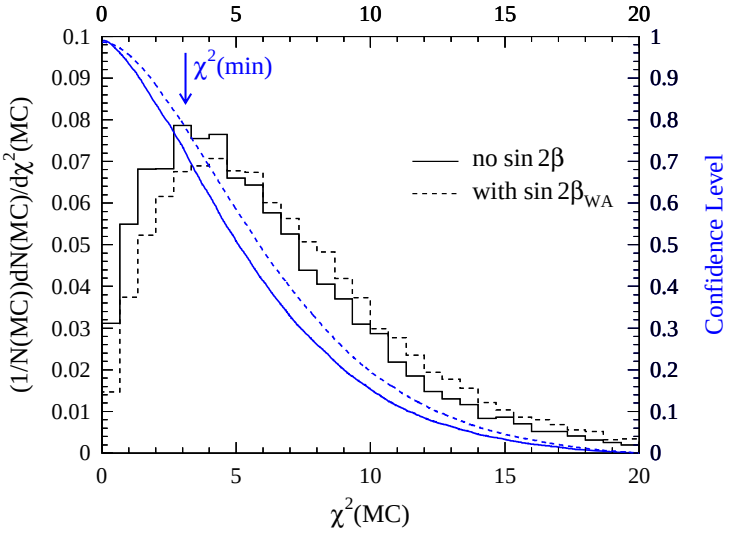


Fig. 5.3: *Determination of CL(SM). The histograms show the test statistics $F(\chi^2)$ built from the Monte Carlo technique as described in the text for two different fits, including or excluding the* $\sin 2\beta$ *measurement. Integration of the distributions above* χ^2_{min} *provides CL(SM).*

The quantity $\chi^2 = -2\ln\mathcal{L}(y_{\rm mod})$ is minimized in the fit, where the likelihood function is defined by $\mathcal{L}(y_{\rm mod}) = \mathcal{L}_{\rm exp}(x_{\rm exp} - x_{\rm theo}(y_{\rm mod}))\cdot\mathcal{L}_{\rm theo}(y_{\rm QCD})$. The experimental part, $\mathcal{L}_{\rm exp}$, depends on measurements, $x_{\rm exp}$, and theoretical predictions, $x_{\rm theo}$, which are functions of model parameters, $y_{\rm mod}$. The theoretical part, $\mathcal{L}_{\rm theo}$, describes our "knowledge" of the theoretical parameters, $y_{\rm QCD}\in\{y_{\rm mod}\}$. We set $\mathcal{L}_{\rm theo}=1$ within an "allowed range" $\mathcal{R}$ provided by a theoretical estimate, and $\mathcal{L}_{\rm theo}=0$ outside $\mathcal{R}$. That is, the $y_{\rm QCD}$ are free to vary within $\mathcal{R}$ without changing the $\mathcal{L}_{\rm theo}$ part of the χ^2 function. It should be kept in mind that the choice of $\mathcal{R}$ is statistically not well-defined and reflects an intrinsic problem of all statistical analyses when dealing with theoretical uncertainties.

It is worthwhile to emphasize that a uniform likelihood function is not identical to a uniform pdf. Whereas a uniform likelihood means that the theoretical parameter is free to vary within $\mathcal{R}$, a uniform pdf states that each value within $\mathcal{R}$ has equal probability and hence introduces a statistical weight. This has important consequences if more than one theoretical parameter enter a constraint or if the constraint depends on a nonlinear function of a theoretical parameter. For example, the ε_K constraint depends on the product $P=\hat{B}_K\cdot|V_{cb}|^4$. The theoretical likelihood for $|V_{cb}|^4$ reads $\mathcal{L}_{\rm theo}(|V_{cb}|^4)=1$ for all theoretically allowed $|V_{cb}|$ values given by $\mathcal{L}_{\rm theo}(|V_{cb}|)=1$. The theoretical likelihood for the product P reads $\mathcal{L}_{\rm theo}(P)=1$ for any value of $\hat{B}_K$ and $|V_{cb}|$ given by $\mathcal{L}_{\rm theo}(\hat{B}_K)=1$ and $\mathcal{L}_{\rm theo}(|V_{cb}|)=1$, respectively. That is, in *R*fit, no statistical weight is introduced for any value of P, independent of the fact whether the single theoretical parameters are bound or unbound and independent of the particular parametrization chosen. On the contrary, the pdf for the theoretical part of $|V_{cb}|^4$ is proportional to $(|V_{cb}|)^{-3/4}$ if the theoretical pdf for $|V_{cb}|$ is chosen to be uniform. The pdf for the product P would be proportional to $-\log|P|$ in leading order if the pdfs for $\hat{B}_K$ and $|V_{cb}|^4$ were chosen to be uniform [3].

The agreement between data and the SM is gauged by the global minimum $\chi^2_{{\rm min};y_{\rm mod}}$, determined by varying freely all model parameters $y_{\rm mod}$. For $\chi^2_{{\rm min};y_{\rm mod}}$, a confidence level CL(SM) is computed by means of a Monte Carlo simulation. For the optimal set of model parameters $y_{\rm mod}$, a large number of pseudo-measurements is generated using the experimental likelihood function $\mathcal{L}_{\rm exp}$. For each set of pseudo-measurements, the minimum $\chi^2_{\rm min}$ is determined and used to build a test statistics $F(\chi^2)$. The CL is then calculated as ${\rm CL(SM)}=\int_{\chi^2_{{\rm min};y_{\rm mod}}}^{\infty}F(\chi^2)d\chi^2$ as illustrated in Fig. 5.3. If there is a hint of an incompatibility between data and the SM one has to investigate in detail which constraint leads to a small value for ${\rm CL(SM)}$.

If the hypothesis "the CKM picture of the SM is correct" is accepted, CLs in parameter subspaces a, *e.g.* $a = (\bar{\rho}, \bar{\eta}), \sin 2\beta, ...$, are evaluated. For a given point in a, one determines the best agreement between data and theory. One calculates $\Delta\chi^2(a) = \chi^2_{\mathrm{min};\mu}(a) - \chi^2_{\mathrm{min};y_{\mathrm{mod}}}$, by varying freely all model parameters μ (including y_{QCD}) with the exception of a. The corresponding CL is obtained from $\mathrm{CL}(a) = \mathrm{Prob}(\Delta\chi^2(a), N_{\mathrm{dof}})$ (see e.g. Fig. 5.3) where N_{dof} is the number of degrees of freedom, in general the dimension of the subspace a. It has to be stressed that $\mathrm{CL}(a)$ depends on the choice of $\mathcal{R}$. The usage of $\mathrm{Prob}(\Delta\chi^2(a), N_{\mathrm{dof}})$ assumes Gaussian shapes for $\mathcal{L}_{\mathrm{exp}}$. The CL obtained has been verified for several examples using a Monte Carlo simulation similar to the one described in the last section.

If the SM cannot accommodate the data, the analysis has to be repeated within extensions of the SM. Even in the case of a good agreement between data and the SM, it is worthwhile to perform the analysis for possible extensions of the SM in order to constrain New Physics parameters, see *e.g.* Ref. [13], or to determine the precision needed to study or exclude certain models.

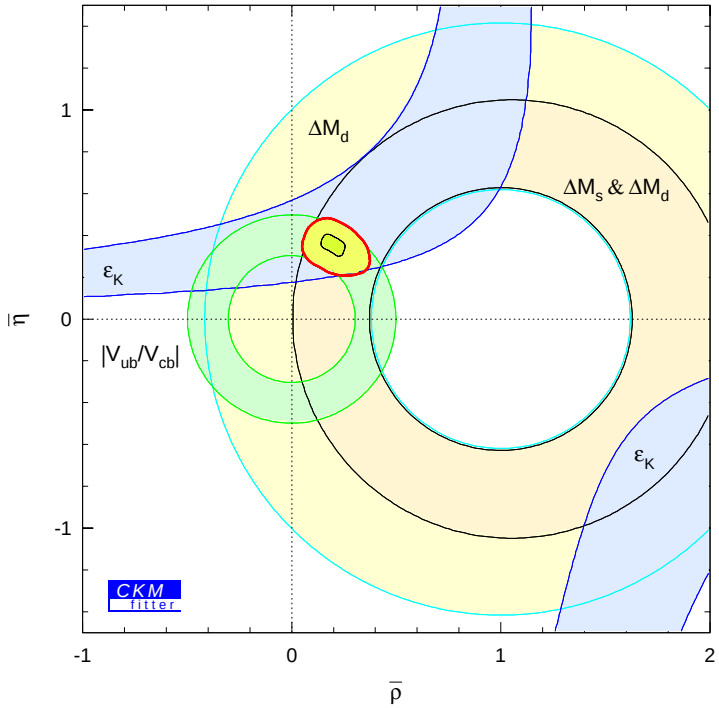


Fig. 5.4: *Different single constraints in the $\bar{\rho} - \bar{\eta}$ plane shown at 95 % CL contours. The 95 % and 10 % CL contours for the combined fit are also shown.*

3.2.2. The Scanning Method

In the scanning method the following procedure is set to deal with "not-statistically" distributed quantities: we select a specific set of theoretical parameters called a "model",

$$\mathcal{M} \equiv \{\mathcal{F}_{D*}(1),\ \tilde{\Gamma}^c_{incl},\ \tilde{\Gamma}^u_{excl}, \tilde{\Gamma}^u_{incl},\ F_{B_d}\sqrt{\hat{B}_{B_d}}, \hat{B}_K,\ \xi,\ \eta_1, \eta_2,\ \eta_3,\ \eta_B\}, \tag{18}$$

where $\mathcal{F}_{D*}(1)$ is the Isgur-Wise function of $\mathrm{B} \to \mathrm{D}^* \ell\nu$ at zero recoil corrected for finite b-quark mass, $\tilde{\Gamma}^c_{incl}$ denotes the reduced decay rate for $b \to c\ell\nu$, $\tilde{\Gamma}^u_{excl}$ ($\tilde{\Gamma}^u_{incl}$) represents the reduced decay rate for $\mathrm{B} \to \rho\ell\nu$ ($b \to u\ell\nu$), F_{B_d} (F_{B_s}) is the B^0_d (B^0_s) decay constant, B_{B_d}, B_{B_s} and $\hat{B}_K$ parameterize the values of the hadronic matrix elements appearing in $\mathrm{B}^0_d - \overline{\mathrm{B}}^0_d$ mixing, $\mathrm{B}^0_s - \overline{\mathrm{B}}^0_s$ mixing and $\mathrm{K}^0 - \overline{\mathrm{K}}^0$ mixing, respectively, $\xi = F_{B_s}/F_{B_d}\sqrt{(\hat{B}_{B_s}/\hat{B}_{B_d})}$, and η_1, η_2, η_3, and η_B denote QCD parameters. Such a set of theoretical parameters carries by definition non-probabilistic uncertainties but still may

involve probabilistic errors. By choosing many different "models" we map out the allowed ranges of these theoretical parameters.

For each "model" $\mathcal{M}$ we construct and minimize the function

$$\chi^2_{\mathcal{M}}(A,\bar{\rho},\bar{\eta}) = \sum_i \left[\frac{E_i - \mathcal{E}_i(A,\bar{\rho},\bar{\eta};C_k;\mathcal{M})}{\sigma_{E_i}}\right]^2, \tag{19}$$

where the E_i are observables based on measured quantities, $\mathcal{E}_i(A,\bar{\rho},\bar{\eta};C_k;\mathcal{M})$ is their parameterization in terms of A, $\bar{\rho}$, and $\bar{\eta}$, C_k denotes measured quantities that possess experimentally derived or other probabilistic uncertainties, such as masses and lifetimes, and the σ_{E_i} denote all measurement uncertainties contributing to both E_i and $\mathcal{E}_i(A,\bar{\rho},\bar{\eta};C_k;\mathcal{M})$. This includes all uncertainties on the theoretical parameters that are statistical in nature.

The inputs used are those given in Table 5.1. To incorporate results on ΔM_s searches we include a χ^2-term defined as the maximum between the log-likelihood ratio used in [2] and 0:

$$-2ln\mathcal{L}_\infty(\Delta M_s) = \max\left(\frac{(1-2\mathcal{A})}{\sigma^2_{\mathcal{A}}}, 0\right) \tag{20}$$

$\mathcal{A}$ is the amplitude spectrum as function of ΔM_s.

The minimization solution $(A,\bar{\rho},\bar{\eta})_{\mathcal{M}}$ for a particular "model" $\mathcal{M}$ incorporates no prior distribution for non-probabilistic uncertainties of the theoretical parameters and meets the frequency interpretation. All uncertainties depend only on measurement errors and other probabilistic uncertainties including any probabilistic component of the uncertainties on the theoretical parameters relevant to each particular measurement. At the moment, for practical reasons, we have treated the comparatively small uncertainties arising from η_1, η_2, η_3, and η_B as probabilistic. The effects of this simplification will be explored in future fits.

A "model" $\mathcal{M}$ and its best-fit solution are kept only if the probability of the fit satisfies $\mathcal{P}(\chi^2_{\mathcal{M}}) > \mathcal{P}_{min}$, which is typically chosen to be 5%. For each "model" $\mathcal{M}$ accepted, we draw a 95% C.L. contour in the $(\bar{\rho},\bar{\eta})$ plane. The fit is repeated for other "models" $\mathcal{M}$ by scanning through the complete parameter space specified by the theoretical uncertainties. This procedure derives from the technique originally described in [14].

The χ^2 minimization thus serves three purposes:

1. If a "model" $\mathcal{M}$ is consistent with the data, we obtain the best estimates for the three CKM parameters, and 95% C.L. contours are determined.
2. If a "model" $\mathcal{M}$ is inconsistent with the data the probability $\mathcal{P}(\chi^2_{\mathcal{M}})$ will be low. Thus, the requirement of $\mathcal{P}(\chi^2_{\mathcal{M}})_{min} > 5\%$ provides a test of compatibility between data and its theoretical description.
3. By varying the theoretical parameters beyond their specified range we can determine correlations on them imposed by the measurements. The first results of this study are shown in Section 4.2.

If no "model" were to survive we would have evidence of an inconsistency between data and theory, independent of the calculations of the theoretical parameters or the choices of their uncertainties. Since the goal of the CKM parameter fits is to look for inconsistencies of the different measurements within the Standard Model, it is important to be insensitive to artificially produced effects and to separate the non-probabilistic uncertainties from Gaussian-distributed errors.

In order to demonstrate the impact of the different theoretical quantities on the fit results in the $(\bar{\rho},\bar{\eta})$ plane, Figs. 5.5a–f show contours for fits in which only one parameter was scanned while the others were kept at their central values. These plots demonstrate the impact of the model dependence in $|V_{ub}|$ and $|V_{cb}|$ as well as that of $F_{B_d}\sqrt{\hat{B}_{B_d}}$ and $\hat{B}_K$, ξ, and η_1, respectively. For each parameter we consider

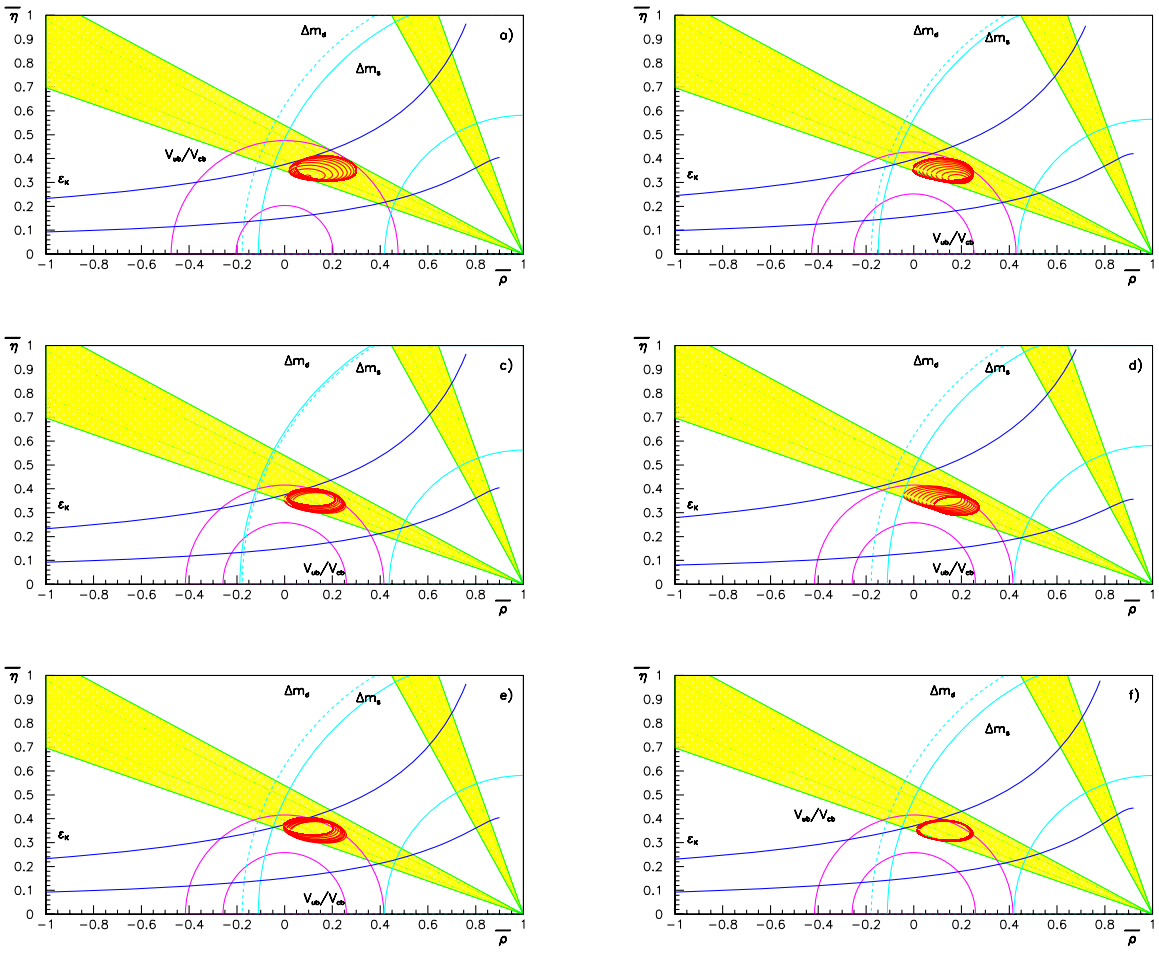


Fig. 5.5: *Contours of different models in the $(\bar{\rho},\bar{\eta})$ plane, by varying only one theoretical parameter at a time, a) $\tilde{\Gamma}_{excl}$, b) $\mathcal{F}_{D^*}(1)$, c) $F_{B_d}\sqrt{\hat{B}_{B_d}}$, d) $\hat{B}_K$, e) ξ (where ΔM_s is included in the fit), and f) η_1. In each plot nine different models are considered by varying the theoretically-allowed range from the minimum value to the maximum value. The figures are arranged with a) in the upper left, b) in the upper right, etc..*

nine different models which span its range equidistantly, starting with the smallest allowed value. Since these plots just serve illustrative purposes we use only the measurements of $|V_{ub}|$, $|V_{cb}|$, ΔM_d, and ϵ_K in the fits, except for Fig. 5.5e where the information of ΔM_s has been included in addition. To guide the eye we show the boundaries of the three bands for $|V_{ub}/V_{cb}|$, $|V_{td}/V_{cb}|$, and ϵ_K. Since the theoretical parameters are kept at their central values except for the one being varied, the bands corresponding to the other parameters reflect only experimental uncertainties.

We now turn to scanning all parameters simultaneously within their theoretically "allowed" ranges. Figure 5.6 shows the resulting contours for a set of representative "models", when all available constraints are included. Note that there is no frequency interpretation for comparing which models are to be "preferred", other than the statement that at most one model is correct. In this analysis we cannot, and do not, give any relative probabilistic weighting among the contours, or their overlap regions. Indeed, the entire purpose of the scanning method is to make clear the relative importance of measurable experimental errors and a-priori unknown theoretical uncertainties.

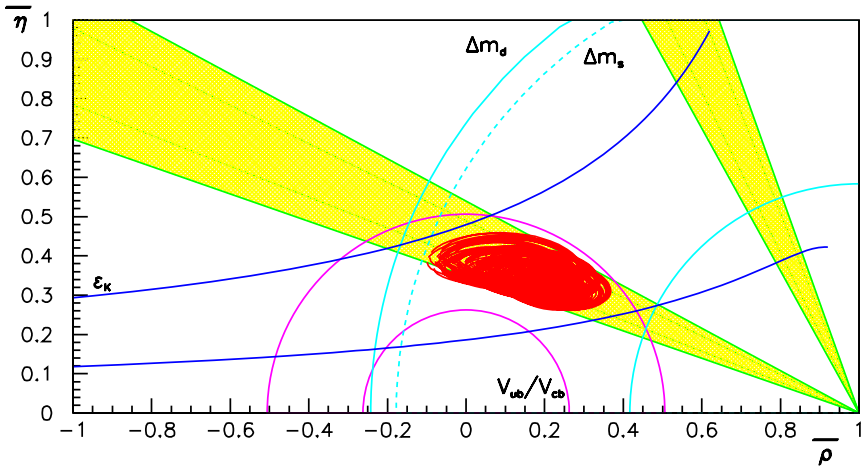


Fig. 5.6: *Contours in the* $(\bar{\rho}, \bar{\eta})$ *plane for different models, scanning theoretical parameters* $\tilde{\Gamma}_{excl}$, $\mathcal{F}_{D^*}(1)$, $F_{B_d}\sqrt{\hat{B}_{B_d}}$, $\hat{B}_K$, *and* ξ, *based on measurements of* $|V_{ub}|$, $|V_{cb}|$, ΔM_d, ϵ_K, *the amplitude for* ΔM_s, *and* $\sin 2\beta$.

4. Impact of the uncertainties on theoretical quantities in CKM Fits

As described in the previous sections, the "correct" way to treat the theoretical parameters is not unambiguously defined but depends on the adopted statistical approach. In this section we will not discuss the problems and the virtues of the different statistical approaches on this point, and concentrate on the impact of the uncertainties on theoretical parameters in constraining $\bar{\rho}$ and $\bar{\eta}$.

Two numerical analyses will be presented: one in the Bayesian framework and one in the frequentist framework. In the first analysis we study the effect on the UT fit from a modification (or a removal) of some theoretical parameter used as input parameter. The second analysis introduces a graphical method to represent, in the space of the theoretical parameters, the goodness of the UT fit and to evaluate the relevance of the knowledge on these parameters.

4.1. Bayesian analysis

In the framework of the Bayesian method the input knowledge is expressed in terms of pdfs for all quantities (theoretical and experimental parameters). Following the procedure described in Section 3.1., the output pdf can be computed for $\bar{\rho}$, $\bar{\eta}$ and for any other quantity of interest.

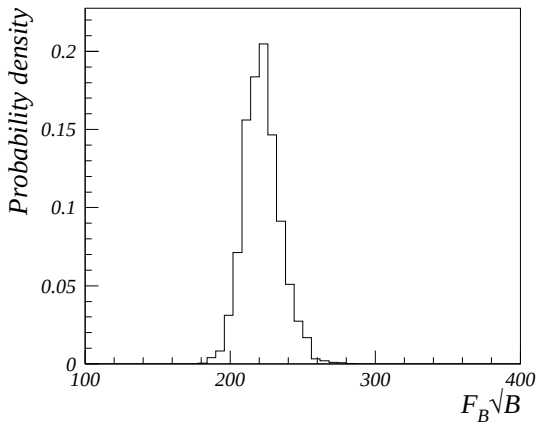
The impact of the uncertainty on a given quantity, which enters as a parameter in a given constraint, is naturally evaluated by comparing the results obtained excluding the corresponding constraint or by varying the error of the input parameter. When the information on a certain quantity is excluded the corresponding input pdf is taken as uniform. The common set of inputs used for this analysis are the ones available at the time of the Workshop (Table 5.1).

4.1.1. Determination of $F_{B_d}\sqrt{\hat{B}_{B_d}}$

First we consider the $F_{B_d}\sqrt{\hat{B}_{B_d}}$ parameter. Quite remarkably the remaining constraints determine precisely this quantity and, from the output distribution shown in the left part of Fig. 5.7, we get

$$F_{B_d}\sqrt{\hat{B}_{B_d}} = (223 \pm 12)\ \text{MeV} \tag{21}$$

This is in perfect agreement with the results from lattice calculation (see Table 5.1) and has a significantly smaller error. This suggests that, unless the lattice error on $F_B\sqrt{\hat{B}_B}$ does not become smaller than 12 MeV, the theoretical knowledge of this quantity is not quite relevant in UT fits. The Table in Fig. 5.7 quantifies the effect of changing the uncertainty on $F_{B_d}\sqrt{\hat{B}_{B_d}}$ (keeping the same central value).



$\sigma \pm \Delta/2$	$\bar{\rho}$	$\bar{\eta}$
$\pm 16 \pm 6$	0.183 ± 0.040	0.355 ± 0.027
$\pm 33 \pm 12$	0.173 ± 0.046	0.357 ± 0.027
$\pm 66 \pm 24$	0.173 ± 0.046	0.355 ± 0.027
∞	0.175 ± 0.049	0.355 ± 0.027

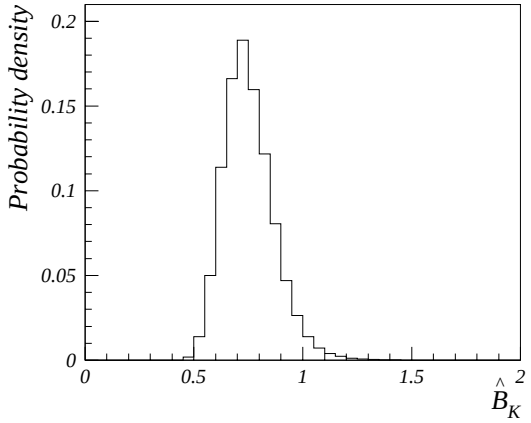
Fig. 5.7: *Left: output distribution for $F_{B_d}\sqrt{\hat{B}_{B_d}}$ assuming a flat input distribution. Right: table reporting the results of the UT fit for $\bar{\rho}$ and $\bar{\eta}$ assuming different input values for the errors on $F_{B_d}\sqrt{\hat{B}_{B_d}}$ (in MeV) (σ is the Gaussian error and $\Delta/2$ is the half-width of the systematic range). The last column ("infinite error") is obtained with a uniform input distribution.*

4.1.2. Determination of $\hat{B}_K$

Here the same exercise has been repeated with the parameter $\hat{B}_K$. Assuming for $\hat{B}_K$ a uniform input distribution between 0 and 2, from the output distribution shown in Fig. 5.8 we obtain

$$\hat{B}_K = 0.73^{+0.13}_{-0.07} \tag{22}$$

The fitted value is again in perfect agreement with the lattice value (see Table 5.1), but in this case the fitted (output) uncertainty is similar to the theoretical (input) one. We then expect that "lattice information" plays a non negligible role, in particular in the determination of $\bar{\eta}$ (because of simple geometrical arguments). Table in Fig. 5.8 shows that, in fact, removing the information coming from Lattice QCD (last row) the error on $\bar{\eta}$ increases by 50%.

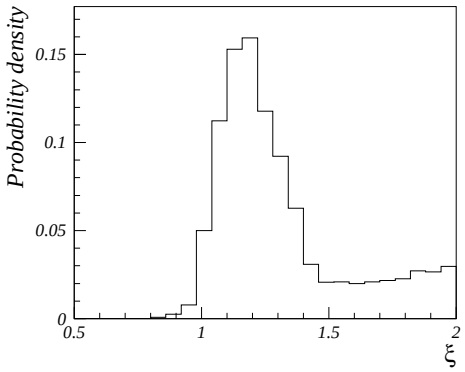


$\sigma \pm \Delta/2$	$\bar{\rho}$	$\bar{\eta}$
$\pm 0.03 \pm 0.065$	0.181 ± 0.040	0.349 ± 0.025
$\pm 0.06 \pm 0.13$	0.173 ± 0.046	0.357 ± 0.027
$\pm 0.12 \pm 0.26$	0.163 ± 0.052	0.365 ± 0.030
∞	0.161 ± 0.055	0.361 ± 0.042

Fig. 5.8: *Left: output distribution for $\hat{B}_K$ assuming a flat input distribution. Right: table reporting the results of the CKM Fits for $\bar{\rho}$ and $\bar{\eta}$ assuming different input values for the errors on $\hat{B}_K$ (σ is the Gaussian error and $\Delta/2$ is the half-width of the systematic range). The last column ("infinite error") is obtained with a flat input distribution.*

4.1.3. Determination of ξ

Since ΔM_s has not yet been measured, ξ cannot be determined by the data. Assuming a uniform distribution between 0.6 and 2 (the upper bound is obviously arbitrary), the output distribution shown in Fig. 5.9 is obtained. The tail on the right part of the plot shows that, at present, ξ is only weakly constrained by experimental data; for this reason the information on the ξ parameter is very important, in particular in the determination of $\bar{\rho}$, as shown in the table in Fig. 5.9.



$\sigma \pm \Delta/2$	$\bar{\rho}$	$\bar{\eta}$
$\pm 0.02 \pm 0.025$	0.183 ± 0.037	0.353 ± 0.025
$\pm 0.04 \pm 0.05$	0.173 ± 0.046	0.357 ± 0.027
$\pm 0.08 \pm 0.10$	0.167 ± 0.062	0.357 ± 0.030
$[0.4, 2]$	0.203 ± 0.080	0.347 ± 0.035

Fig. 5.9: *Left: output distribution for ξ assuming a flat input distribution between 0.6 and 2. Right: table reporting the results of the CKM Fits for $\bar{\rho}$ and $\bar{\eta}$ assuming different input values for the errors on ξ (σ is the Gaussian error and $\Delta/2$ is the half-width of the systematic range). The last column is obtained with a flat input distribution between 0.6 and 2.*

4.2. Scan analysis

In the study of the sensitivity of the UT fit to a given theoretical parameter one should define how to treat the remaining parameters. In the Bayesian approach (described in the previous section) the remaining parameters are integrated using their input pdf while in the standard frequentist approach the confidence level for a parameter is computed irrespectively of the values of all the remaining parameters (logical "OR" over the values of the parameters).

The technique presented here aims at studying and visualizing the sensitivity of UT fits to theoretical uncertainties, in the theoretical parameters space (T), minimizing a priori inputs and intermediate combinations of parameters. The method tries to represent pairs or triplets of theoretical parameters while keeping some information on the remaining (undisplayed) parameters. The input knowledge on a theoretical parameter is described by a "nominal central value" and a "theoretical preferred range". In two dimensions the procedure is as follows:

- Pick two of the parameters T for display. Call these the *primary parameters*, T_1 and T_2.
- Pick a third T parameter, the *secondary parameter* T_s. This parameter is singled out for special attention to the effects of projecting over it.
- Call all the other T parameters the *undisplayed parameters*, T_X.
- For each point P in the grid of scanned values of $T_1 \otimes T_2$, a number of fits will have been attempted, covering all the scanned values of T_s and T_X. For each P, evaluate the following hierarchy of criteria against the ensemble of results of these fits, deriving for the point a value, we call it the "Level", which is an integer between 0 and 5 inclusive:

  1. Define a minimum acceptable value for $P(\chi^2)$. Did any of the fits for P pass this cut? If not, assign Level = 0 and stop; otherwise assign Level = 1 and continue.
  2. Did any of the remaining fits lie within the "theoretically preferred region" for *all* the undisplayed parameters T_X? If not, stop; if yes, assign Level = 2 and continue.

3. Did any of the remaining fits have the secondary parameter T_s within its "theoretically preferred region"? If not, stop; if yes, assign Level = 3 and continue.
4. Did any of the remaining fits have T_s equal to its "nominal central value"? (That value must have been included in the scan grid for this to make sense.) If not, stop; if yes, assign Level = 4 and continue.
5. Did any of the remaining fits have *all* the undisplayed parameters T_X also at their "nominal central values"? If not, stop; if yes, assign Level = 5 and stop.

- Now display contours of the quantity Level over the grid in the $T_1 \otimes T_2$ plane. Assign a unique color to each parameter T, so the contours for T_s at Level = 3,4 are drawn in the color corresponding to that parameter. The contours for Level = 4,5, which correspond to restrictions of parameters exactly to their central values, are also drawn distinctively, with dashing.
  The Level 3 contour (solid, colored), in particular, displays the allowed region, at the selected confidence level, for T_1 and T_2, based on the experimental data and on limiting all other theoretical parameters to their preferred ranges. Study of the relative spacing of the Level 2, 3, and 4 contours readily reveals the effects of the application of the T_s bounds on the fit results.
- Overlay the contours with straight lines showing the theoretically preferred ranges and nominal central values for T_1 and T_2, in their respective unique colors, again with dashing for the central value. This allows the theoretical bounds on T_1 and T_2 to be evaluated directly for consistency against all other available data, yet avoiding any convoluted use of priors for these two parameters. Comparison of these theoretical bounds for T_1 and T_2 with the Level 3 contour that shows the experimental information, constrained by the application of the theoretical bounds on T_s and the T_x, allows a direct visual evaluation of the consistency of all available information, with the effects of the application of all theoretical bounds manifest, not obscured by convolutions performed in the fit itself.

Figure 5.10 shows the results of the previous procedure, using $F_B\sqrt{\hat{B}_B}$ and V_{ub} as primary parameters, $\hat{B}_K$ as secondary parameter, while the undisplayed parameter (there is just one in this case) is ξ. What can be seen immediately is that the entire theoretically allowed region for the primary parameters, shown by the crossing of the solid lines, is consistent with all the other data, including the theoretical bound on $\hat{B}_K$, and that even when all parameters are constrained to their central values the resulting fit (there can be only one at that point) is fully consistent. Changing the role of primary, secondary, and undisplayed parameters in many different ways, helps to understand the role of these parameters in the fit.

These plots can be extended in three dimensions by drawing the three bi-dimensional projections of the allowed region. Several three dimensional plots and further details can be found in [15].

5. Fit comparison

In this section we compare the results on the CKM quantities obtained following two approaches: Bayesian and *R*fit. The common set of inputs are the ones available at the time of the Workshop (Table 5.1). The Scan method has not been included in the comparison because it does not evaluate overall allowed regions for the CKM parameters. As explained in the previous sections, the main difference between the Bayesian and the *R*fit analyses originates from the computation of the Likelihood functions (identified with pdfs in the Bayesian case) for each parameter and in the treatment of the Likelihood fit.

5.1. Input likelihoods and combination of likelihoods

In general a determination of a given quantity is characterized by a central value, a statistical error and a systematical error. Starting from such a splitting of the errors Bayesian and frequentist approaches may describe this quantity according to different likelihood functions.

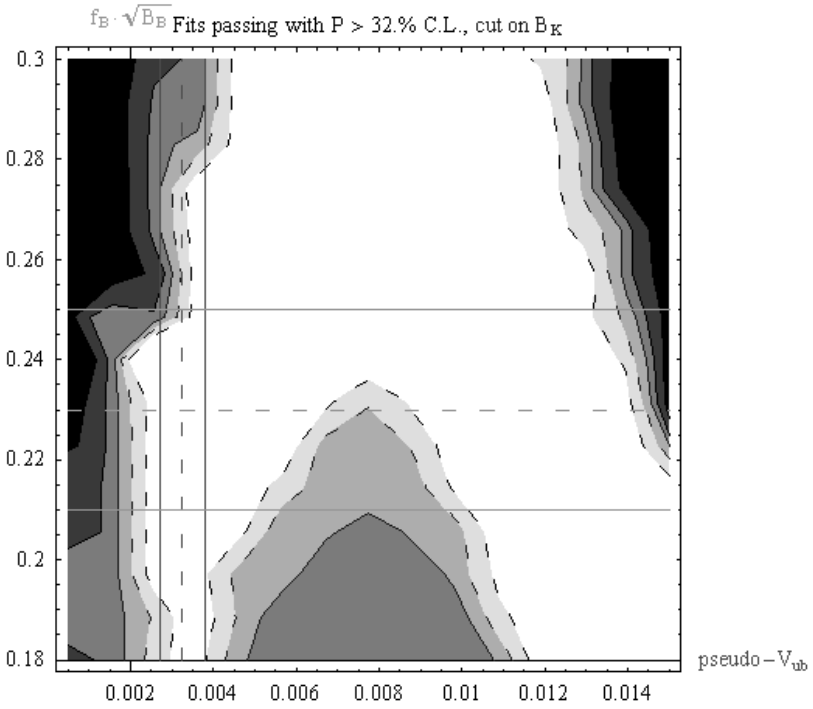


Fig. 5.10: *Results of the procedure described in the text, using $F_{B_d}\sqrt{\hat{B}_{B_d}}$ and V_{ub} as primary parameters, $\hat{B}_K$ as secondary parameter, and ξ as undisplayed parameter.*

In the Bayesian approach, the basic assumption is that the value of any quantity is distributed according to a pdf. The final pdf of a certain quantity is obtained by convoluting the pdfs corresponding to the different uncertainties affecting the quantity. In particular, the uncertainty on a quantity is usually splitted in two parts: a statistical part which can be described by a Gaussian pdf, $\mathcal{G}(x - x_0)$ (this part may contain many sources of uncertainties which have been already combined into a single pdf) and another part which is usually of theoretical origin and is often related to uncertainties due to theoretical parameters. In the following we will denote it as theoretical systematics. It is often described using an uniform pdf: $\mathcal{U}(x) = 1/2\sigma_{\rm theo}$ for $x \in [x_0 - \sigma_{\rm theo}, x_0 + \sigma_{\rm theo}]$ and $\mathcal{U}(x) = 0$ elsewhere. The combined pdf $\mathcal{P}$ is then obtained by convoluting the Gaussian pdf $\mathcal{G}$ with the uniform pdf $\mathcal{U}$: $\mathcal{P} = \mathcal{G} \otimes \mathcal{U}$.

In the frequentist analysis, no statistical meaning is attributed to the uncertainties related to theoretical systematics. The likelihood function $\mathcal{L}$ for the quantity x contains a statistical part, $\mathcal{L}_{\rm exp}(x - x'_0)$, described by a Gaussian with mean value x'_0, and a "not-statistical" part, $\mathcal{L}_{\rm theo}(x'_0)$. The function $\mathcal{L}_{\rm theo}(x'_0)$, as denoted *R*fit likelihood, is a uniform function $\mathcal{L}_{\rm theo}(x'_0) = 1$ for $x'_0 \in [x_0 - \sigma_{\rm theo}, x_0 + \sigma_{\rm theo}]$ and $\mathcal{L}_{\rm theo}(x'_0) = 0$ elsewhere. The final likelihood is given by the product $\mathcal{L} = \mathcal{L}_{\rm exp}(x - x'_0) \cdot \mathcal{L}_{\rm theo}(x'_0)$. In conclusion, when a quantity contains an uncertainty to which the frequentists do not attribute any statistical meaning, the likelihood which describes this quantity is obtained as a product between this uncertainty and the statistical one.

When several determinations of the same quantity are available one may combine them to obtain a single input for a quantity entering the fit (these considerations apply for example to the determinations of $|V_{ub}|$ and $|V_{cb}|$). We suppose, in the following, that these determinations are not correlated. In addition, it is assumed that the various determinations of these quantities are compatible. Then, for the combination, the Bayesian approach calculates the product of the single pdfs, whereas the frequentist approach calculates the product of the individual likelihoods. Hence, the mathematical concepts for the

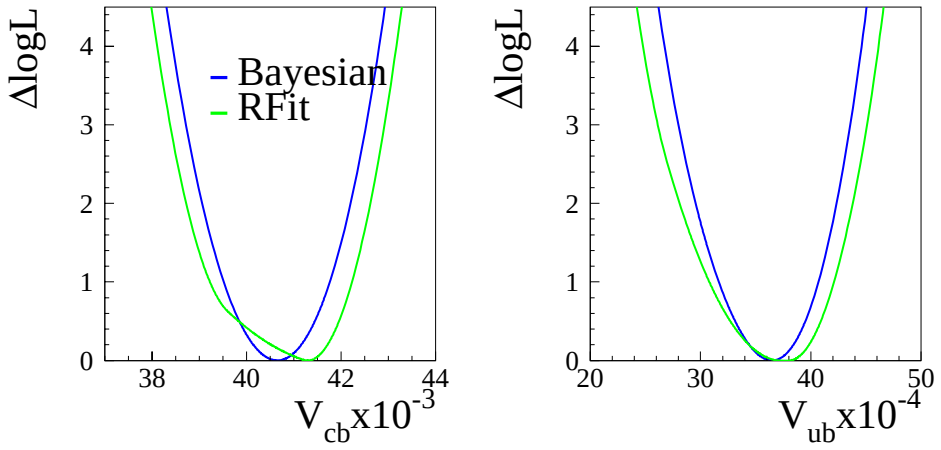


Fig. 5.11: *The ΔLikelihood for $|V_{cb}|$ and $|V_{ub}|$ using the Bayesian and frequentist approaches when combining the inclusive and the exclusive determinations.*

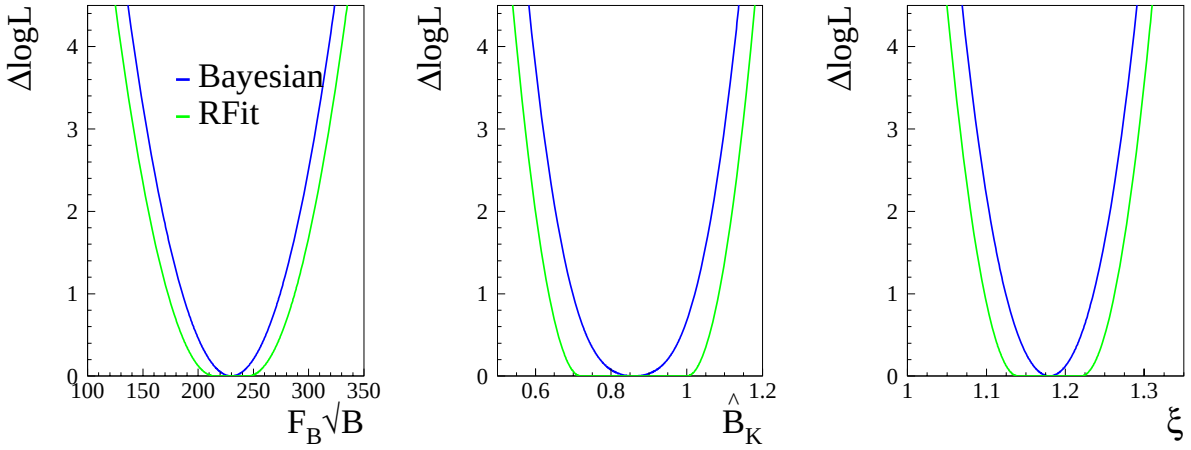


Fig. 5.12: *The ΔLikelihood for $F_{B_d}\sqrt{\hat{B}_{B_d}}$, $\hat{B}_K$ and ξ using the Bayesian and frequentist approaches.*

combination procedure of the two statistical approaches are identical.

5.2. Distributions for the relevant quantities in the CKM fits

The relevant quantities entering the fit are summarized in Table 5.1 given at the beginning of this Chapter. Figures. 5.11 and 5.12 show the Δ Log(Likelihood) for $|V_{cb}|$,$|V_{ub}|$ and for the non-perturbative QCD parameters, $F_{B_d}\sqrt{\hat{B}_{B_d}}$, ξ and $\hat{B}_K$ as obtained following the Bayesian and the frequentist methods. To be more explicit, in Table 5.2. we show the 68% and 95% ranges as obtained following the Bayesian and the *R*fit methods. It can be noticed that differences on the input quantities between the two approaches can be important and depend upon the chosen splitting of the errors. In the Bayesian approach the splitting of the total error in two errors is not really important, since, the two errors are often, already, the results of the convolution of many different source of errors. It has been also shown that the choice of the shape of the pdf to be attributed to the error has a moderate impact on the final results [16], once the central value and the r.m.s. of the pdf has been fixed. In the *R*fit this splitting is crucial and a careful breakdown of the sources of the errors which contribute to it should be done. For this comparison we

have decided to keep this splitting and to classify certain errors as a flat pdf and "non statistical" for the Bayesian and *R*fit approaches, respectively.

Parameters	68% range	95% range
$\hat{B}_K$ *R*fit (Bayes) [ratio R/B]	0.68-1.06 (0.76-0.98) [1.70]	0.62-1.12 (0.67-1.06) [1.25]

Table 5.2: *68% and 95% ranges for some relevant quantities used in the CKM fits in the* R*fit and Bayes approaches.*

5.3. Results and Comparison

*R*fit Method			
Parameter	$\leq 5\%$ CL	$\leq 1\%$ CL	$\leq 0.1\%$ CL
$\bar{\rho}$	0.091 - 0.317	0.071 - 0.351	0.042 - 0.379
$\bar{\eta}$	0.273 - 0.408	0.257 - 0.423	0.242 - 0.442
$\sin 2\beta$	0.632 - 0.813	0.594 - 0.834	0.554 - 0.855
γ°	42.1 - 75.7	38.6 - 78.7	36.0 - 83.5

Bayesian Method			
Parameter	5% CL	1% CL	0.1% CL
$\bar{\rho}$	0.137 - 0.295	0.108 - 0.317	0.045 - 0.347
$\bar{\eta}$	0.295 - 0.409	0.278 - 0.427	0.259 - 0.449
$\sin 2\beta$	0.665 - 0.820	0.637 - 0.841	0.604 - 0.863
γ°	47.0 - 70.0	44.0 - 74.4	40.0 - 83.6

Ratio *R*fit/Bayesian Method			
Parameter	5% CL	1% CL	0.1% CL
$\bar{\rho}$	1.43	1.34	1.12
$\bar{\eta}$	1.18	1.12	1.05
$\sin 2\beta$	1.17	1.18	1.16
γ°	1.46	1.31	1.09

Table 5.3: *Ranges at difference C.L for $\bar{\rho}$, $\bar{\eta}$, sin 2β and γ. The measurements of $|V_{ub}|\,/\,|V_{cb}|$ and ΔM_d, the amplitude spectrum for including the information from the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$oscillations, $|\varepsilon_K|$ and the measurement of sin 2β have been used.*

For the comparison of the results of the fit we use $\bar{\rho}$, $\bar{\eta}$, sin 2β and γ. Those quantities are compared at the 95%, 99% and 99.9% C.L. It has to be stressed that in the frequentist approach those confidence levels correspond to $\geq$95%, $\geq$99% and $\geq$99.9%. All the available constraints have been used: the measurements of $|V_{ub}|\,/\,|V_{cb}|$, ΔM_d, the amplitude spectrum for including the information from the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations, $|\varepsilon_K|$ and the measurement of sin 2β. It has to be stressed once more that the inputs used are the same in the two approaches (in term of Gaussian and uniform uncertainties), but

they correspond to different input likelihoods, for $|V_{cb}|$, $|V_{ub}|$, $F_{B_d}\sqrt{\hat{B}_{B_d}}$, $\hat{B}_K$ and ξ as shown in the previous figures. Figure 5.13 shows the comparison on the $(\bar{\rho}, \bar{\eta})$ plane. The numerical results are given in Table 5.3. Figure 5.14 shows the comparison between the allowed regions obtained using Bayesian or *R*fit methods if the constraint from the direct measurement of sin2β is removed from the fit.

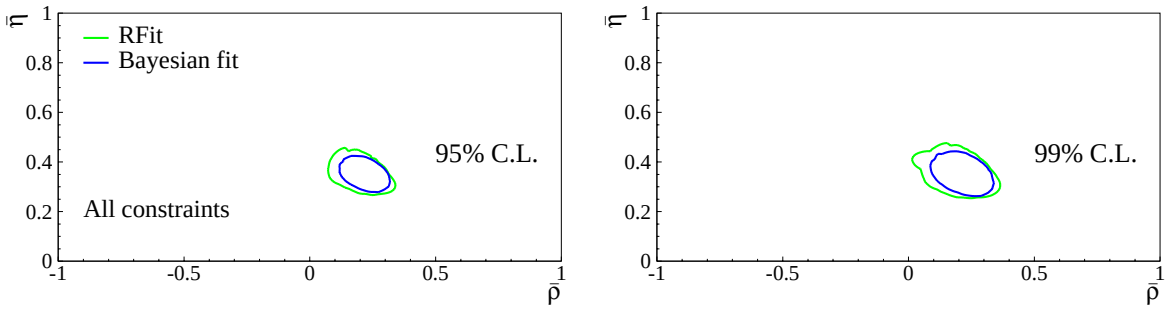


Fig. 5.13: *Comparison Bayesian/*R*fit Methods. Allowed regions for $\bar{\rho}$ and $\bar{\eta}$ at 95% (left plot) and 99% (right plot) using the measurements of $|V_{ub}|$ / $|V_{cb}|$, ΔM_d, the amplitude spectrum for including the information from the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations, $|\varepsilon_K|$ and the measurement of sin 2β.*

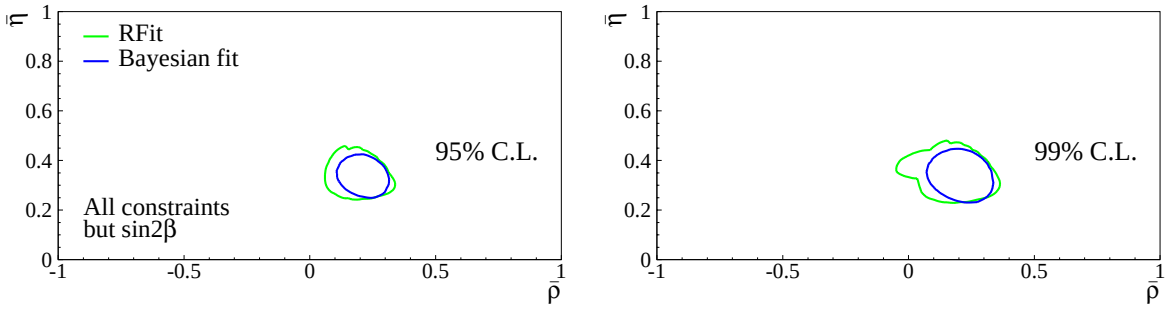


Fig. 5.14: *Comparison Bayesian/*R*fit Methods. Allowed regions for $\bar{\rho}$ and $\bar{\eta}$ at 95% (left plot) and 99% (right plot) using the measurements of $|V_{ub}|$ / $|V_{cb}|$, ΔM_d, the amplitude spectrum for including the information from the $\mathrm{B}_s^0 - \overline{\mathrm{B}}_s^0$ oscillations and $|\varepsilon_K|$*

5.3.1. Further comparisons

To further the orogin of the residual difference between the two methods, we have performed the following test: both methods use the distributions as obtained from *R*fit or from the Bayesian method to account for the information on input quantities. The results of the comparison using the input distributions as obtained from *R*fit are shown in Figs. 5.15 (Table 5.4). In some cases (0.1% C.L.) the ranges selected by the Bayesian approach are wider. The comparison using the input distributions, as obtained from the Bayesian method, give a maximal difference of 5%. These two tests show that, if same input likelihood are used, the results on the output quantities are very similar. The main origin of the residual difference on the output quantities, between the Bayesian and the *R*fit method comes from the likelihood associated to the input quantities.

5.3.2. Some conclusions on the fit comparison

The Bayesian and the *R*fit methods are compared in an agreed framework in terms of input and output quantities. For the input quantities the total error has been splitted in two errors. The splitting and the p.d.f distribution associated to any of the errors is not really important in the Bayesian approach. It becomes central in the *R*fit approach where the systematic errors are treated as "non statistical" errors.

Parameter	5% CL	1% CL	0.1% CL
$\bar{\rho}$	1.20	1.13	0.96
$\bar{\eta}$	1.03	0.99	0.94
$\sin 2\beta$	1.07	1.08	1.07
γ°	1.24	1.12	0.95

Table 5.4: *Comparison. Ratio for confidence levels* R*fit/Bayesian using the distributions as obtained from* R*fit to account for the information on input quantities*

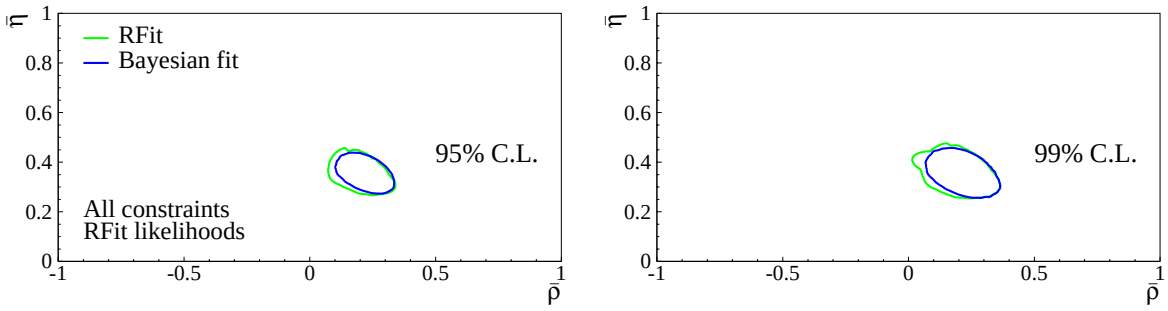


Fig. 5.15: *Comparison Bayesian/*R*fit Methods using the distributions as obtained from* R*fit to account for the information on input quantities. Allowed regions for $\bar{\rho}$ and $\bar{\eta}$ at 95% (left plot) and 99%(right plot) using the measurements of $|V_{ub}|$ / $|V_{cb}|$ and ΔM_d the amplitude spectrum for including the information from the $\mathrm{B}^0_s - \overline{\mathrm{B}}^0_s$ oscillations, $|\varepsilon_K|$ and the measurement of sin 2β.*

The result is that, even if the same central values and errors are used in the two methods, the likelihood associated to the input parameters, which are entering in the fitting procedure, can be different. The output results ($\bar{\rho}$,$\bar{\eta}$, sin2 β, γ) differ by 15%-45%, 10%-35% and 5-15% if the 95%, 99% and 99.9% confidence regions are compared, respectively, with ranges from the frequentist method being wider. If the same likelihoods are used the output results are very similar.

6. Test of the CKM picture in the Standard Model

After comparing different statistical methods, in this final Section we show how the present data can be used to test the CKM picture of the Standard Model. The results presented here have been obtained with a Bayesian fit to the latest inputs of Table 5.1. The central values, errors and 68% (95%) [and 99%] C.L. ranges obtained for various quantities of interest are collected in Table 5.5.

The most crucial test is the comparison between the UT parameters determined with quantities sensitive to the sides of the UT (semileptonic B decays and oscillations) with the measurement of CP violation in the kaon sector ($|\varepsilon_K|$) and, also with the one in the B (sin2β) sector. This test is shown in Fig. 5.16. It can be translated quantitatively into a comparison between the values of sin2β obtained from the measurement of the CP asymmetry in the $J/\psi \mathrm{K}^0_\mathrm{S}$ decays and the one determined from "sides" measurements:

$$\begin{array}{ll} \sin 2\beta = & 0.685 \pm 0.052\ (0.581, 0.789) \quad \text{indirect} - \text{sides only} \\ \sin 2\beta = & 0.734 \pm 0.054\ (0.628, 0.840) \qquad \mathrm{B}^0 \to \mathrm{J}/\psi \mathrm{K}^0_\mathrm{S}\ , \end{array}$$

where, within parentheses, we give also the 95% probability region. The spectacular agreement between these values shows the consistency of the Standard Model in describing the CP violation phenomena

$\|V_{cb}\| \times 10^3$	41.5 ± 0.8	(39.9,43.1) [39.1,43.9]
$\bar{\eta}$	0.341 ± 0.028	(0.288,0.397) [0.271,0.415]
$\bar{\rho}$	0.178 ± 0.046	(0.085,0.265) [0.052,0.297]
$\sin 2\beta$	0.705 ± 0.037	(0.636,0.779) [0.612,0.799]
$\sin 2\alpha$	-0.19 ± 0.25	(–0.62,0.33) [–0.75,0.47]
γ(degrees)	61.5 ± 7.0	(49.0,77.0) [44.3,82.1]
$\Delta M_s(ps^{-1})$	18.3 ± 1.7	(15.6,22.2) [15.1,27.0]

Table 5.5: *Values and errors for different quantities using the present knowledge summarized in Table 5.1. Within parentheses and brackets the 95% and 99% probability regions are, respectively, given.*

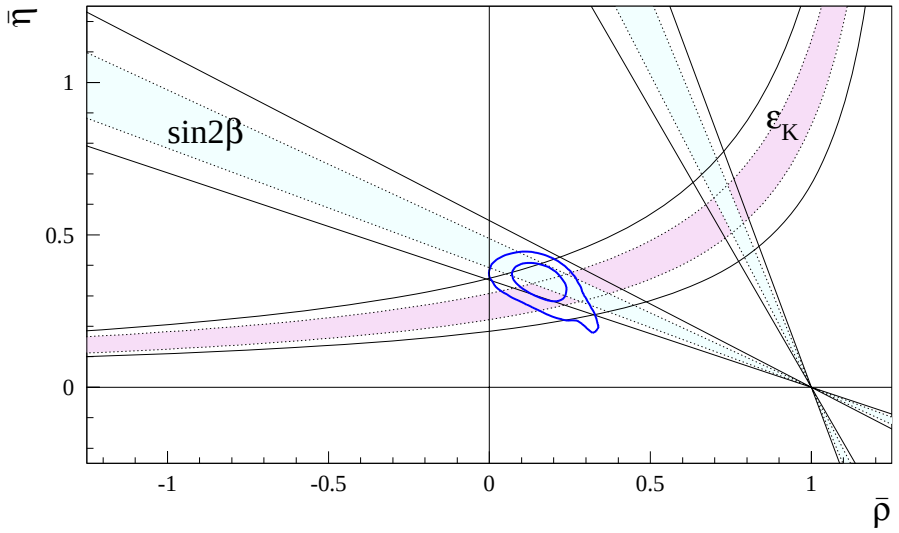


Fig. 5.16: *The allowed regions for $\overline{\rho}$ and $\overline{\eta}$ (contours at 68%, 95%) as selected by the measurements of $|V_{ub}|$ / $|V_{cb}|$, ΔM_d, and by the limit on $\Delta M_s/\Delta M_d$ are compared with the bands (at 68% and 95% C.L.) from CP violation in the kaon (ϵ_K) and in the B (sin2β) sectors.*

in terms of one single complex parameter η. Conversely, assuming the validity of the SM, this is also an important test of the OPE, HQET and LQCD theories which have been used to extract the CKM parameters. It has to be noted that the test is significant provided that the errors on sin 2β from the two determinations are comparable. Presently, the accuracy of both is at the 10% level. It is also of interest to explicitly make predictions for quantities which will be measured in the next future. We concentrate on ΔM_s which will be soon measured at Tevatron. The results obtained by excluding (or including) the information from the $B_s^0 - \bar{B}_s^0$ analyses are:

$$\begin{array}{rc} \Delta M_s(\text{with } \Delta M_s \text{ included}) = & 18.3 \pm 1.7\ (15.6, 22.2)\ [15.1, 27.0]\ \text{ps}^{-1} \\ \Delta M_s(\text{without } \Delta M_s\ ) = & 20.6 \pm 3.5\ (14.2, 28.1)\ [12.8, 30.7]\ \text{ps}^{-1}\ . \end{array}$$

where, within parentheses, we give the 95% and the 99% regions.

It will be interesting to compare these results with future measurements with the goal of identifying new physics contributions. Moreover a precise measurement of ΔM_s will reduce significantly the uncertainties in the output quantities in Table 5.5.

References

[1] A. Ali and D. London, Eur. Phys. J. C **18** (2001) 665;
S. Mele, Phys. Rev. D **59** (1999) 113011;
D. Atwood and A. Soni, Phys. Lett. B **508** (2001) 17.

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

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

[4] Y. Grossman, Y. Nir, S. Plaszczynski and M.-H. Schune, Nucl. Phys. B **511** (1998) 69;
S. Plaszczynski and M.-H. Schune, hep-ph/9911280.

[5] L. Wolfenstein, Phys. Rev. Lett. **51** (1983) 1945.

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

[7] S. Herrlich and U. Nierste, Nucl. Phys. B **419** (1994) 192, Phys. Rev. D **52** (1995) 6505 and recent update by M. Jamin and U. Nierste.

[8] Y. Nir, hep-ph/0208080.

[9] B. Aubert *et al.*, BaBar Collaboration, hep-ex/0207042.

[10] K. Abe *et al.*, Belle Collaboration, hep-ex/0208025.

[11] G. D'Agostini, CERN Report 99–03.

[12] "*CKMfitter*: code, numerical results and plots", *http://ckmfitter.in2p3.fr*

[13] S. Laplace, Z. Ligeti, Y. Nir, G. Perez, Phys. Rev. D **65**:094040 (2002)

[14] The BaBar Physics Book, BaBar Collaboration, (P.F. Harrison and H.R. Quinn, eds.). SLAC-R-0504 (1998).

[15] G.P. Dubois-Felsmann D.G. Hitlin, F.C. Porter and G. Eigen, CALT 68-2396 (2002).

[16] F. Parodi, P. Roudeau and A. Stocchi, Nuovo Cim. **112A** (1999) 833.