Eur. Phys. J. C 21, 225–259 (2001)
Digital Object Identifier (DOI) 10.1007/s100520100729

THE EUROPEAN PHYSICAL JOURNAL C

A new approach to a global fit of the CKM matrix

A. Höcker[1], H. Lacker[1], S. Laplace[1], F. Le Diberder[2]

[1] Laboratoire de l'Accélérateur Linéaire, IN2P3-CNRS et Université de Paris-Sud, BP 34, 91898 Orsay Cedex, France
[2] Laboratoire de Physique Nucléaire et des Hautes Energies, 4, place Jussieu, Tour 33 r.d.c., 75005 Paris, Cedex 05, France

Received: 18 April 2001 /
Published online: 10 August 2001 – © Springer-Verlag / Società Italiana di Fisica 2001

Abstract. We report on a new approach to a global CKM matrix analysis taking into account most recent experimental and theoretical results. The statistical framework (*R*fit) developed in this paper advocates frequentist statistics. Other approaches, such as Bayesian statistics or the 95% CL scan method are also discussed. We emphasize the distinction of a model testing and a model dependent, metrological phase in which the various parameters of the theory are estimated. Measurements and theoretical parameters entering the global fit are thoroughly discussed, in particular with respect to their theoretical uncertainties. Graphical results for confidence levels are drawn in various one and two-dimensional parameter spaces. Numerical results are provided for all relevant CKM parameterizations, the CKM elements and theoretical input parameters. Predictions for branching ratios of rare K and B meson decays are obtained. A simple, predictive SUSY extension of the Standard Model is discussed.

Contents

1 Introduction . . . 225
2 The CKM matrix . . . 226
  2.1 The standard parameterization . . . 227
  2.2 The Wolfenstein parameterization . . . 227
  2.3 Phase invariance . . . 227
  2.4 The unitarity triangle . . . 227
3 The statistical framework . . . 229
  3.1 The likelihood function . . . 229
  3.2 Metrology . . . 231
  3.3 Probing the SM . . . 233
  3.4 Probing new physics . . . 234
4 Alternative statistical treatments . . . 234
  4.1 Reminder: the *R*fit scheme . . . 234
  4.2 The 95% scan method . . . 234
  4.3 The extended conservative method (*ER*fit) . . . 235
  4.4 The Bayesian treatment . . . 235
5 Likelihoods and systematic errors . . . 236
  5.1 The hat function . . . 236
  5.2 Combining statistical and systematic uncertainties . . . 237
6 Fit ingredients . . . 237
  6.1 The CKM matrix elements . . . 237
  6.2 CP observables and mixing . . . 241
  6.3 Future prospects: rare decays of K and B-mesons . . . 243
7 Constrained fits within the SM . . . 245
  7.1 Two dimensional parameter spaces . . . 245
  7.2 One dimensional parameter spaces . . . 246
  7.3 Probing the standard model . . . 250
8 Supersymmetric extensions of the standard model . . . 252
  8.1 Supersymmetric fits . . . 252
9 Conclusions . . . 253

Appendices
A Critical Issues of the Bayesian Approach . . . 253
B Critical Issues of the 95% CL Scan Method . . . 254
C Comments on Statistics of Normal Ratios . . . 255
  C.1 A Numerical Example . . . 255
  C.2 Probability Density Functions . . . 255
  C.3 Confidence Levels . . . 256

1 Introduction

Within the Standard Model (SM), CP violation is generated by a single non-vanishing phase in the unitary Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix V [1,2]. A useful parameterization [3,4] follows from the observation that the elements of V exhibit a hierarchy in terms of the parameter $\lambda = |V_{us}|$. Other parameters of this representation are A, ρ and η, where CP violation necessarily requires $\eta \neq 0$. The parameters λ and A are obtained from measurements of semileptonic decay rates of K mesons and B meson decays involving $b \to c$ transitions, respectively. Constraints on ρ and η are obtained from measurements of semileptonic B decays yielding $|V_{ub}|$ and the ratio $|V_{ub}/V_{cb}|$. Standard Model predictions of B_d^0 and B_s^0 oscillations, and of indirect CP violation in the neutral kaon sector, depend on CKM parameters; therefore measurements of these observables provide constraints in the $\rho - \eta$ plane, albeit being limited by theoretical uncertainties coming mainly from long distance QCD effects. Finally, in the era of the B-factories, it will be possible, for the first time, to assess the CP-violating angles α, β and γ of the *Unitarity Triangle* (UT) express-

ing the unitarity relation between the first and the third column of V.

The first goal of a global CKM fit is to *probe the validity of the SM*, that is to quantify the agreement between the SM and the experimental information. Furthermore, one intends to perform a detailed *Metrology*, that is to find allowed ranges for CKM matrix elements and related quantities, assuming the SM to be correct. Finally, within an extended theoretical framework, one may search for specific signals of *new physics*, by estimating the additional theoretical parameters.

Analyzing data in a well defined theoretical framework ceases to be a straightforward task when one moves away from Gaussian statistics. This is the case for the theoretically limited precision on the SM predictions of the neutral K and B mixing observables and, to a lesser extent, for the semileptonic decay rates of B decays to charmed and charmless final states. The statistical approach (Rfit) developed in this analysis allows a non-Bayesian treatment of the, *a priori* unknown (*i.e.*, not statistically distributed), theoretical parameters and theoretical systematics of measurements. The ensemble of the statistical analysis is realized in the program package *CkmFitter*[1]. A detailed description of the methods it uses, with emphasis on the new method denoted Rfit which is proposed here, and the presentation of state-of-the-art results are the subject of this paper[2].

The paper is organized as follows: after recalling the most common CKM parameterizations, we comprehensively discuss the statistical framework of the analysis, starting with the introduction of the relevant likelihoods in Sect. 3.1, followed by a definition of the three analysis steps: metrology, model testing and probing for new physics. We then recall the principles of the 95% CL Scan scheme [5] and of the Bayesian approach [6,7] in Sect. 4 (see also [8] and references therein for a tentative collection of publications on the CKM matrix and related topics). We work out their limitations and motivate further going ideas, while never leaving non-Bayesian grounds. This is followed by a discussion of the treatment of experimental and theoretical systematics in Sect. 5. In Sect. 6 we present a compendium of the input measurements, their predictions in the framework of the SM, and discuss the theoretical parameters and their uncertainties used in the analysis. We display our fit results as confidence levels in various parameter spaces in Sect. 7 and produce tables of constraints on all relevant CKM parameters, constraining measurements and theoretical inputs, and predictions of rare K and B decays. Within our statistical approach, we perform a test of the goodness of the theory and discuss the effect of a simple, predictive Minimal Supersymmetric extension of the SM. Deepening statistical discussions on some crucial issues of the analysis are given in the appendix.

2 The CKM matrix

Invariance under local gauge transformation prevents the bare masses of the leptons and quarks to appear in the $SU(3) \times SU(2) \times U(1)$ Lagrange density of the SM. Instead, the spontaneous breakdown of electroweak symmetry dynamically generates masses for the fermions due to the Yukawa coupling of the fermion fields to the Higgs doublet. Since the latter has a non-vanishing vacuum expectation value, the Yukawa couplings g give rise to the 3×3 mass matrices

$$M_i = \frac{v g_i}{\sqrt{2}} \ , \tag{1}$$

with $i = u(d)$ for up(down)-type quarks and $i = e$ for the massive leptons. The transformation of the M_i from the basis of the flavor eigenstates to the basis of the mass eigenstates is realized by unitary rotation matrices U_i, where

$$U_{u(d,e)} M_{u(d,e)} U^{\dagger}_{u(d,e)} = \mathrm{diag}\left(m_{u(d,e)}, m_{c(s,\mu)}, m_{t(b,\tau)}\right) \ . \tag{2}$$

For the Lagrange density in the basis of the mass-eigenstates the neutral-current part remains unchanged (*i.e.*, there are no flavor changing neutral currents present at tree level), whereas the charged current part of the quark sector is modified by the product of the up-type and down-type quark mass matrices,

$$V = U_u U_d^{\dagger} \ , \tag{3}$$

which is the CKM mixing matrix. By convention, V operates on the $-1/3$ charged down-type quark mass eigenstates

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{4}$$

and, being the product of unitary matrices, V itself is unitary:

$$V V^{\dagger} = \mathrm{Id} \ . \tag{5}$$

There exists a hierarchy between the elements of V both for their value (the diagonal elements dominate) and their errors (since they dominate, they are better known). Unitarity and the phase arbitrariness of fields reduce the initially nine complex parameters of V to three real numbers and one phase, where the latter accounts for CP violation. It is therefore interesting to over-constrain V since deviations from unitarity would reveal the existence of new generation(s) or new couplings.

The charged current couplings among left-handed quark fields are proportional to the elements of V. For right-handed quarks there exist no W boson interaction

[1] *CkmFitter* is a framework package that hosts several statistical approaches to a global CKM fit, such as Rfit, Bayesian techniques and the 95% CL Scan method. It is available as public share ware. Please, contact the authors for more information

[2] Visit the *CkmFitter* web page to find plots, reference links, detailed descriptions and more:
http://www.slac.stanford.edu/∼laplace/ckmfitter.html

$$V \simeq \begin{pmatrix} 1-\frac{1}{2}\lambda^2-\frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda\left[1+A^2\lambda^4\left(\rho+i\eta-\frac{1}{2}\right)\right] & 1-\frac{1}{2}\lambda^2-\frac{1}{8}\lambda^4(1+4A^2) & A\lambda^2 \\ A\lambda^3\left[1-(\rho+i\eta)\left(1-\frac{1}{2}\lambda^2\right)\right] & -A\lambda^2\left[1+\lambda^2\left(\rho+i\eta-\frac{1}{2}\right)\right] & 1-\frac{1}{2}A^2\lambda^4 \end{pmatrix} \tag{7}$$

in the SM and the Z, photon and gluon couplings are flavor diagonal. For left-handed leptons the analysis proceeds similar to the quarks.

There are many ways of parameterizing the CKM matrix in terms of four parameters. It is the purpose of this section to summarize the most popular representations.

2.1 The standard parameterization

The Standard Parameterization of V is taken to be the one proposed by Chau and Keung [9], and advocated by the PDG [10]. It is obtained by the product of three complex rotation matrices, where the rotations are characterized by the Euler angles θ_{12}, θ_{13} and θ_{23}, which are the mixing angles between the generations, and an overall phase δ:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23}-c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23}-s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23}-c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23}-s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{6}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ for $i < j = 1,2,3$. This parameterization has the considerable advantage of being exact in the sense that it strictly satisfies the unitarity relation (5).

2.2 The Wolfenstein parameterization

Following the observation of a hierarchy between the different matrix elements, Wolfenstein [3] proposed a simple expansion of the CKM matrix in terms of the four parameters λ, A, ρ and η ($\lambda = |V_{us}| \sim 0.22$ being the expansion parameter), which is widely used in contemporary literature. Using the convention of [4] one has (see (7) on top of the page) It is obtained from (6) *via* the definitions

$$\begin{aligned} s_{12} &= \lambda \ , \\ s_{23} &= A\lambda^2 \ , \\ s_{13}e^{-i\delta} &= A\lambda^3(\rho - i\eta) \ , \end{aligned} \tag{8}$$

and is valid to the order $O(|\lambda|^6) \simeq 0.01\%$.

2.3 Phase invariance

It was shown by Jarlskog [11] that the determinant of the commutator of the up-type and down-type unitary mass matrices (1) reads

$$\det[M_u, M_d] = -2iF_uF_dJ \ , \tag{9}$$

with F_u, F_d, being

$$\begin{aligned} F_{u(d)} = &(m_{t(b)} - m_{c(s)})(m_{t(b)} - m_{u(d)}) \\ &\times(m_{c(s)} - m_{u(d)})/m^3_{t(b)} \ . \end{aligned} \tag{10}$$

The phase-convention independent measure of CP violation, J, is given by

$$\mathrm{Im}\left[V_{ij}V_{kl}V^*_{il}V^*_{kj}\right] = J\sum_{m,n=1}^{3}\epsilon_{ikm}\epsilon_{jln} \ , \tag{11}$$

with the CKM matrix elements V_{ij} and ϵ_{ikm} being the total antisymmetric tensor. One representation of (11) reads, for instance, $J = \mathrm{Im}[V_{ud}V_{cs}V^*_{us}V^*_{cd}]$. A non-vanishing CKM phase and hence CP violation necessarily requires $J \neq 0$. The Jarlskog parameter expressed in the Standard Parameterization (6) reads

$$J = c_{12}c_{23}c^2_{13}s_{12}s_{23}s_{13}\sin\delta \ , \tag{12}$$

and, using the Wolfenstein approximation (7), valid to the order $O(|\lambda|^{10})$, one finds

$$J = A^2\lambda^6\eta \ \sim 10^{-5} \ . \tag{13}$$

The empirical value of J is small compared to its maximum of $1/(6\sqrt{3}) \simeq 0.1$ showing that CP violation is suppressed as a consequence of the strong hierarchy exhibited by the CKM matrix elements. It is the remarkable outcome of (9) that CP violation requires not only J to be non-zero, but also the existence of a non-degenerated mass hierarchy. Equal masses between at least two generations of up-type or down-type quarks would necessarily remove the CKM phase.

Phase convention invariance of the V-transformed quark wave functions is a requirement for physically meaningful quantities. Such invariants are the moduli $|V_{ij}|^2$ and the quadri-products $V_{ij}V_{kl}V^*_{il}V^*_{kj}$ (*c.f.*, the Jarlskog invariant J (11)). Non-trivial higher invariants can be reformulated as functions of moduli and quadri-products (see, *e.g.*, [13]). Indeed, (11) expresses the fact that, owing to the orthogonality of any pair of different rows or columns of V, the imaginary parts of all quadri-products are equal up to their sign. We will use phase-invariant representations and formulae throughout this paper[3].

2.4 The unitarity triangle

The allowed region in $\rho - \eta$ space can be elegantly displayed using the unitarity triangle (UT) described by the *rescaled* unitarity relation between the first and the third column of the CKM matrix

$$\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}} + \frac{V_{cd}V^*_{cb}}{V_{cd}V^*_{cb}} + \frac{V_{td}V^*_{tb}}{V_{cd}V^*_{cb}} = 0 \ . \tag{14}$$

Note that twice the area of the *non-rescaled* UT corre-

[3] We are indebted to K. Schubert for drawing our attention to this point

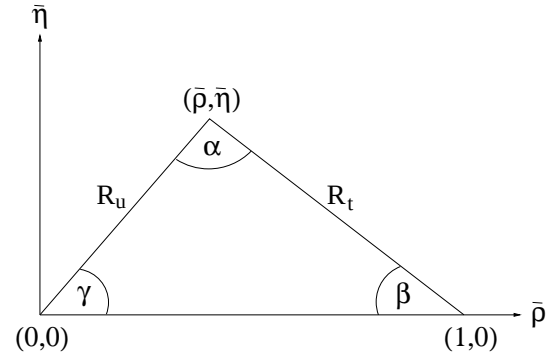

Fig. 1. The rescaled Unitarity Triangle in the Wolfenstein parameterization

sponds to the Jarlskog parameter J. This identity provides a geometrical interpretation of the phase convention invariance of J: a rotation of the CKM matrix rotates the UT accordingly while leaving its area, and thus J, invariant. It is the remarkable property of the UT (14) that its three sides are governed by the same power of λ and A

$$\frac{A\lambda^3}{A\lambda^3} + 1 + \frac{A\lambda^3}{A\lambda^3} \simeq 0 \ , \tag{15}$$

which predicts large CP asymmetries in the B sector [12]. As a comparison, the corresponding UT for the kaon sector is almost flat

$$0 = \frac{V_{ud}V_{us}^*}{V_{cd}V_{cs}^*} + \frac{V_{cd}V_{cs}^*}{V_{cd}V_{cs}^*} + \frac{V_{td}V_{ts}^*}{V_{cd}V_{cs}^*} \simeq \frac{\lambda}{\lambda} + 1 + \frac{A^2\lambda^5}{\lambda} \ , \tag{16}$$

exhibiting small CP asymmetries. The UT (14) is sketched in Fig. 1 in the complex $(\bar{\rho}, \bar{\eta})$ plane ($\bar{\rho} = \rho(1-\lambda^2/2)$, $\bar{\eta} = \eta(1-\lambda^2/2)$) of the Wolfenstein parameterization[4]. The sides R_u and R_t (the third side being normalized to unity) are given by

$$R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\bar{\rho}^2 + \bar{\eta}^2} \ , \tag{17}$$

$$R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\bar{\rho})^2 + \bar{\eta}^2} \ . \tag{18}$$

and the three angles, α, β, γ, read

$$\alpha = \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right] \ , \qquad \beta = \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right] \ ,$$
$$\gamma = \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] \ , \tag{19}$$

where $\gamma \equiv \delta$ in the Standard Parameterization. The angles and sides of the UT obey the trigonometric relation, $\sin\alpha : \sin\beta : \sin\gamma = 1 : R_u : R_t$.

[4] The length of the vector to the triangle apex is given by $\left|1 + \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\rho^2+\eta^2}\left(1 - \frac{\lambda^2}{2}\right) + O(|\lambda|^4)$, so that the replacements $\rho \to \rho(1-\lambda^2/2)$ and $\eta \to \eta(1-\lambda^2/2)$, where $V_{ud} = 1 - \lambda^2/2 - O(\lambda^4)$, improves the precision of the apex coordinate in the Wolfenstein approximation [4]

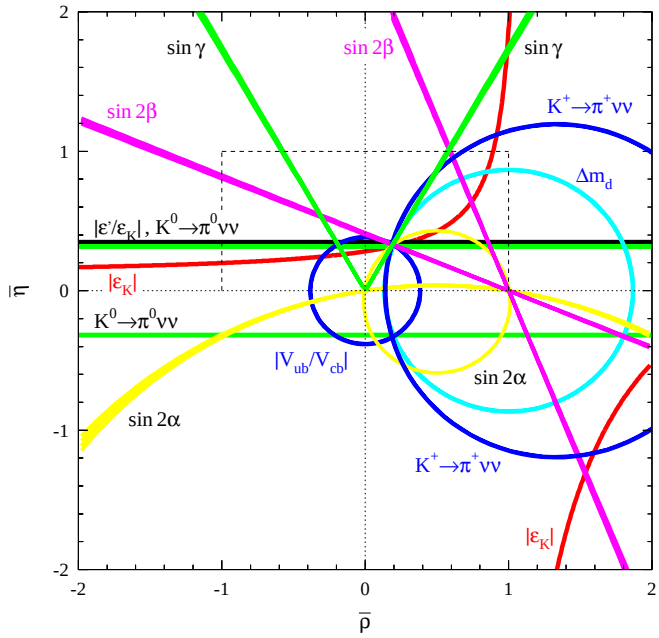

Fig. 2. Constraints in the $(\bar{\rho}, \bar{\eta})$ plane for the most relevant observables. The theoretical parameters used correspond to some "standard" set chosen to reproduce compatibility. The dashed lines indicate the rectangle on which we concentrate in the following for the global fit

The relations between the angles and the $\bar{\rho}$, $\bar{\eta}$ coordinates are given by

$$\sin 2\alpha = \frac{2\bar{\eta}(\bar{\eta}^2 - \bar{\rho}(1-\bar{\rho}))}{(\bar{\eta}^2 + (1-\bar{\rho})^2)(\bar{\eta}^2 + \bar{\rho}^2)} \ , \tag{20}$$

$$\sin 2\beta = \frac{2\bar{\eta}(1-\bar{\rho})}{\bar{\eta}^2 + (1-\bar{\rho})^2} \ , \tag{21}$$

$$\tan\gamma = \frac{\bar{\eta}}{\bar{\rho}} \ . \tag{22}$$

A graphical compilation of the most relevant present and future constraints sensitive to the CP violating phase δ is displayed in Fig. 2. We simplify the representation by assuming a measurement of $\sin 2\alpha$ whereas, in principle, the UT angle α can be directly determined from $B \to 3\pi$ decays. For the third UT angle γ, we assume a measurement of $\sin\gamma$, even though charmless B decays may allow a non-ambiguous determination of γ. A more detailed elaboration of future measurements is given in [14]. Some "standard" set of theoretical parameters is used for this exercise in order to reproduce compatibility between the constraints. The present experimental values for the observables and their dependence on the CKM matrix elements in the framework of the SM are discussed in Sect. 6.

Over-constraining the unitary CKM matrix aims at validating or not the SM with three generations. The interpretation of these constraints requires a robust statistical framework which protects against misleading conclusions. The next section describes to some detail the statistical approach applied for the analysis reported in this work.

3 The statistical framework

We are considering an analysis involving a set of $N_{\rm exp}$ measurements collectively denoted by $x_{\rm exp} = \{x_{\rm exp}(1), \ldots, x_{\rm exp}(N_{\rm exp})\}$, described by a set of corresponding theoretical expressions $x_{\rm theo} = \{x_{\rm theo}(1), \ldots, x_{\rm theo}(N_{\rm exp})\}$. The theoretical expressions $x_{\rm theo}$ are functions of a set of $N_{\rm mod}$ parameters $y_{\rm mod} = \{y_{\rm mod}(1), \ldots, y_{\rm mod}(N_{\rm mod})\}$. Their precise definition is irrelevant for the present discussion (*c.f.*, Sect. 6 for details) beside the fact that:

- a subset of $N_{\rm theo}$ parameters within the $y_{\rm mod}$ set are fundamental, and free, parameters of the theory (*i.e.*, the four CKM unknowns in the SM, the top quark mass, etc.) and are denoted $y_{\rm theo}$, where $y_{\rm theo} = \{y_{\rm theo}(1), \ldots, y_{\rm theo}(N_{\rm theo})\}$.
- the remaining $N_{\rm QCD} = N_{\rm mod} - N_{\rm theo}$ parameters, which appear due to our present inability to compute precisely strong interaction quantities (*e.g.*, f_{B_d}, B_d, etc.), are denoted $y_{\rm QCD}$, where $y_{\rm QCD} = \{y_{\rm QCD}(1), \ldots, y_{\rm QCD}(N_{\rm QCD})\}$.

There are three quite different goals the analysis aims at:

1. Within the SM, to achieve the best estimate of the $y_{\rm theo}$ parameters: that is to say to perform a careful metrology of the theoretical parameters, for later use.
2. Within the SM, to set a confidence level (CL) which quantifies the agreement between data and the theory, as a whole.
3. Within an extended theoretical framework, *e.g.*, Supersymmetry, to search for specific signs of new physics by pinning down additional fundamental, and free, parameters of the theory.

These three goals imply three statistical treatments all of which rely on a likelihood function meant to gauge the agreement between data and theory.

3.1 The likelihood function

We adopt a χ^2-like notation and denote

$$\chi^2(y_{\rm mod}) \equiv -2\ln(\mathcal{L}(y_{\rm mod}))\ , \tag{23}$$

where $\mathcal{L}$, the likelihood function (it is defined below) receives contributions of two types

$$\mathcal{L}(y_{\rm mod}) = \mathcal{L}_{\rm exp}(x_{\rm exp} - x_{\rm theo}(y_{\rm mod}))\ \mathcal{L}_{\rm theo}(y_{\rm QCD})\ . \tag{24}$$

The first term, the experimental likelihood $\mathcal{L}_{\rm exp}$, measures the agreement between $x_{\rm exp}$ and $x_{\rm theo}$, while the second term, the theoretical likelihood $\mathcal{L}_{\rm theo}$, expresses our present knowledge on the $y_{\rm QCD}$ parameters.

It has to be recognized from the outset that the notation χ^2 of (23) is a commodity which can be misleading. In general, denoting "Prob" the well known routine from the CERN library, one *cannot* infer a CL from the above χ^2 value using

$$\mathrm{CL} = \mathrm{Prob}(\chi^2(y_{\rm mod}), N_{\rm dof})\ , \tag{25}$$

$$= \frac{1}{\sqrt{2^{N_{\rm dof}}}\Gamma(N_{\rm dof}/2)} \int_{\chi^2(y_{\rm mod})}^{\infty} e^{-t/2} t^{N_{\rm dof}/2-1}\, dt\ . \tag{26}$$

This is because neither $\mathcal{L}_{\rm exp}$ nor $\mathcal{L}_{\rm theo}$ (they are further discussed in the sections below) are built from purely Gaussian measurements:

- In most cases $\mathcal{L}_{\rm exp}$ should handle experimental systematics, and, in some instance, it has to account for inconsistent measurements.
- In practice, $\mathcal{L}_{\rm theo}$ relies on hard-to-quantify educated guessworks, akin to the ones used to define experimental systematics, but in most cases even less well-defined.

The first limitation is not specific to the present analysis and is not the main source of concern, here. The second limitation is the most challenging: its impact on the analysis is particularly strong with the data presently available. The statistical treatment advocated below, denoted *R*fit, is designed to cope with both of the above limitations. Notwithstanding its attractive features, the *R*fit scheme does not offer a treatment of the problem at hand free from any assumption: an ill-defined problem cannot be dealt with rigorously. However, the *R*fit scheme extracts the most out of simple and clear-cut *a priori* assumptions.

The alternative statistical treatments discussed in Sect. 4 differ from the *R*fit scheme by the procedure used to define the CL from the above χ^2, or by the content and interpretation of $\mathcal{L}_{\rm theo}$.

3.1.1 The experimental likelihood

The experimental component of the likelihood is given by the product

$$\mathcal{L}_{\rm exp}(x_{\rm exp} - x_{\rm theo}(y_{\rm mod})) = \prod_{i=1}^{N_{\rm exp}} \mathcal{L}_{\rm exp}(i)\ , \tag{27}$$

where the individual likelihood components account for independent measurements[5].

The likelihood components: Ideally, the individual likelihood components $\mathcal{L}_{\rm exp}(i)$ would be pure Gaussians

$$\mathcal{L}_{\rm exp}(i) = \frac{1}{\sqrt{2\pi}\sigma_{\rm exp}(i)} \exp\left[-\frac{1}{2}\left(\frac{x_{\rm exp}(i) - x_{\rm theo}(i)}{\sigma_{\rm exp}(i)}\right)^2\right]\ , \tag{28}$$

each with a standard deviation given by the experimental statistical uncertainty $\sigma_{\rm exp}(i)$ of the $i^{\rm th}$ measurement. However, in practice, one has to deal with additional experimental and theoretical systematic uncertainties.

Experimental systematics: An experimental systematics is assumed to take the form of a possible biasing offset, the measurement could be corrected for, were it be known. Their precise treatment is discussed in Sect. 5. There, a natural extension of the usual method of adding linearly or in quadrature statistical and systematic uncertainties is proposed.

[5] Features marked by $^{(*)}$ in the following item list are not issued in the analysis presented in this work, but may become important in future CKM profiles

Theoretical systematics: Theoretical systematics, when they imply small effects, are treated as the experimental ones. However, because most theoretical systematics imply large effects and affect in a non-linear way the $x_{\rm theo}$ prediction, most of them are dealt with through the theoretical likelihood component $\mathcal{L}_{\rm theo}$ (*c.f.*, Sect. 3.1.2).

Model dependent measurements$^{(*)}$: When theoretical systematics cannot be expressed as a dependence on not perfectly known parameters, but are expressed as a set of measurements based on alternative models, labelled by the index m, (*e.g.*, the exclusive $|V_{ub}/V_{cb}|$ measurement exhibits such a model dependence) $\mathcal{L}_{\rm exp}(i)$ is defined by

$$\mathcal{L}_{\rm exp}(i) = \mathcal{L}_{\rm exp}(i,m) \ , \tag{29}$$

and m is treated as an additional $y_{\rm QCD}$ parameter, taking only discrete values.

Identical observables and consistency: When several measurements refer to the same observable (*e.g.*, various measurements of Δm_d) they have to be consistent, independently of the theoretical framework used for the analysis.

Similarly, when several measurements refer to different observables which are linked to the same $y_{\rm theo}$ parameter, *e.g.*., $|V_{ud}|$ and $|V_{us}|$, or determinations of $|V_{ub}|$ stemming from different observables, or measurements of $\sin 2\beta$ obtained from similar B decays, one may *decide* to overrule possible disagreement by imposing the measurements to be consistent. By doing so, one is deliberately blinding oneself from possible new physics effects which may have revealed themselves otherwise. Clearly, such overruling should be applied with great caution, and it should be well advertized whenever it occurs.

The method to deal with this *imposed* consistency is to account for the measurements at once, by merging them into a single component, usually obtained from their weighted mean. A more refined treatment is needed when this set of measurements is clearly inconsistent. A general method to handle a set of measurements, whether they are consistent or not, is proposed in [15,16]. Yet, for "not too large" inconsistencies, the proposed method yields similar results as the χ^2 rescaling approach adopted by the PDG [10]. To clarify the presentation of the *R*fit scheme, we use the latter: when N inconsistent measurements appear in this analysis, the error obtained for their weighted mean is rescaled by the factor $\sqrt{\chi^2_{\rm wm}/(N-1)}$, where $\chi^2_{\rm wm}$ is the weighted mean χ^2.

Related observables and consistency$^{(*)}$: In some instances, several observables, although not identical, are functionally related in a way independent of the theoretical framework used for the analysis. The number of such instances is denoted $N_{\rm cst}$, and the effective number of measurements, the one to be used to compute degrees of freedom, is defined by

$$N^{\rm eff}_{\rm exp} = N_{\rm exp} - N_{\rm cst} \ . \tag{30}$$

An example is provided by the set of measurements yielding separately $|V_{ub}|$, $|V_{cb}|$ and $|V_{ub}/V_{cb}|$. The ratio of the first two should be compatible with the third measurement, whether or not the SM is valid. Since the measurements are not referring to a unique observable they cannot be merged simply into a single component, as above. One should normalize their contribution to ensure that they do not contribute to the overall χ^2 value, if they are in the best possible mutual agreement, independently of the theoretical framework used for the analysis. This normalization is in fact what is done in the case of identical observables. It is irrelevant for the metrological phase of the analysis, and for the third phase, where one searches for specific sign of new physics: then, any constant can be added to the χ^2 without affecting the result. However, it is relevant for the second phase, where one probes the SM: a statistical fluctuation in the set $|V_{ub}|$, $|V_{cb}|$ and $|V_{ub}/V_{cb}|$ which makes them violate their functional relation should not trigger a claim for new physics. In this example, the normalization constant is obtained as the maximal value of the function of the two variables $|V^{\rm theo}_{ub}|$, $|V^{\rm theo}_{cb}|$

$$\begin{aligned}\mathcal{L}_{\rm exp}(|V^{\rm theo}_{ub}|,|V^{\rm theo}_{cb}|) = {}& \mathcal{L}_{\rm exp}\left(|V_{ub}| - |V^{\rm theo}_{ub}|\right) \\ & \times\mathcal{L}_{\rm exp}\left(|V_{cb}| - |V^{\rm theo}_{cb}|\right) \\ & \times\mathcal{L}_{\rm exp}\left(\left|\frac{V_{ub}}{V_{cb}}\right| - \frac{|V^{\rm theo}_{ub}|}{|V^{\rm theo}_{cb}|}\right) \ .\end{aligned} \tag{31}$$

Here as well, care should be taken not to normalize that way the contributions of observables the functional connection of which is model dependent. For instance, the measurements leading to $\sin\gamma$ and $\pi-\beta-\alpha$ are sensitive to new physics because their measured values may violate the SM functional relation $\sin\gamma = \sin(\pi-\beta-\alpha)$ [6].

Normalization: More generally, the normalization of each individual likelihood component is chosen such that its maximal value is equal to one. This is not important for the analysis, but it is convenient: it ensures that a measurement does not contribute numerically to the overall χ^2 value if it is in the best possible agreement with theory, and that the (so-called) χ^2 takes only positive values. In the pure Gaussian case, it implies simply to drop the normalization constant of (28): one is thus recovering the standard χ^2 definition.

3.1.2 The theoretical likelihood

The theoretical component of the likelihood is given by the product

$$\mathcal{L}_{\rm theo}(y_{\rm QCD}) = \prod_{i=1}^{N_{\rm QCD}} \mathcal{L}_{\rm theo}(i) \ , \tag{32}$$

where the individual likelihood components $\mathcal{L}_{\rm theo}(i)$ account for the partial knowledge available on the $y_{\rm QCD}$ parameters (*e.g.*, f_{B_d}) including more or less accurately

[6] It is worth pointing out that an apparent functional violation is present (since long ago) in the available data (*c.f.*. Sect. 6.1): $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 < 1$.

known correlations between them (*e.g.*, f_{B_d}/f_{B_s}). Ideally, one should incorporate in $\mathcal{L}_{\rm exp}$ measurements from which constraints on $y_{\rm QCD}$ parameters can be derived. By doing so, one could remove altogether the theoretical component of the likelihood. However, this is not what is done, because usually there is no such measurement: the *a priori* knowledge on the $y_{\rm QCD}$ stems rather from educated guesswork[7]. As a result, the $\mathcal{L}_{\rm theo}(i)$ components are incorporated by hand in (32) and they can hardly be considered as issued from probability distribution functions (PDF). In effect, their mere presence in the discussion is a clear sign that the problem at hand is ill-defined. It demonstrates that a (here critical) piece of information is coming neither from experimental, nor from statistically limited computations, but from somewhere else: from the mind of physicists. At present, these components play a leading role in the analysis and it is mandatory to handle them with the greatest caution.

The default scheme – range fit (*R*fit): In the scheme we propose, the theoretical likelihoods $\mathcal{L}_{\rm theo}(i)$ do not contribute to the χ^2 of the fit while the corresponding $y_{\rm QCD}$ parameters take values within ranges, thereafter termed "allowed ranges" and denoted $[y_{\rm QCD}]$. The numerical derivation of these ranges is discussed in Sects. 5 and 6. Most of them are identified to the ranges $[y_{\rm QCD} - \sigma_{\rm syst}\ , y_{\rm QCD} + \sigma_{\rm syst}]$, where $\sigma_{\rm syst}$ is the theoretical systematics evaluated for $y_{\rm QCD}$. Hence $y_{\rm QCD}$ values are treated on an equal footing, irrespective of how close they are from the edges of the allowed range. Instances where even only one of the $y_{\rm QCD}$ trespasses its range are not considered[8].

This is the unique, simple, and clear-cut assumption made in the *R*fit scheme: $y_{\rm QCD}$ parameters are bound to remain within *predefined* allowed ranges. The *R*fit scheme departs from a perfect frequentist analysis only because the allowed ranges $[y_{\rm QCD}]$ do not extend to the whole physical space where the parameters could *a priori* take their values[9]. This should *not* be understood as implying that a uniform PDF is ascribed to each $y_{\rm QCD}$ parameter. This important remark is further discussed in Sect. 4.4 and in Appendix A.

This unique and minimal assumption, is nevertheless a strong assumption: all the results obtained should be understood as valid only if all the assumed allowed ranges contain the true values of their $y_{\rm QCD}$ parameters. But there is no guarantee that this is the case, and this restriction should be kept in mind. On the other hand, also the contrary is true: if the ranges are chosen too big, one may miss a discovery.

3.2 Metrology

For metrology, one is not interested in the quality of the agreement between data and the theory as a whole. Rather, taking for granted that the theory as a whole is correct, one is only interested in the quality of the agreement between data and various realizations of the theory, specified by distinct sets of $y_{\rm mod}$ values. More precisely, as discussed in Sect. 3.2.1, the realizations of the theory one considers are under-specified by various subsets of so-called relevant parameters values. In the following we denote

$$\chi^2_{{\rm min};y_{\rm mod}}\ , \tag{33}$$

the absolute minimal value of the χ^2 function of (23), obtained when letting all $N_{\rm mod}$ parameters free to vary.

In principle, this absolute minimal value does not correspond to a unique $y_{\rm mod}$ location. This is because measurements (resp. theoretical predictions) entering in the analysis are all affected by more or less important experimental (resp. theoretical) systematics. These systematics being handled by means of allowed ranges, there is always a multi-dimensional degeneracy for any value of χ^2.

In practice, with the presently available observables, theoretical systematics play a prominent role. If one does not incorporate significant $\sin 2\beta$ measurements in the analysis, the domain where $\chi^2 = \chi^2_{{\rm min};y_{\rm mod}}$ is noticeably wide. For convenience, in the following we refer to this domain as $y^{\rm opt}_{\rm mod}$, as if a unique point in the $y_{\rm mod}$ space were leading to $\chi^2_{{\rm min};y_{\rm mod}}$. The projections of the $y^{\rm opt}_{\rm mod}$ domain onto one dimensional spaces result in finite intervals within which the data analysis cannot make distinction (and similarly for two-dimensional spaces). When $\sin 2\beta$ is incorporated in the analysis, one adds a measurement with negligible systematics which lifts partially the degeneracy and makes some projections of $y^{\rm opt}_{\rm mod}$ become point-like.

This degeneracy should be treated carefully when one is exploring a sub-space a of $y_{\rm mod}$: points widely apart in a can lead to the same χ^2, provided the other parameter values are changed accordingly. However, except for numerical accidents, identical $\chi^2(y_{\rm mod}) = \chi^2_{{\rm min};y_{\rm mod}}$ values imply identical $\mathcal{L}_{\rm exp}$ components, and hence identical predictions: $x_{\rm theo}(y_{\rm mod})$ values are constant within $y^{\rm opt}_{\rm mod}$.

Ideally, for metrological purposes, one should attempt to estimate as best as possible the complete $y_{\rm mod}$ set. In that case, one should use the offset-corrected χ^2

$$\Delta\chi^2(y_{\rm mod}) = \chi^2(y_{\rm mod}) - \chi^2_{{\rm min};y_{\rm mod}}\ , \tag{34}$$

where $\chi^2(y_{\rm mod})$ is the χ^2 for a given set of model parameters $y_{\rm mod}$. The minimal value of $\Delta\chi^2(y_{\rm mod})$ is zero, by

[7] The same remark applies to experimental systematics, but, since these are usually not the dominant part of the experimental uncertainties, the problem is less acute

[8] In the case of model dependence (*c.f.*, Sect. 3.1.1) the allowed values for the discrete parameter m labelling the models correspond to the set of models deemed acceptable. *R*fit is allowed to select at will any one within this set, in the same way that it is allowed to select a $y_{\rm QCD}$ parameter at will within its allowed range. In practice, when $y_{\rm QCD}$ parameters cannot be handled beforehand as explained in Sect. 5, the actuation of the allowed range in *CkmFitter* is obtained using the `Set Limit` option of the *MINUIT* package. It is equivalent to set the component $\mathcal{L}_{\rm theo}(i)$ to unity when the corresponding $y_{\rm QCD}(i)$ parameter is within $[y_{\rm QCD}(i)]$, and to zero otherwise

[9] Not all $y_{\rm QCD}$ parameters need to be given an *a priori* allowed range: *e.g.*, values taken by final state strong interaction phases (FSI) appearing in B decays are not necessarily theoretically constrained

construction. This ensures that, to be consistent with the assumption that the SM is correct, CLs equal to unity are obtained when exploring the $y_{\rm mod}$ space (namely, once $y_{\rm mod}$ enters the $y_{\rm mod}^{\rm opt}$ domain). In a Gaussian situation, one would then directly obtain the CL for a particular set of $y_{\rm mod}$ values as

$$\mathcal{P}(y_{\rm mod}) = {\rm Prob}(\Delta\chi^2(y_{\rm mod}), N_{\rm dof}) \ , \tag{35}$$

with $N_{\rm dof} = {\rm Min}(N_{\rm exp}^{\rm eff}, N_{\rm mod})$, where $N_{\rm exp}^{\rm eff}$ is defined in (30).

3.2.1 Relevant and less relevant parameters

However, one is not necessarily interested in all the $y_{\rm mod}$ values, but only in a subset of them. This can be for two distinct reasons:

- The other parameters being deemed less relevant. For instance, in the SM, CP violation can be summarized by the value taken by the Jarlskog parameter J, or by the value taken by η (in the Wolfenstein parameterization): the other CKM parameters and the $y_{\rm QCD}$ parameters may thus conceivably be considered of lower interest. More generally, one is rarely considering a CKM fit as a means to pin down anything else than CKM parameters, least of all $y_{\rm QCD}$ parameters[10].
- Parameters that cannot be significantly constrained by the input data of the CKM fit. This is the case for most of the non-CKM parameters: $y_{\rm QCD}$ parameters, but also the top quark mass, etc.

In practice, the $y_{\rm mod}$ parameters usually retained as relevant for the discussion are $\bar\rho$ and $\bar\eta$. The other parameters λ, A, the top quark mass (etc.) and all the $y_{\rm QCD}$ are considered as subsidiary parameters, merely to be taken into account in the analysis, but irrelevant for the discussion. In that case, the aim of the metrological stage of the analysis is to set CLs in the $(\bar\rho, \bar\eta)$ plane.

We denote by a the subset of N_a parameters under discussion (*e.g.*, $a = \{\bar\rho, \bar\eta\}$ and μ the N_μ remaining $y_{\rm mod}$ parameters[11].

The goal is to set CLs *in the a space, irrespective of the μ values.*

The smaller the region in the a space where the CL is sizeable (above ${\rm CL_{cut}} = 0.05$, say) the stronger the constraint is. The ultimate (and unattainable) goal being to make this allowed region shrink to a point: it would then correspond to the 'true' a. This means that one seeks to exclude the largest possible region of the a space. To do so, for a fixed value of a, one has to find the μ values which maximize the agreement between data and theory, and set the CL on a at the value corresponding to this optimized μ

$${\rm CL}(a) = {\rm Max}_\mu\{{\rm CL}(a,\mu)\} \ . \tag{36}$$

Proceeding that way, one uses the most conservative estimate for a given a point: this point will be engulfed in the excluded region only if ${\rm CL}(a,\mu) < {\rm CL_{cut}}$, $\forall\mu$. Stated differently, the CLs one is interested in are upper bounds of confidence levels. In effect, as discussed in Sect. 3.2.3, this is the standard procedure one uses to obtain CLs for a sub-set of fitted parameters.

3.2.2 Metrology of relevant parameters

According to the above discussion, we denote

$$\chi^2_{{\rm min};\,\mu}(a) \ , \tag{37}$$

the minimal value of the χ^2 function of (23), for a fixed value of a, when letting all μ parameters free to vary. For metrological purposes, one uses the offset-corrected χ^2

$$\Delta\chi^2(a) = \chi^2_{{\rm min};\,\mu}(a) - \chi^2_{{\rm min};y_{\rm mod}} \ , \tag{38}$$

the minimal value of which is zero, by construction: it is reached when a enters the $y_{\rm mod}^{\rm opt}$ domain.

Since only the minimal value of the χ^2 with respect to μ enters the *R*fit analysis, when μ contains a $y_{\rm QCD}(j)$ parameter which appears only in one measurement i, it is advisable to absorb its effect by computing beforehand

$$\mathcal{L}_{\rm exp}(i)_{{\rm max}:y_{\rm QCD}(j)} = {\rm Max}_{y_{\rm QCD}(j)}\{\mathcal{L}_{\rm exp}(i)\mathcal{L}_{\rm theo}(y_{\rm QCD})\} \ , \tag{39}$$

to clarify the analysis[12] (see Sect. 5).

Gaussian case: In a purely Gaussian situation one would directly obtain the CL for a as

$$\mathcal{P}(a) = {\rm Prob}(\Delta\chi^2(a), N_{\rm dof}) \ , \tag{40}$$

where $N_{\rm dof} = {\rm Min}(N_{\rm exp}^{\rm eff} - N_\mu, N_a)$. Equivalently, one may derive the same CL from the covariance matrix obtained from the fit leading to the absolute minimum, if in the a-region under consideration, the χ^2 is parabolic.

Non-Gaussian case: In a non-Gaussian situation, one has to consider $\Delta\chi^2(a)$ as a test statistics, and one has to rely on a Monte Carlo simulation to obtain its expected distribution in order to compute $\mathcal{P}(a)$. As further discussed in Sect. 3.3, this does not imply taking a Bayesian approach and to make use of PDFs for the unknown theoretical parameters μ.

[10] Although it can be argued that, while theoretical uncertainties dominate, pinning down $y_{\rm QCD}$ parameters might turn out to be the main (and not so interesting) achievement of a CKM analysis...

[11] It is worth to stress that this splitting is arbitrary and that it can be changed at will: for instance one may decide to focus only on $a = \{J\}$, or to consider $a = \{\sin 2\alpha, \sin 2\beta\}$, etc.

[12] This should be done for instance for FSI phases, most notably for the determination of the angles α and γ. But this cannot be done for the product $f_{B_d}\sqrt{B_d}$, because it appears in Δm_d but also indirectly in Δm_s, since $f_{B_d}\sqrt{B_d}$ and $f_{B_s}\sqrt{B_s}$ are theoretically linked

For the sake of simplicity, we use (40) in the present work. This implies that the experimental component $\mathcal{L}_{\rm exp}$ $(x_{\rm exp} - x_{\rm theo}(y_{\rm mod}))$ is free from non Gaussian contributions and inconsistent measurements. However, the $\Delta\chi^2(a)$ function itself does not have to be parabolic. What matters is that the $\mathcal{L}_{\rm exp}$ components are derived from Gaussian measurements (*c.f.*, Sect. 3.2.3 for an example), being understood that no $\mathcal{L}_{\rm theo}$ components are present.

3.2.3 Illustrations

To illustrate the above definitions we consider two specific examples in this section[13].

Standard situation: Consider an analysis which depends on only two quantities: the first is a fundamental parameter a, and the second is a QCD parameter $y_{\rm QCD}$. We assume here that the situation is a standard one, where it turns out that both quantities are simultaneously measurable: the full $\chi^2(a, y_{\rm QCD})$ function takes the form

$$\begin{aligned}\chi^2(a, y_{\rm QCD}) = &\left(\frac{a-a^0}{\sigma[a]}\right)^2 + \left(\frac{y_{\rm QCD}-y^0_{\rm QCD}}{\sigma[y_{\rm QCD}]}\right)^2 \\ &-2c\left(\frac{a-a^0}{\sigma[a]}\right)\left(\frac{y_{\rm QCD}-y^0_{\rm QCD}}{\sigma[y_{\rm QCD}]}\right) \\ &+\chi^2_{{\rm min};\,a,y_{\rm QCD}}\,, \end{aligned} \tag{41}$$

where c is a correlation coefficient. Applying the *R*fit scheme, the 95% CL interval for a is obtained as follows. One first computes the offset-corrected χ^2

$$\Delta\chi^2(a) = \chi^2_{{\rm min};\,y_{\rm QCD}}(a) - \chi^2_{{\rm min};\,a,y_{\rm QCD}}\,, \tag{42}$$

$$= \left(\frac{a-a^0}{\sigma[a]}\right)^2 (1-c^2)\,. \tag{43}$$

The limits $a_\pm$ of the 95% CL interval are such that

$$\mathrm{Prob}(\Delta\chi^2(a_\pm), 1) = 0.05 \rightarrow \Delta\chi^2(a_\pm) = 3.84\,, \tag{44}$$

and hence

$$a_\pm = a^0 \pm 1.96\,\frac{\sigma[a]}{\sqrt{1-c^2}}\,, \tag{45}$$

which is just the standard answer for the 95% CL interval of a, if one disregards information on $y_{\rm QCD}$.

Measurement of sin2β: If one uses $a = \{\bar\rho, \bar\eta\}$, a measurement of sin2$\beta$ alone yields a double infinite degeneracy corresponding to the solutions of (21), namely

$$\bar\eta = (1-\bar\rho)\frac{1 \pm \sqrt{1-(\sin 2\beta_{\rm exp})^2}}{\sin 2\beta_{\rm exp}}\,. \tag{46}$$

Along the two above straight lines in the $(\bar\rho, \bar\eta)$ plane, $\chi^2(a) = \chi^2_{{\rm min};y_{\rm mod}} = 0$. There is no μ parameters here, and hence

$$\begin{aligned}\Delta\chi^2(a) &= \chi^2_{{\rm min};\,\mu}(a) - \chi^2_{{\rm min};y_{\rm mod}} \\ &= \left(\frac{\sin 2\beta_{\rm exp} - \sin 2\beta_{\rm theo}}{\sigma[\sin 2\beta]}\right)^2\,. \end{aligned} \tag{47}$$

Using (40) one gets the CL in the $(\bar\rho, \bar\eta)$ plane

$$\mathcal{P}(\bar\rho, \bar\eta) = \mathrm{Prob}(\Delta\chi^2(a), 1)\,. \tag{48}$$

While the double infinite degeneracy of $\Delta\chi^2(a) = 0$ clearly precludes a parabolic behavior for this function, (48) remains exact due to the right hand side of (47).

3.3 Probing the SM

By construction, the metrological phase is unable to detect a failure of the SM to describe the data. This is because (38) wipes out the information contained in $\chi^2_{{\rm min};y_{\rm mod}}$. This value is a measure (a test statistics) of the best possible agreement between data and theory. Ideally, in a pure Gaussian case, this quantity could be turned into a CL referring to the SM as a whole in a straightforward way

$$\mathcal{P}(\mathrm{SM}) \le \mathrm{Prob}(\chi^2_{{\rm min};y_{\rm mod}}, N_{\rm dof})\,, \tag{49}$$

with $N_{\rm dof} = N^{\rm eff}_{\rm exp} - N_{\rm mod}$, were it be a positive number. The whole Standard Model being at stake one should not rely on (49), but use a Monte Carlo simulation to obtain the expected distribution of $\chi^2_{{\rm min};y_{\rm mod}}$. The Monte Carlo simulation is built as follows[14]:

- One selects a set of $y_{\rm mod}$ values within $y^{\rm opt}_{\rm mod}$ and assumes it to be the true one[15].
- Then one generates all $x_{\rm exp}(i)$, following the distribution of individual experimental likelihood component $\mathcal{L}_{\rm exp}(i)$, having reset their central values to the values $x_{\rm exp}(i) = x_{\rm theo}(i)$ computed with the above $y_{\rm mod}$ set. In case of significant experimental systematics, this implies the use of appropriate likelihoods as discussed in Sect. 5.
- In contrast to the above, one does not modify the $\mathcal{L}_{\rm theo}$ component of the likelihood: their central values are kept to their original settings. This is because these central values are not random numbers, but parameters contributing to the definition of $\mathcal{L}$.
- Then one computes the minimum of the χ^2 by letting all $y_{\rm mod}$ free to vary, as is done in the actual data analysis.
- From this sample of Monte Carlo simulations, one builds $\mathcal{F}(\chi^2)$, the distribution of $\chi^2_{{\rm min};y_{\rm mod}}$, normalized to unity.

[13] A more involved example is discussed in Appendix C

[14] For the sake of generality, the theoretical likelihood is not assumed to be necessarily the trivial *R*fit one (*c.f.*, Sect. 3.1.2)

[15] As discussed above, the various optimal realizations yield identical theoretical predictions, the choice made for a particular $y_{\rm mod}$ within $y^{\rm opt}_{\rm mod}$ is thus irrelevant. It was explicitly checked that the outcome of the analysis does not depend on this choice

The CL referring to the SM as a whole is then

$$\mathcal{P}(\mathrm{SM}) \leq \int_{\chi^2 \geq \chi^2_{\mathrm{min};y_{\mathrm{mod}}}} \mathcal{F}(\chi^2)\, d\chi^2 \ , \tag{50}$$

which is the upper bound to the CL one may set on the SM.

3.4 Probing new physics

If the above analysis establishes that the SM cannot accommodate the data, the next step is to probe for the new physics revealed by the observed discrepancy. The goal is akin to metrology: it is to measure new physical parameters y_{NP} (whose values, for example, are zero if the SM holds) complementing the set of y_{theo} parameters of the SM. The treatment is identical to the one of Sect. 3.2, using $a = \{y_{\mathrm{NP}}\}$. The outcome of the analysis is for example a 95% CL domain of allowed values for y_{NP} defined, in a first approximation, from (40)

$$\mathcal{P}(y_{\mathrm{NP}}) = \mathrm{Prob}(\Delta\chi^2(y_{\mathrm{NP}}), N_{\mathrm{NP}}) \geq 0.05 \ . \tag{51}$$

Even if the SM cannot be said to be in significant disagreement with data, it remains worthwhile to perform this metrology of new physics, for two reasons:

- It might be able to faster detect first signs of discrepancy between data and the SM, if the theoretical extension used in the analysis turns out to be the right one. The two approaches are complementary, the first (*c.f.*, Sect. 3.3) leads to a general statement about the validity of the SM, independently of any assumption for the new physics, the second is specific to a particular extension of the SM. In that sense it is less satisfactory. Being complementary, the two approaches can disagree: the first may conclude that the SM is in acceptable agreement with data, while the second may exclude the SM value $y_{\mathrm{NP}} = 0$, and, conversely, the first may invalidate the SM, while the second may lead to a fairly good value of $\mathcal{P}(y_{\mathrm{NP}} = 0)$, if the extension of the SM under consideration is not on the right track.
- The most sensitive observables, and the precision to be aimed at for their determination cannot be derived by any other means than by this type of analysis. When considering new experiments, it is therefore particularly valuable to have a sensitive model of new physics, to prioritize the efforts and set the precisions to be achieved.

4 Alternative statistical treatments

Several alternative statistical treatments are available. Three of them are briefly discussed below: however not all variations are considered. The relative merits and limitations of the three treatments will not be discussed extensively here, except to point out features of the *R*fit scheme.

4.1 Reminder: the *R*fit scheme

Let us briefly re-sketch the main steps of an analysis in the *R*fit scheme: for a given point a in the parameter space (*e.g.*, $a = \{\bar{\rho}, \bar{\eta}\}$) *R*fit proceeds to:

- Find the overall minimal $\chi^2_{\mathrm{min};y_{\mathrm{mod}}}$ with respect to all theoretical parameters.
- Perform a discrete, although fine scan of the a space, and minimize $\chi^2_{\mathrm{min};\mu}(a)$ with respect to the remaining parameters μ, for each point: the y_{QCD} parameters being allowed to vary freely within their $[y_{\mathrm{QCD}}]$ ranges.
- Calculate the offset-corrected CL ($\mathcal{P}(a)$ of (40)), for each point. It is the upper bound of the confidence levels one may set on a, which corresponds to the best possible set of theoretical parameters μ.
- Compute the CL of the overall $\chi^2_{\mathrm{min};y_{\mathrm{mod}}}$ by means of a Monte Carlo Simulation. It is the upper bound of the confidence levels one may set on the SM, which corresponds to the best possible set of y_{QCD} parameters.

The *R*fit scheme suffers from two drawbacks:

- It relies on *a priori* allowed ranges for the y_{QCD} parameters.
- In the hopeful case where data are such that the method is lead to "rule out" the SM, it provides no indication as to which y_{QCD} parameter(s) should be allowed to trespass its allowed range, and by how much, to rescue the SM.

4.2 The 95% scan method

The 95% CL Scan method [5] does not incorporate the theoretical component $\mathcal{L}_{\mathrm{theo}}$ except to define allowed ranges for the y_{QCD} values: in effect, this is equivalent to the *R*fit scheme. These y_{QCD} values are equidistantly scanned within their allowed ranges. For each set of y_{QCD} values (denoted *model* in the 95% CL Scan method terminology) three operations are performed:

1. One determines

$$\chi^2_{\mathrm{min};y_{\mathrm{theo}}}(y_{\mathrm{QCD}}) \ , \tag{52}$$

the minimal value of the χ^2 function of (23), from $\mathcal{L}_{\mathrm{exp}}$ only, for a fixed set of y_{QCD} values, when letting all y_{theo} parameters free to vary. One then computes the confidence level

$$\mathcal{P}(y_{\mathrm{QCD}}) = \mathrm{Prob}(\chi^2_{\mathrm{min};y_{\mathrm{theo}}}(y_{\mathrm{QCD}}), N_{\mathrm{dof}}) \ , \tag{53}$$

where $N_{\mathrm{dof}} = N^{\mathrm{eff}}_{\mathrm{exp}} - N_{\mathrm{theo}}$. If $\mathcal{P}(y_{\mathrm{QCD}})$ is above a threshold $\mathrm{CL}_{\mathrm{cut}}$ (usually $\mathrm{CL}_{\mathrm{cut}} = 0.05$) the *model* is considered as acceptable, and selected. The SM is ruled out if no *model* is selected.

2. Among the y_{theo} values, a subset a is retained as central values to be displayed for the current *model* (if selected) and CLs in the a space are derived using

$$\Delta\chi^2(a, y_{\mathrm{QCD}}) = \chi^2_{\mathrm{min};y_{\mathrm{theo}} \neq a}(a, y_{\mathrm{QCD}}) - \chi^2_{\mathrm{min};y_{\mathrm{theo}}}(y_{\mathrm{QCD}}) \ , \tag{54}$$

and

$$\mathcal{P}(a, y_{\rm QCD}) = {\rm Prob}(\Delta\chi^2(a, y_{\rm QCD}), N_{\rm dof}) \ , \qquad (55)$$

where $N_{\rm dof} = {\rm Min}(N_{\rm exp}^{\rm eff} - N_{\rm theo} + N_a, N_a)$.

3. The method concludes by a graphical display, for all selected models, of the contours in the a space defined by $\mathcal{P}(a, y_{\rm QCD}) = {\rm CL}_{\rm cont}$ (with ${\rm CL}_{\rm cont} = 0.05$).

4.2.1 Comparison with *R*fit

Although the outcome of the 95% CL Scan method, the graphical display, is quite different from the *R*fit one, both schemes are close in nature: they are frequentist approaches, flawed by the same double drawback mentioned in the previous section. In addition, while it is built on rather firm ground, the 95% CL Scan method presents several unwelcome features which are reviewed in Appendix B. The main differences between the two methods are:

- whereas *R*fit seeks for a statistical statement pertaining to a, and only to a, the 95% CL Scan method leads to statements on a, for given values of μ which take the form of the 95% CL contours.
- to correct for this, the 95% CL Scan method may conclude by an envelope, which delimits an allowed region in the a space with at least 95% CL.
- *R*fit draws a single smooth CL surface. From this surface one can read off the 95% CL contour, or define a family of contours, each corresponding to a given CL. These contours encircle a domain, the plateau of the CL surface, where the CL is essentially equal to unity.
- *R*fit is flexible. The default treatment can be extended to atone for the second fundamental drawbacks, and to accommodate for a smooth transition toward the Bayesian treatment, while, nevertheless, keeping part of the virtues of the 95% CL Scan method and of the *R*fit scheme. This is discussed in the next section.

4.3 The extended conservative method (*ER*fit)

The *R*fit scheme uses $\mathcal{L}_{\rm theo}(i)$ functions which trivially take only two values: either 1 within the allowed range, or 0 outside, thereby strictly forbidding any $y_{\rm QCD}$ to trespass $[y_{\rm QCD}]$. Instead, the extended *ER*fit scheme uses for $\mathcal{L}_{\rm theo}(i)$ functions which take values between 1 and 0. They are equal to 1 within $[y_{\rm QCD}]$ (there, they do not contribute at all to the full χ^2, and one recovers the *R*fit scheme) and drop smoothly to 0 outside. These functions are not PDFs: they are not combined the ones with the others through convolutions, and hence (see next section) the *ER*fit scheme is not a Bayesian scheme.

The precise way the functions decrease down to zero is obviously arbitrary: one needs to define a standard. The proposed expressions for $\mathcal{L}_{\rm theo}(i)$ are presented in Sect. 5. Their relevant characteristic is the following: they use two continuously varying parameters, denoted ζ and κ. The first parameter is a scale factor which fixes the allowed range where $\mathcal{L}_{\rm theo}(i) = 1$. The second parameter determines the transition to zero. The parameter values permit to cover a large spectrum of schemes, ranging from *R*fit ($\zeta = 1, \kappa = 0$), to a Gaussian scheme ($\zeta = 0, \kappa = 1$) and defining a standard[16], denoted *ER*fit, for which $\zeta = 1$, and $\kappa \simeq 0.8$.

Because *ER*fit acknowledges the fact that the allowed ranges should not be taken literally, it offers two advantages over *R*fit:

- *ER*fit is more conservative than *R*fit: by construction, a *ER*fit CL is always larger than the corresponding *R*fit one, and in the $(\bar{\rho}, \bar{\eta})$ plane its CL surface exhibits the same plateau at CL = 1 (i.e., the *R*fit and *ER*fit$y_{\rm mod}^{\rm opt}$ spaces are identical).
- In case the SM tends to be ruled out by *R*fit, the *ER*fit scheme is able to detect the eventual $y_{\rm QCD}$ parameter(s) which, if allowed to trespass its allowed range, would restore an acceptable agreement between data and theory, and which value it should take.

Despite the two above arguments in favor of *ER*fit, *R*fit *is chosen as the scheme advocated in this paper rather than* *ER*fit: because it uses a simpler and unique prescription to incorporate theoretical systematics, it is less prone to be confused with a Bayesian treatment. Moreover, *ER*fit does not provide a clear-cut distinction between statistical and theoretical systematic errors in the fit. Finally, in cases where one determines theoretical parameters *via* the fit, as it is the case, *e.g.*, for the quantity $f_{B_d}\sqrt{B_d}$, *R*fit is the natural choice. But, obviously, if *R*fit concludes to a SM failure, then *ER*fit should be used.

4.4 The Bayesian treatment

The Bayesian treatment [6,7] considers $\mathcal{L}$ as a PDF, from which is defined $\mathcal{F}(a)$, the PDF of a, through the convolution

$$\mathcal{F}(a) = C \int \mathcal{L}(y_{\rm mod})\ \delta(a - a(y_{\rm mod}))\ dy_{\rm mod} \ , \qquad (56)$$

where the constant C is computed *a posteriori* to ensure the normalization to unity of $\mathcal{F}(a)$. In practice, the integral can be obtained very conveniently by Monte Carlo techniques. For each point in the a space one sets a confidence level ${\rm CL}(a)$, for example, according to

$${\rm CL}(a) \ = \int\limits_{\mathcal{F}(a') \le \mathcal{F}(a)} \mathcal{F}(a')\ da' \ , \qquad (57)$$

but another definition for the domain of integration can de chosen. The method concludes by a graphical display of CL. In particular, the 95% CL contour can be read-off among others. New physics is not meant to be detected by the Bayesian treatment: it is aimed at metrology mostly.

[16] Obviously, it would be better if theorists, and not experimentalists, choose for these two parameters the values which appear the most adequate for each of their predictions

4.4.1 Comparison with *R*fit

Although their graphical display appear similar, the Bayesian treatment and the *R*fit scheme depart significantly: the meaning attached to a given CL value are not the same. For the Bayesian treatment, the CL is a quantity *defined*, using (56), for example by (57). The justification of this definition lies in the understanding that a CL value is meant to provide a quantitative measure of our qualitative *degree of belief*. Whereas one understands qualitatively well what is meant by *degree of belief*, because of its lack of formal definition, one cannot check that it is indeed well measured by the CL: the argument is thus circular. One is left with the sheer definition of (57), which, being just a definition, suffers no discussion.

The key point in the Bayesian treatment is the use of (56), even though the likelihood contains theoretical components. This implies that the $y_{\rm QCD}$ parameters, which stem from theorist computations, are to be considered as random realizations (*sic*) of their true values. The PDFs of these 'random' numbers are then drawned from guesswork (The $[y_{\rm QCD}]$ ranges do not fare better with respect to that.). For self-consistency, if one assumes that a large number of theorists perform the same $y_{\rm QCD}$ computation, the distribution of their results should then be interpreted as a determination of the $y_{\rm QCD}$ PDF. Once injected in (56), this PDF, the shape of which contains no information on nature, but information on the way physicist mind work, will be transformed into information pertaining to nature. This entails to a surprising confusion between what is an experimental result and what is a thinking result. As illustrated in Appendix A and in Sect. 7.2.1, it is less the *ad hoc* shapes of the PDFs which are at stake than the implication of using (56).

5 Likelihoods and systematic errors

In Sect. 3 we have defined the basic formalism of the *R*fit scheme. The treatment of experimental and theoretical systematics is the subject of this section.

Let x_0 be a quantity, which is not a random variable, but which is not perfectly known. We will consider in turn two quantities of this type:

- A theoretical parameter which is not well determined (*e.g.*, $x_0 = f_{B_d}$): the theoretical prediction of an observable depends on x_0 (*e.g.*, ΔM_{B_d}).
- An experimental bias due to detector/analysis defects: the measurement should be corrected for this bias.

It is the purpose of this section to suggest a prescription of how to incorporate the limited knowledge of such quantities into the analysis. The standard treatment of this problem relies on a χ^2 analysis

$$\chi^2 = \left(\frac{x_{\rm exp} - x_{\rm theo}}{\sigma_{\rm exp}}\right)^2 + \left(\frac{x_0 - \bar{x}_0}{\sigma_o}\right)^2 , \qquad (58)$$

where

- $x_{\rm theo}$ (resp. $x_{\rm exp}$) depends on x_0, if the latter is a theoretical parameter (resp. experimental systematics),
- $\bar{x}_0$ is the expected central value of x_0,
- σ_o is the *uncertainty* on x_0.

This standard treatment is satisfactory as long as the degree of belief we put on the knowledge of the value of x_0 is peaked at $\bar{x}_0$ and distributed like a Gaussian. This is usually summarized by

$$x_0 = \bar{x}_0 \pm \sigma_o \ . \qquad (59)$$

However, this is not necessarily what is intended to be meant by (59). Rather, the theorist (resp. the experimentalist) *may* mean that the prediction (resp. the measurement) can take any value obtained by varying x_0 at will within the range $[\bar{x}_0 - \zeta\sigma_o, \bar{x}_0 + \zeta\sigma_o]$ (denoted *the allowed range* below, where ζ is a constant scale factor of order unity), but that it is unlikely that x_0 takes its true value outside the allowed range. This does *not* imply that the possible values are equally distributed within the allowed range: they are not distributed *at all*[17]. If (59) is given such a meaning, then the statistical analysis should treat all x_0 values within the allowed range on the same footing (which again does not imply with equal *probability*): this corresponds to the *R*fit scheme (with $\zeta = 1$).

This is not the case for the χ^2 expression of (58) since the farther x_0 moves away from $\bar{x}_0$, the larger becomes the related component of χ^2.

On the other hand, it might also be useful to define specific tails instead of sharp cuts, thus allowing the theoretical parameters to leave their allowed ranges, if needed: this corresponds to the *ER*fit scheme.

The idea is to move from a pure χ^2 analysis to a log-likelihood one, redefining the χ^2 to be

$$\chi^2 = \left(\frac{x_{\rm exp} - x_{\rm theo}}{\sigma_{\rm exp}}\right)^2 - 2\ln\mathcal{L}_{\rm syst}(x_0) \ , \qquad (60)$$

where $\mathcal{L}_{\rm syst}(x_0)$, hereafter termed the *Hat* function, is a function equal to unity for x_0 within the allowed range. Its precise definition is given below.

5.1 The hat function

The Hat function $\mathcal{L}_{\rm syst}(x_0, \kappa, \zeta)$ is a continuous function defined as

$$-2\ln\mathcal{L}_{\rm syst}(x_0,\kappa,\zeta) = \begin{cases} 0 \ , & \forall x_0 \in [\bar{x}_0 \pm \zeta\sigma_o] \\ \left(\dfrac{x_0 - \bar{x}_0}{\kappa\sigma_o}\right)^2 - \left(\dfrac{\zeta}{\kappa}\right)^2 \ , & \forall x_0 \notin [\bar{x}_0 \pm \zeta\sigma_o] \end{cases} \qquad (61)$$

where the constant κ determines the behavior of the function outside the allowed range. For the *R*fit scheme $\kappa =$

[17] In some cases (*e.g.*, lattice QCD) statistical fluctuations may enter in the computation. In such instances one may reliably define a Gaussian likelihood for this component of the theoretical uncertainty

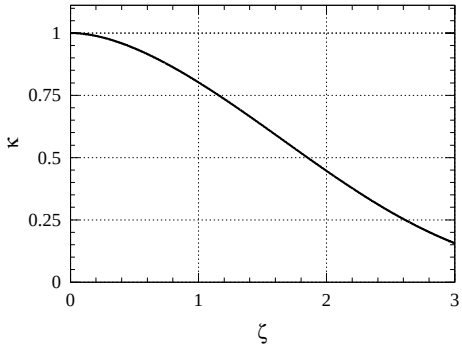

Fig. 3. The κ parameter as a function of ζ (see text)

0 is used. To define a standard κ can be chosen to be a function of ζ such that the relative normalizations of $\mathcal{L}_{\rm syst}(x_0,\kappa,\zeta)$ (briefly viewed here, for the purpose of defining a standard, as a PDF) within and outside the allowed range equal those of a Gaussian of width σ_o

$$\int_{-\infty}^{+\infty} \mathcal{L}_{\rm syst}(x_0,\kappa,\zeta)\; dx_0 \cdot \int_0^{\zeta/\sqrt{2}} e^{-t^2}\; dt = \sqrt{\pi}\,\zeta\sigma_0 \;. \qquad (62)$$

The parameter κ is numerically computed as a function of ζ. The result is shown in Fig. 3, in the range of interest $0 \le \zeta \le 3$. For the limit $\zeta \to 0$ one obtains $\kappa \to 1$, and the Hat becomes a pure Gaussian. The *ER*fit scheme is defined by $\zeta = 1$, for which one obtains $\kappa \simeq 0.8$.

Examples of Hat functions with $\bar{x}_0 = 0$ and $\sigma_o = 1$ are shown on the left plot of Fig. 4. Being a function, and not a PDF, $\mathcal{L}_{\rm syst}(x_0)$ needs not be normalized to unity.

5.2 Combining statistical and systematic uncertainties

Having defined $\mathcal{L}_{\rm syst}(x_0)$, and following (39), one proceeds with the minimization of the χ^2 of (60) by allowing x_0 to vary freely.

For theoretical systematics, the result depends on the way x_0 enters $x_{\rm theo}$, and not much more can be said in generality.

For experimental and theoretical systematics where x_0 can be assumed to be an unknown offset[18]: the quantity to be confronted to the theoretical prediction $x_{\rm theo}$ is simply $x_{\rm exp} + x_0$. Omitting the details of straightforward calculations, and assuming that $\bar{x}_0 = 0$ (otherwise $x_{\rm exp}$ should be corrected for it), one obtains, after minimization of the χ^2 with respect to x_0:

- $|\, x_{\rm exp} - x_{\rm theo} \,| \le \zeta\sigma_o$: $\quad \chi^2_{{\rm min};\,x_0} = 0$.
- $\zeta\sigma_o \le |\, x_{\rm exp} - x_{\rm theo} \,| \le \zeta\sigma_o(1 + (\frac{\sigma_{\rm exp}}{\kappa\sigma_o})^2)$:
  $\chi^2_{{\rm min};\,x_0} = \left(\frac{|x_{\rm exp}-x_{\rm theo}|-\zeta\sigma_o}{\sigma_{\rm exp}}\right)^2$.
- $|\, x_{\rm exp} - x_{\rm theo} \,| \ge \zeta\sigma_o(1 + (\frac{\sigma_{\rm exp}}{\kappa\sigma_o})^2)$:
  $\chi^2_{{\rm min};\,x_0} = \frac{(x_{\rm exp}-x_{\rm theo})^2}{\sigma^2_{\rm exp}+(\kappa\sigma_o)^2} - \left(\frac{\zeta}{\kappa}\right)^2$.

[18] If systematics take the form of an unknown multiplicative factor, and this is often the case for theoretical uncertainties, a treatment similar to the one discussed here applies

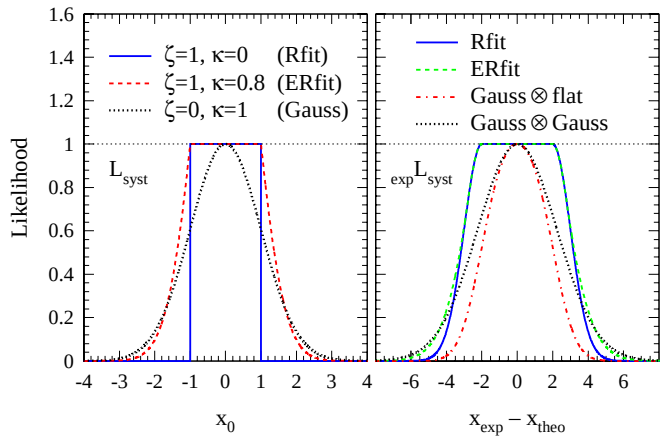

Fig. 4. The left hand plot shows the Hat functions ($\bar{x}_0 = 0$ and $\sigma_o = 1$) used for the *R*fit scheme, the *ER*fit scheme and the Gaussian treatment. The right hand plot shows the combined likelihood ${}_{\rm exp}\mathcal{L}_{\rm syst}$ (with $\bar{x}_0 = 0$ and $\sigma_{\rm exp} = \sigma_o = 1$) obtained from (39) for the *R*fit scheme, the *ER*fit scheme, a convolution of a Gaussian with a uniform distribution (hence taken as a PDF, following the Bayesian approach) and a convolution of two Gaussians (see Appendix A)

In the limit $\zeta \to 0$ (and hence, $\kappa \to 1$) only the third instance is met, and one recovers the usual rule of adding in quadrature the statistical and the systematic uncertainties. Otherwise, the result is non trivial. An example of the effective likelihood ${}_{\rm exp}\mathcal{L}_{\rm syst}(x_{\rm exp} - x_{\rm theo}) \equiv -\frac{1}{2}\chi^2_{{\rm min};\,x_0}$ (with $\bar{x}_0 = 0$ and $\sigma_{\rm exp} = \sigma_o = 1$) is shown in the right hand plot of Fig. 4 for the *R*fit scheme, the *ER*fit scheme, a convolution of a Gaussian with a uniform distribution (hence taken as a PDF, following the Bayesian approach) and a convolution of two Gaussians.

6 Fit ingredients

This section provides a compendium of the measurements and SM predictions entering the overall constrained CKM fit. In some cases, we pre-combine compatible measurements by means of a simple weighted mean in order to speed up the fit by reducing the effective number of degrees of freedom (see the discussion in Sect. 3.1.1). Since the solution of a χ^2 minimization of Gaussian distributed measurements corresponds to the weighted mean, this entails no loss of information. Below, we give a status of the input quantities used. The corresponding numerical values that enter the fit and the treatment of their uncertainties within *R*fit are summarized in Table 1.

6.1 The CKM matrix elements

- $|V_{ud}|$ **:** The matrix element $|V_{ud}|$ has been extracted by means of three different methods: super-allowed nuclear β-decays, neutron β-decay and pion β-decay. Using the lifetime measurements of super-allowed nuclear β-decays with pure Fermi-transitions ($0^+ \to 0^+$),

Table 1. Inputs to the global CKM fit. If not stated otherwise: for two errors given, the first is statistical and accountable systematic and the second stands for systematic theoretical uncertainties. The fourth, fifth and sixth columns indicate the treatment of the parameters within *R*fit: measurements dominated by experimental errors (or statistical components of theoretical parameters) are marked as "Gauss." by an asterisk; parameters dominated by systematic theoretical uncertainties, treated as ranges in *R*fit, are marked as "Theo."; for parameters that have experimental and systematic theoretical errors, treated in the fit according to (61), both fields, "Exp." and "Theo.", are marked; parameters with small errors, marked as "Prop.", have their uncertainties propagated to the corresponding measurements to whose errors they are added in quadrature. The last column indicates whether or not the parameter is floating in the fit. In general, measurements with non-vanishing systematic theoretical errors have a floating theoretical component. Theoretical parameters with significant errors are necessarily floating. *Upper part:* experimental determinations of the CKM matrix elements. *Middle upper part:* CP-violating and mixing observables. *Middle lower part:* parameters of the SM predictions obtained from experimental data. *Lower part:* parameters of the SM predictions obtained from theory

			Error treatment in *R*fit:			
Parameter	Value ± Error(s)	Reference(s)	Gauss.	Theo.	Prop.	Float.
$\|V_{ud}\|$	0.97394 ± 0.00089	see text	*	-	-	-
$\|V_{us}\|$	0.2200 ± 0.0025	see text	*	-	-	-
$\|V_{ub}\|$	$(3.49 \pm 0.23 \pm 0.55) \times 10^{-3}$	see text	*	*	-	*
$\|V_{cd}\|$	0.224 ± 0.014	see text	*	-	-	-
$\|V_{cs}\|$	0.969 ± 0.058	see text	*	-	-	-
$\|V_{cb}\|$	$(40.76 \pm 0.50 \pm 2.0) \times 10^{-3}$	see text	*	*	-	*
$\|\epsilon_K\|$	$(2.271 \pm 0.017) \times 10^{-03}$	see text	*	-	-	-
Δm_d	(0.487 ± 0.014) ps^{-1}	see text	*	-	-	-
Δm_s	Amplitude spectrum	[94], see text	*	-	-	-
$\sin 2\beta_{\mathrm{WA}}$	0.48 ± 0.16	see text	*	-	-	-
m_c	(1.3 ± 0.1) GeV	[10]	-	*	-	*
$m_t(\overline{\mathrm{MS}})$	(166.0 ± 5.0) GeV	[10]	*	-	-	*
m_K	(493.677 ± 0.016) MeV	[10]	-	-	*	-
Δm_K	$(3.4885 \pm 0.0008) \times 10^{-15}$ GeV	[10]	-	-	*	-
m_{B_d}	(5.2794 ± 0.0005) GeV	[10]	-	-	*	-
m_{B_s}	(5.3696 ± 0.0024) GeV	[10]	-	-	*	-
m_W	(80.419 ± 0.056) GeV	[10]	-	-	*	-
G_F	$(1.16639 \pm 0.00001) \times 10^{-5}$ GeV^{-2}	[10]	-	-	-	-
f_K	(159.8 ± 1.5) MeV	[10]	-	-	*	-
B_K	$0.87 \pm 0.06 \pm 0.13$	see text	*	*	-	*
η_{cc}	1.38 ± 0.53	see text	-	*	-	*
η_{ct}	0.47 ± 0.04	see text	-	-	*	-
η_{tt}	0.574 ± 0.004	see text	-	-	*	-
$\eta_B(\overline{\mathrm{MS}})$	0.55 ± 0.01	see text	-	*	-	*
$f_{B_d}\sqrt{B_d}$	$(230 \pm 28 \pm 28)$ MeV	see text	*	*	-	*
ξ	$1.16 \pm 0.03 \pm 0.05$	see text	*	*	-	*

$|V_{ud}|$ can be extracted with high precision. Averaging the values of nine different super-allowed nuclear β-decays [17] the result including nucleus-independent and nucleus-dependent radiative corrections as well as charge-dependent corrections (see [18] for a compendium of references) is: $|V_{ud}| = 0.97400 \pm 0.00014_{\mathrm{exp}} \pm 0.00048_{\mathrm{theo}}$ [18,17]. The precision of $|V_{ud}|$ from nuclear β-decays is often questioned in light of the observed '2σ' deviation from the unitarity condition when combining this value with the best knowledge for $|V_{us}|$ and $|V_{ub}|$. A possible enhancement of $|V_{ud}|$ is predicted by a quark-meson coupling model due to a change of charge symmetry violation for quarks inside bound nucleons compared to unbound nucleons [19]. Since the status here is unclear, the error has been enlarged by the amount of the possible correction using the PDG rescaling [10]: $|V_{ud}| = 0.9740 \pm 0.0010$.[19]

[19] It should be kept in mind that such a procedure might hide a possible violation of the unitarity condition in the first family

For the neutron β-decay no nuclear structure effects play any role. However, $|V_{ud}|$ has to be extracted from two quantities, the neutron lifetime and the ratio g_A/g_V. In contrast to nuclear β-decays these measurements are not dominated by theoretical uncertainties. The weighted mean for the neutron lifetime measurements is $\tau_n = (885.7 \pm 1.0)$ s [20] and the average value for g_A/g_V reads -1.2699 ± 0.0029 [21] where the error was rescaled by a factor of two due to inconsistencies in the data set. Combining these numbers we get: $|V_{ud}| = 0.9738 \pm 0.0020_{g_A/g_V,\tau_n} \pm 0.0004_{\rm rad}$, where the error is dominated by the experimental error on g_A/g_V.

The pion β-decay $\pi^+ \to \pi^0 e^+ \nu_e$ is a very attractive candidate to extract $|V_{ud}|$ from the branching ratio and the pion lifetime, since it is mediated by a pure vector transition and does not suffer from nuclear structure effects. However, due to the small branching ratio, BR $= (1.025 \pm 0.034) \times 10^{-8}$ [10], the present statistical accuracy is not competitive with the other methods: $|V_{ud}| = 0.967 \pm 0.016_{\rm BR} \pm 0.0009_{\rm theo}$.

Assuming all three measurements to be Gaussian distributed[20] we combine them *via* a weighted mean, yielding: $|V_{ud}| = 0.97394 \pm 0.00089$.

– $\boldsymbol{|V_{us}|}$ **:** The analyses of kaon and hyperon semileptonic decays provide the best determination of $|V_{us}|$ directly related to λ in the Wolfenstein parameterization. However, due to theoretical uncertainties from the breakdown of SU(3) flavor symmetry, the hyperon decay data are less reliable [22,23]. As a consequence, we use only the value obtained from the vector transitions $K^+ \to \pi^0 e^+ \nu_e$ and $K_L \to \pi^- e^+ \nu_e$ [10]. Owing to the small electron mass, only one form factor plays a role in $K_{\rm e3}$ decays the functional dependence of which can be extracted from data. The form factor value at zero recoil, $f_1(0)$, is calculated within the framework of chiral perturbation theory and is found to be: $f_1^{K^0\pi^-}(0) = 0.961 \pm 0.008$ [24]. The error estimate for this value was criticized in [25]. However, a relativistic constituent quark model, successful in the description of electroweak properties of light mesons gives a consistent result: $f_1^{K^0\pi^-}(0) = 0.963 \pm 0.004$ [26]. Channel-independent and channel-dependent radiative corrections [27,28] as well as charge symmetry (K^+/K_L) and charge independence (π^-/π^0) breaking corrections [24] are applied in order to compare the results from both channels: $f_1^{K^0\pi^-}(0)|V_{us}| = 0.2134 \pm 0.0015_{\rm exp} \pm 0.0001_{\rm rad}$ ($K^+ \to \pi^0 e^+ \nu_e$) and $f_1^{K^0\pi^-}(0)|V_{us}| = 0.2101 \pm 0.0013_{\rm exp} \pm 0.0001_{\rm rad}$ ($K_L \to \pi^- e^+ \nu_e$). The weighted average is: $f_1^{K^0\pi^-}(0)|V_{us}| = 0.2114 \pm 0.0016$ where the error was rescaled by a factor of 1.6 to account for inconsistencies.

The result then reads $|V_{us}| = 0.2200 \pm 0.0017_{\rm exp} \pm 0.0018_{\rm theo} = 0.2200 \pm 0.0025$. As in the case for V_{ud} all uncertainties were considered as Gaussian errors.

– $\boldsymbol{|V_{ub}|}$ **:** Both, inclusive B-decays ($b \to X_u \ell^- \bar{\nu}_\ell$), measured at LEP [29–31], and exclusive B-decays ($B^0 \to \pi^- \ell^+ \nu_\ell$, $B^0 \to \rho^- \ell^+ \nu_\ell$), measured by CLEO [32], allow an extraction of the third column element $|V_{ub}|$[21]. The exclusive measurements are dominated by the theoretical uncertainty due to the model dependence in the determination of the form factor. The exclusive CLEO measurements give:

$$|V_{ub}| = (3.25 \pm 0.14_{\rm stat}{}^{+0.21}_{-0.29}{}_{\rm sys} \pm 0.55_{\rm theo}) \times 10^{-3} \ ,$$

where the error is dominated by the theoretical uncertainties. We add the statistical and experimental systematic error in quadrature and consider the theoretical error as a range: $|V_{ub}| = (3.25 \pm 0.29_{\rm exp} \pm 0.55_{\rm theo}) \times 10^{-3}$. There is some hope that exclusive measurements in the future may take advantage of unquenched lattice QCD calculations and thus reduce the model dependent error.

The three inclusive LEP measurements rely on different techniques and are combined in [40], taking into account all uncorrelated and correlated errors: $|V_{ub}| = (4.04^{+0.41}_{-0.46}({\rm stat+det})^{+0.41}_{-0.46}(b \to c)^{+0.24}_{-0.25}(b \to u) \pm 0.02(\tau_b) \pm 0.19({\rm HQE})) \times 10^{-3}$. The theoretical uncertainty from the Heavy Quark Expansion (HQE) is a combination of three sources: the neglect of higher terms including $1/m_b^3$, the uncertainty in the b-quark mass m_b and from perturbative corrections. The extraction of $|V_{ub}|$ from the inclusive semileptonic BR($b \to {\rm u}\, \ell^- \bar{\nu}_\ell$) relies on the validity of quark-hadron duality. Although quark-hadron duality can not be expected to be exact, there are good reasons that inclusive semileptonic decays of beauty hadrons are described quite accurately by HQE [41]. However, since the analyses have to apply cuts in order to suppress the background from $b \to c$ transitions only a part of the total semileptonic rate is measured which could lead to sizable effects from quark-hadron duality violation which are difficult to quantify. In the case of the inclusive measurements we combine the theoretical uncertainties from HQE and the large systematic uncertainties due to $b \to c$ and $b \to u$ transitions by adding them in quadrature and obtain: $|V_{ub}| = (4.04 \pm 0.44_{\rm exp} \pm 0.54_{\rm model}) \times 10^{-3}$.

For the combined result of inclusive and exclusive measurements we obtain: $|V_{ub}| = (3.49 \pm 0.24 \pm 0.55_{\rm theo}) \times 10^{-3}$ where only the first error was used for the weighted mean. The maximum of both single ranges was assigned as the final systematic theoretical uncertainty.

[20] Being consequent, also in this case, one should treat the theoretical errors as ranges. However, as long as the relative uncertainty on λ from $|V_{ud}|$ and $|V_{us}|$ is much smaller than what one obtains for $(\bar{\rho}, \bar{\eta})$ from constraints like $|V_{ub/V_{cb}}|$, $|\epsilon_K|$, Δm_d and $\Delta m_s/\Delta m_d$, the procedure used is certainly not a critical issue

[21] The determination of $|V_{ub}|$ from the lepton endpoint spectrum, obtained by ARGUS [33] and CLEO [34,35], suffers from large model dependencies [36–39]. In addition, possible violations of quark-hadron duality might be enhanced in this small part of the phase space. Hence, these results were not taken into account for this analysis

– $|V_{cd}|$ **:** Both matrix elements, $|V_{cd}|$ and $|V_{cs}|$, can be determined from di-muon production in deep inelastic scattering (DIS) of neutrinos and anti-neutrinos on nucleons. In an analysis performed by CDHS [42], $|V_{cd}|$ and $|V_{cs}|$ are extracted by fitting the data of three experiments, CDHS [42], CCFR [43] and CHARM II [44], giving: $|V_{cd}|^2 \times B_c = (4.63 \pm 0.34) \times 10^{-3}$, where $B_c = 0.0919 \pm 0.0094$ [45–47] is the weighted average of semileptonic branching ratios of charmed hadrons produced in neutrino-nucleon DIS. This results in: $|V_{cd}| = 0.224 \pm 0.014$ [18].

– $|V_{cs}|$ **:** Besides DIS, the matrix element $|V_{cs}|$ can be obtained from D_{e3} decays, charm tagged W decays and hadronic W decays.
The average DIS result from CDHS, CCFR and CHARM II is $\kappa|V_{cs}|^2 B_c = (4.53 \pm 0.37) \times 10^{-2}$ where $\kappa = 0.453 \pm 0.106^{+0.028}_{-0.096}$ is the relative size of strange quarks in the sea compared to $\bar{u}$ and $\bar{d}$ resulting in $|V_{cs}| = 1.04 \pm 0.16$ [10].
Similar to K_{e3}, decays $|V_{cs}|$ can be also extracted from D_{e3} decays. However, the theoretical uncertainty in the form factor calculation $f_1(0) = 0.7 \pm 0.1$ [48] limits the precision: $|V_{cs}| = 1.04 \pm 0.16$ [18], in perfect agreement with $|V_{cs}|$ from DIS.
Under the assumption that unitarity holds for three families, the ratio $R_c = \Gamma(W^+ \to c\bar{q})/\Gamma(W^+ \to \text{hadrons}) = \sum_{i=d,s,b} |V_{ci}|^2 / (\sum_{i=d,s,b;j=u,c} |V_{ji}|^2)$ for W decays is expected to be 1/2. The results of all LEP2 experiments are consistent with this expectation and give $|V_{cs}| = 0.97 \pm 0.09_{\text{stat}} \pm 0.07_{\text{sys}}$ [10,49–52]. The ratio of hadronic and leptonic W decays measured by LEP2 provides the tightest bound on $|V_{cs}|$ if unitarity for three families is assumed: $R_c = \Gamma(W^+ \to \text{hadrons}/\Gamma(W^+ \to \text{leptons}) = \sum_{i=d,s,b;j=u,c} |V_{ji}|^2 \times (1 + \alpha_s(m_W)/\pi)$. From the four LEP experiments $|V_{cs}| = 0.989 \pm 0.016$ [10] is found where the errors on the single measurements are dominated by statistical errors[22].
Very recently, the OPAL collaboration has presented a new direct determination of $|V_{cs}|$ from $W \to X_c X$ resulting in $|V_{cs}| = 0.969 \pm 0.058$ [53], which we use in the fit.

– $|V_{cb}|$ **:** In the Wolfenstein parameterization, $|V_{cb}|$ determines the parameter A the precision of which plays an important part for the constraints $|V_{ub}/V_{cb}|$, $|\epsilon_K|$ and Δm_d. It is obtained from exclusive $B \to D^{(*)}\ell\bar{\nu}_\ell$ and inclusive semileptonic b decays to charm, $b \to \ell^- \bar{\nu}_\ell X_c$, both measured by CLEO and the LEP experiments. The theoretical framework for extracting numerical values for $|V_{cb}|$ from the measured decay rates is Heavy Quark Effective Theory (HQET) [54] for exclusive measurements and HQE [55] for inclusive measurements.
The exclusive results are given in the form $\mathcal{F}_{D^*}(1)|V_{cb}|$, so that they must be divided by the value of the Isgur-Wise function at zero-recoil. The form factor at zero-recoil, $\mathcal{F}_{D^*}(1)$, is 1 in the heavy quark limit. For finite quark masses the corrections can be calculated in HQET. There are open discussions in the literature concerning the calculation of the $1/m_Q^2$ corrections for $\mathcal{F}_{D^*}(1)$. Here we use the value $\mathcal{F}_{D^*}(1) = 0.913 \pm 0.042$ [5] for which several references have been taken into account. In the future, the most accurate determinations of $\mathcal{F}_{D^*}(1)$ are expected to come from lattice QCD. A first result with a quite small error reads $\mathcal{F}_{D^*}(1) = 0.935 \pm 0.030$ [56]. The results from LEP and CLEO read

$$\begin{aligned}
\mathcal{F}_{D^*}(1)|V_{cb}|_{\text{LEP}} &= 0.0350 \pm 0.0007_{\text{stat}} \pm 0.0015_{\text{sys}} \\
&\quad [40,57] \\
\mathcal{F}_{D^*}(1)|V_{cb}|_{\text{CLEO}} &= 0.0424 \pm 0.0018_{\text{stat}} \pm 0.0019_{\text{sys}} \\
&\quad [58] \\
\mathcal{F}_{D^*}(1)|V_{cb}| &= 0.0373 \pm 0.0013 \\
|V_{cb}| &= 0.0409 \pm 0.0014 \pm 0.0019_{\text{theo}}.
\end{aligned}$$

The combined fit of the CLEO and LEP numbers results in a confidence level of 7%.
The theoretical error for the inclusive measurement contains the uncertainty in the kinetic energy μ_π^2 of the b-quark inside the b-hadron and uncertainties from perturbative corrections, the b-quark mass and the neglect of higher order terms in the $1/m_b$ expansion including $1/m_b^3$ terms [59]. As in the case of the inclusive determination of V_{ub} possible violations of quark-hadron duality could imply sizable effects. Future experimental investigations should aim to shed more light on this topic. The most recent inclusive results read

$$\begin{aligned}
|V_{cb}|_{\text{LEP}} &= 0.04076 \pm 0.00050_{\text{exp}} \\
&\quad \pm 0.00204_{\text{theo}} \qquad [40,57] \\
|V_{cb}|_{\text{CLEO}} &= 0.041 \pm 0.0010_{\text{stat}} \pm 0.0020_{\text{sys}} \\
&\quad \pm 0.00205_{\text{theo}} \qquad [60,61]
\end{aligned}$$

The weighted mean of exclusive and inclusive results is $|V_{cb}| = (40.76 \pm 0.50 \pm 2.0_{\text{theo}}) \times 10^{-3}$, and is dominated by the inclusive measurement. In light of the controversial experimental and theoretical situation in the exclusive sector and possible violations of quark-hadron duality the theoretical uncertainty was not further reduced.

– $|V_{tb}|$ **:** Assuming unitarity for three families, one obtains $|V_{tb}|$ from the ratio of the bottom quark production in top decays to the total top decay width: $|V_{tb}| = 0.99 \pm 0.15$ [62]. The unitarity assumption, here explicitly used, could be removed. However, owing to the poor precision presently achieved, we do not use this measurement in the global CKM fit.

– $|V_{ts}V_{tb}/V_{cb}|$ **:** The inclusive ratio of $b \to s\gamma$ to $b \to c\,\ell^-\bar{\nu}_\ell$ production provides a measure of the third row CKM elements. Present accuracy is only fair: $|V_{ts}V_{tb}/$

[22] The measurement of R_c should be used in the fit, rather than the quoted $|V_{cs}|$ determination which is derived from it. This piece of information is not used here, for the sake of simplicity

$V_{cb}| = 0.93 \pm 0.14 \pm 0.08$ [63–65]. Hence, the value is not used in the fitting procedure.

6.2 CP observables and mixing

Constraints on the CKM phase are obtained by the CP-violating observables in the $K^0 - \bar{K}^0$ and $B_d^0 - \bar{B}_d^0$ systems, and by $B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$ mixing.

– $\boldsymbol{|\epsilon_K|}$ **:** Indirect CP violation in the $K^0 - \bar{K}^0$ system is measured by

$$\epsilon_K = \frac{2}{3}\eta_{+-} + \frac{1}{3}\eta_{00} \ , \tag{63}$$

with η_{+-} (η_{00}) being the ratio of the amplitudes of the long-lived and short-lived neutral kaons decaying into two charged (neutral) pions. They have been measured to an accuracy of 1% [66–69]. The averaged value of [10] is used in this analysis. Within the SM, CP violation is induced by $\Delta S = 2$ transitions owing to box diagrams. It can be related to the CKM-matrix elements by means of the vacuum insertion approximation, used to determine the hadronic matrix element

$$\langle \bar{K}^0 | (\bar{s}\gamma^\mu(1-\gamma^5)d)^2 | K^0 \rangle = \frac{8}{3} m_K^2 f_K^2 B_K \ . \tag{64}$$

Neglecting the real part of the non-diagonal element of neutral kaon mass matrix M_{12}, one obtains

$$\begin{aligned} |\epsilon_K| = \frac{G_F^2 m_W^2 m_K f_K^2}{12\sqrt{2}\pi^2 \Delta m_K} B_K \Big(& \eta_{cc} S(x_c, x_c) \mathrm{Im}\big[(V_{cs}V_{cd}^*)^2\big] \\ & + \eta_{tt} S(x_t, x_t) \mathrm{Im}\big[(V_{ts}V_{td}^*)^2\big] + 2\eta_{ct} S(x_c, x_t) \\ & \times \mathrm{Im}\big[V_{cs}V_{cd}^* V_{ts}V_{td}^*\big] \Big) \end{aligned} \tag{65}$$

Here, the $S(x_i, x_j)$ are the Inami-Lim functions [70]

$$\begin{aligned} S(x) &\equiv S(x_i, x_j)_{i=j} \\ &= x\left(\frac{1}{4} + \frac{9}{4(1-x)} - \frac{3}{2(1-x)^2}\right) \\ &\quad - \frac{3}{2}\left(\frac{x}{1-x}\right)^3 \ln(\mathrm{x}) \ , \\ S(x_i, x_j)_{i\neq j} &= x_i x_j \left[\left(\frac{1}{4} + \frac{3}{2(1-x_i)} - \frac{3}{4(1-x_i)^2}\right) \right. \\ &\quad \times \frac{1}{x_i - x_j} \ln(\mathrm{x_i}) + (x_i \leftrightarrow x_j) \\ &\quad \left. - \frac{3}{4} \frac{1}{(1-x_i)(1-x_j)} \right] \ , \end{aligned} \tag{66}$$

depending on the masses of the virtual charm and top quarks in the box diagrams ($x_i = m_i^2/m_W^2$). The QCD corrections to the Inami-Lim functions have been calculated to next-to-leading order: $\eta_{cc} = 1.38 \pm 0.53$ [72], $\eta_{tt} = 0.574 \pm 0.004$ [73] and $\eta_{ct} = 0.47 \pm 0.04$ [72] (for a compendium see also [71]). The kaon decay constant has been extracted from the leptonic decay rate: $f_K = (159.8 \pm 1.4_{|V_{us}|} \pm 0.44_{\mathrm{theo}})$ MeV [10]. The $K_S - K_L$ mass difference is known with excellent accuracy, $\Delta m_K = (3.4885 \pm 0.0008) \times 10^{-15}$ GeV [10]. The main uncertainty in (65) originates from the bag parameter B_K which cannot be measured but has to be predicted by theory. The most reliable calculations of B_K are supposed to come from lattice QCD. Currently, these calculations are performed only under the assumption of $SU(3)$ symmetry using the quenched approximation, *i.e.*, using quarks with infinite masses and neglecting the contribution of sea-quarks in closed loops, which leads to a substantial reduction in computing time. The world average reads: $B_K = 0.87 \pm 0.06 \pm 0.14_{\mathrm{quench}}$ [75], where the first error combines statistical and accountable systematic uncertainties while the second stands for an estimate of the error introduced by the quenched approximation and $SU(3)$ breaking effects.

– $\boldsymbol{\epsilon'/\epsilon_K}$ **:** In terms of the neutral kaon amplitude ratios η_{+-} and η_{00}, one finds to a very good approximation

$$\mathrm{Re}(\epsilon'/\epsilon_K) = \frac{1}{6}(1 - |\eta_{00}/\eta_{+-}|^2) \ . \tag{67}$$

The first evidence of direct CP violation in the neutral kaon system has been found by the NA31 Collaboration at CERN [76] which was not confirmed by the E731 Collaboration [77]. Since then, measurements at KTeV [78] and NA48 [79] have obtained significant positive results while some inconsistencies about the central value remain. The experimental values are

$$\mathrm{Re}(\epsilon'/\epsilon_K) \times 10^4 = \begin{cases} 23.0 \pm 6.5 \ (\mathrm{NA31}[76]) \\ 7.4 \pm 5.9 \ (\mathrm{E731}\ [77]) \\ 28.0 \pm 4.1 \ (\mathrm{KTeV}[78]) \\ 14.0 \pm 4.3 \ (\mathrm{NA48}[79]) \end{cases} \tag{68}$$

for which the weighted mean of $\langle \mathrm{Re}(\epsilon'/\epsilon_K) \rangle = (19.2 \pm 2.5) \times 10^{-4}$ has a consistency of $\chi^2/3 = 3.5$. The experimental situation being somewhat inconsistent, the theoretical prediction in the framework of the SM is still under strong investigations. The basic expression is of the form [80]

$$\mathrm{Re}(\epsilon'/\epsilon_K) = \mathrm{Im}\left[V_{ts}^* V_{td} V_{us}^* V_{ud}\right] \Sigma\left(B_6, B_8, m_s \langle s\bar{s} \rangle\right) \tag{69}$$

with $\Sigma(\ldots)$ being a function of the hadronic matrix elements, B_6 and B_8, of the dimension-6 and dimension-8 non-perturbative power corrections which contribute to the effective Hamiltonian. The quantitative size of these operators, in particular B_6, and the exact ingredients of $\Sigma(\ldots)$ are still under consideration [80].
As a consequence, we shall not use ϵ'/ϵ_K in the current analysis.

– $\boldsymbol{\Delta m_d}$ **:** The frequency of $B_d^0 - \bar{B}_d^0$ oscillation is given by the mass difference, Δm_d, between the two B_d^0 mass eigenstates, B_{H} and B_{L}. It has been measured to an

accuracy of 3%[23] (see Table 1). In analogy to $|\epsilon_K|$, $B_d^0 - \bar{B}_d^0$ oscillation in the SM is driven by effective flavor-changing neutral current (FCNC) processes through $\Delta B = 2$ box diagrams. In contrary to $|\epsilon_K|$, where the large hierarchy in the Inami-Lim functions is partly compensated by the CKM matrix elements, the $\Delta B = 2$ box diagrams are dominated by top quark exchange between the virtual $W^\pm$ boson lines. This simplifies the theoretical prediction which then reads

$$\Delta m_d = \frac{G_F^2}{6\pi^2}\eta_B m_{B_d} f_{B_d}^2 B_d m_W^2 S(x_t)\,|V_{td}V_{tb}^*|^2 \ , \quad (70)$$

where $\eta_B = 0.55 \pm 0.01$ (for a review see [71]) is a correction to the Inami-Lim function $S(x_t)$ (see (66)) from perturbative QCD. The leptonic decay constant f_{B_d} has not been measured so far and, like the bag parameter B_d, has to be determined by theory, in particular lattice QCD. Up to now, calculations are mainly performed in the quenched approximation where the different groups find consistent results. The most recent world averages for the decay constant is $f_{B_d} = (175 \pm 20)$ MeV [83] where the error includes statistical and accountable systematic uncertainties. The unquenched result is estimated to be about 10% higher than the quenched value, $f_{B_d} = (200 \pm 23^{+27}_{-17})$ MeV [83, 84]. Recently, first (partly) unquenched calculations with two degenerate sea quarks were published and are, within the given uncertainties, in agreement with the expected increase [85,86]. The world average for the bag parameter in the quenched approximation is $B_d = 1.30 \pm 0.12 \pm 0.13$, where the second error is the estimated uncertainty due to the quenched approximation [83]. In the present work we use $f_{B_d}\sqrt{B_d} = (230 \pm 28 \pm 28)$ MeV [83], where the second error has been symmetrized.

– $\boldsymbol{\Delta m_s}$ **:** Although Δm_s itself has only a weak dependence on the CKM phase the ratio $\Delta m_s/\Delta m_d$ introduces a strong constraint since the dependence of the SM prediction on the parameters η_B and m_t cancel in the ratio. Furthermore, the ratio $\xi = f_{B_s}\sqrt{B_s}/f_{B_d}\sqrt{B_d}$ can be calculated more reliably from lattice QCD than $f_{B_d}\sqrt{B_d}$ alone since most of the systematics cancel. For ξ, we are combining the average values for quenched calculations from [83,6] and choose:

$$\xi = 1.16 \pm 0.03_{\text{stat,sys}} \pm 0.05_{\text{quench}} \ .$$

Limits on $B_s^0 - \bar{B}_s^0$ oscillation governed by the mass difference Δm_s have been obtained by ALEPH [87], DELPHI [88], OPAL [89], SLD [90] and CDF [91]. A convenient approach to average various results on Δm_s is the Amplitude Method [92] (see also the exhaustive study in [93]), which consists of a likelihood fit to the measured proper time distribution with the amplitude of the oscillating term being the free parameter at given frequency Δm_s. Figure 5 shows the average of the measured amplitude spectrum [94] with the expected spectra for different true Δm_s superimposed. The latter have been obtained, following the prescription of [93], from a toy simulation in which the decay length and momentum resolutions are tuned to reproduce the measured errors on the amplitudes (the RMS of the relative difference between measured and simulated errors is smaller than 2% in the relevant sensitive region of Δm_s). Shown in addition are the experimental sensitivity defined as $1.645 \times \sigma_A$ for a given Δm_s and the 95% CL limit which is given by the sum of the sensitivity and the measured central amplitude. A sensitivity of 18.0 ps^{-1} and a lower limit for Δm_s of 14.9 ps^{-1} at 95% CL is obtained [94].

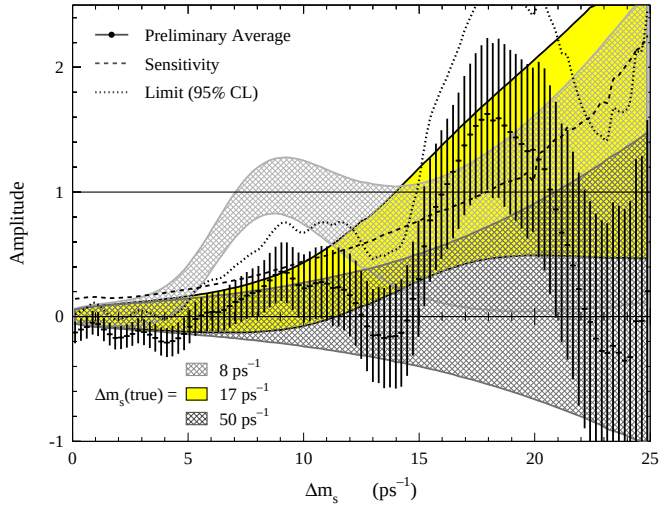

Fig. 5. Preliminary average of the experimental amplitude spectrum using results from ALEPH, DELPHI, OPAL, SLD and CDF [94]. The dashed and dotted lines give the sensitivity ($1.645 \times \sigma_A$) and the 95% limit $A + 1.645 \times \sigma_A$), respectively. Shown in addition are the distributions obtained from a toy Monte Carlo simulation using as true values Δm_s = 8, 17, 50 ps^{-1}

The information from $B_s^0 - \bar{B}_s^0$ oscillations is usually implemented into χ^2 fits using [92] $\chi^2_{|1-A|} = ((1-A)/\sigma_A)^2$ and $\text{CL}(\chi^2_{|1-A|}) = \text{Erfc}(|1-A|/\sigma_A/\sqrt{2}))$. However, this procedure does not properly interpret the information of the amplitude spectrum. For instance, two measured amplitudes A_1 and A_2, where $A_1 > 1$ and $A_2 < 1$ but $A_1 - 1 = 1 - A_2$, result in the same confidence level in this approach although it would be natural to assign a larger likelihood for an oscillation to A_1 than to A_2. We propose an alternative procedure which exploits the information from the sign of $1 - A$ by suppressing the module in the above definition of $\chi^2_{|1-A|}$:

$$\chi^2_{1-A} = 2 \cdot \left[\text{Erfc}^{-1}\left(\frac{1}{2}\,\text{Erfc}\left(\frac{1-A}{\sqrt{2}\sigma_A}\right)\right)\right]^2 \ . \quad (71)$$

It has been pointed out that the maximum information from the fit to the proper time distributions of mixed and unmixed $B_s^0(\bar{B}_s^0)$ decays is obtained from the ratio

[23] The consistent, though preliminary measurements of BA BAR[81] and Belle [82], were not considered in the average quoted in Table 1

of the likelihood at given frequency Δm_s, $\mathcal{L}(\Delta m_s)$, to the likelihood at infinity, $\mathcal{L}(\Delta m_s = \infty)$ [92,95,6]. The logarithm of this ratio reads

$$2\Delta\ln\mathcal{L}^{\infty}(\Delta m_s) = \frac{(1-A)^2}{\sigma_A^2} - \frac{A^2}{\sigma_A^2} \ , \qquad (72)$$

which is assigned *via* (23) to χ^2_∞. The behaviour of the above defined χ^2 and likelihood functions versus Δm_s for the measured amplitude spectrum is plotted in Fig. 6. The dashed line shows the drawback of using $\chi^2_{|1-A|}$: the strongest signal yield is obtained at the crossings $A = 1$ and not at the maximum amplitude situated around 17 ps^{-1}. This drawback is cured for χ^2_{1-A} (solid line). The dotted line shows the ratio χ^2_∞, providing a significantly stronger constraint. For the current analysis we decided not to use the likelihood ratio since the validity of the normalisation of the likelihood which allows to identify $\mathcal{L}^\infty$ with a probability density function is questionable. This problem could be circumvented by means of a realistic Monte Carlo simulation which permits the conversion of likelihoods to confidence levels which however is currently not available.

– **sin2β** The first measurements for sin2β in B decays to CP eigenstates containing charmonium from the B factories, CDF and LEP give results which are compatible with both, the SM expectation and zero:

$$\sin 2\beta = \begin{cases} 0.34 \pm 0.21 \ \text{(BABAR [96])} \\ 0.58 \pm 0.34 \ \text{(Belle [97])} \\ 0.79 \pm 0.43 \ \text{(CDF [98])} \\ 0.84 \pm 0.93 \ \text{(ALEPH [99])} \\ 3.2 \pm 2.0 \ \text{(OPAL [100])} \end{cases} \qquad (73)$$

From these measurements (asymmetric errors have been averaged) we obtain the weighted mean $\sin 2\beta_{\rm WA} = 0.48 \pm 0.16$, where the (small) effect from using different Δm_d values in the single analyses has not been taken into account.

6.3 Future prospects: rare decays of K and B-mesons

Theoretically clean measurements of CKM matrix elements are obtained by virtue of rare K and B decays. The countours in the $(\bar\rho, \bar\eta)$ plane expected from rare K decays are drawned on Fig.2.

– The decay $\boldsymbol{K^0_{\rm L} \to \pi^0\nu\bar\nu}$ has not been observed yet. The current upper limit reads [10]

$$\mathrm{BR}(K^0_{\rm L} \to \pi^0\nu\bar\nu) < 5.9 \times 10^{-7} \quad (\mathrm{CL} = 90\%) \ , \qquad (74)$$

while the expected SM branching ratio is of the order of 2×10^{-11}. The decay proceeds *via* a loop induced FCNC transition at short distance and is greatly dominated by a direct CP-violating amplitude in the SM [102], $A(K^0_{\rm L} \to \pi^0\nu\bar\nu) \propto \mathrm{Im}[V_{td}V^*_{ts}]\langle\pi^0|(\bar s d)_{V-A}|K^0\rangle$, due to the cancellation of the charm contributions. The

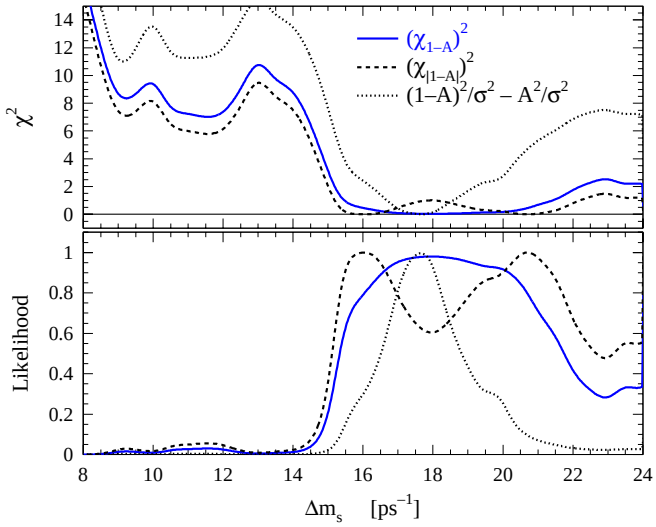

Fig. 6. The measured χ^2 (upper plot) and likelihood (lower plot) functions (defined in the text) versus the frequency of $B^0_s - \bar B^0_s$ oscillation, Δm_s

SM prediction for the branching fraction of the decay reads [102,103]

$$\begin{aligned}&\mathrm{BR}(K^0_{\rm L} \to \pi^0\nu\bar\nu) \\ &= r_{K_{\rm L}} \frac{\tau_{K_{\rm L}}}{\tau_{K^+}} \frac{3\alpha^2}{2\pi^2} \frac{\mathrm{BR}(K^+ \to \pi^0 e^+\nu)}{|V_{ud}V^2_{us}|^2 \sin^4\theta_{\rm W}} \\ &\quad \times \left(\eta_X X_0(x_t) \mathrm{Im}\,[V^*_{ts}V_{td}V^*_{us}V_{ud}]\right)^2 . \end{aligned} \qquad (75)$$

Here, $r_{K_{\rm L}} = 0.944$ corrects for isospin breaking effects and the different phase space [104] involved in the relation between the $K^0_{\rm L}$ and the K^+ branching fractions. The other parameters in (75) are the kaon lifetimes, the QED running fine structure constant and the Weinberg angle. The Inami-Lim function $X(x_t)$ for $x_t = (m_t/m_W)^2$ is defined as

$$X_0(x) = \frac{x}{8}\left(\frac{x+2}{x-1} + \frac{3x-6}{(x-1)^2}\ln(x)\right) \qquad (76)$$

for which $\eta_X = 0.994$ ((75)) represents the NLO correction [103]. In [105] the CP conserving contribution to $K^0_{\rm L} \to \pi^0\nu\bar\nu$ has been found to be suppressed by a factor of 6×10^{-5} with respect to the CP-violating rate. Expressed in the Wolfenstein parameterization, the SM prediction corresponds to

$$\mathrm{BR}(K^0_{\rm L} \to \pi^0\nu\bar\nu) \propto \lambda^8 A^4 \bar\eta^2 \ , \qquad (77)$$

showing that a measurement would provide a pair of horizontal lines in the $(\bar\rho, \bar\eta)$ plane. Proposed experiments that could measure $\mathrm{BR}(K^0_{\rm L} \to \pi^0\nu\bar\nu)$ are the KOPIO experiment at BNL [106] and KAMI at FNAL [107] expecting 60 and 120 events, respectively. The experiment(s) will not start before 2005 and have to take data for several years. The expected precision on the branching ratio in 2010 is of the order 5% to 10%,

thereby yielding a precision of a few percent on the measurement of $\bar{\eta}$.

– $\boldsymbol{K^+ \to \pi^+\nu\bar{\nu}}$ **:** The SM prediction of the rare decay $K^+ \to \pi^+\nu\bar{\nu}$ is given by [103]

$$\begin{aligned}&\mathrm{BR}(K^+ \to \pi^+\nu\bar{\nu})\\&= r_{K^+}\frac{3\alpha^2}{2\pi^2}\frac{\mathrm{BR}(K^+ \to \pi^0 e^+\nu)}{|V_{us}|^2\sin^4\theta_\mathrm{W}}\\&\quad\times\sum_{i=e,\mu,\tau}\left|\eta_X X_0(x_t)V_{td}V_{ts}^* + X_\mathrm{NL}^{(i)}V_{cd}V_{cs}^*\right|^2 \quad (78)\end{aligned}$$

where the function $X_0(x_t)$ is given by (76) and where the $X_\mathrm{NL}^{(\ell)}$ terms account for the charm contributions (not suppressed here) and are calculated in [108]. The values depend on the QCD scale $\Lambda_{\overline{\mathrm{MS}}}^{(4)}$ and the running charm quark mass: $X_\mathrm{NL}^{(e)} = (8-13)\times 10^{-4}$ and $X_\mathrm{NL}^{(\tau)} = (5-9)\times 10^{-4}$. A detailed discussion of the theoretical uncertainties connected with this and the above SM predictions is provided in [103]. The BNL experiment E787 has observed one event resulting in $\mathrm{BR}(K^+ \to \pi^+\nu\bar{\nu}) = (1.5^{+3.4}_{-1.2})\times 10^{-10}$ [109]. About 5 - 10 events are expected to be observed by the successor E949 [110] whereas the CKM project at FNAL [111], starting about 2005, expects to collect about 100 events within some years of data taking. A similar precision for the branching ratio measurement as in the case of $K_\mathrm{L}^0 \to \pi^0\nu\bar{\nu}$ may be achieved in the year 2010. A theoretical uncertainty of the order 5% but likely not well below this value might be possible [103].

– The $\boldsymbol{B^+ \to \tau^+\nu}$ decay has not been observed yet. The current upper limit for its branching fraction reads [10]

$$\mathrm{BR}(B^+ \to \tau^+\nu) < 5.7\times 10^{-4} \quad (\mathrm{CL} = 90\%)\ . \quad (79)$$

In the SM the branching ratio is given by

$$\begin{aligned}\mathrm{BR}(B^+ \to \tau^+\nu) = &\frac{G_\mathrm{F}^2 m_B m_\tau^2}{8\pi}\left(1-\frac{m_\tau^2}{m_B^2}\right)\\&\times f_{B_d}^2|V_{ub}|^2\tau_B\ , \quad (80)\end{aligned}$$

with the B meson decay constant f_{B_d} (see Table 1) and the lifetime of the charged B, $\tau_B = 1.653\pm 0.028$ ps [10]. Depending on the precision of the lattice calculation of f_{B_d}, a measurement of $\mathrm{BR}(B^+ \to \tau^+\nu)$ may either yield a direct measurement of $|V_{ub}|$, or may improve the prediction of Δm_d through the constraint of f_{B_d} which is the more likely way to proceed. Additional information may be obtained by measuring the radiative decay $B^+ \to \ell^+\nu_\ell\gamma$ in which the helicity suppression is circumvented due to the emission of the photon from the primary u-quark (see, e.g., [112]). Although the calculation of the branching ratio is model dependent a measurement possibly provides a useful experimental check of lattice calculations.

– CP-violating **Partial Rate Asymmetries** (PRA) of inclusive $b \to s(d)\gamma$ decays can be calculated in the SM [113–117]. They are defined by the ratio

$$A_\mathrm{CP}^{b\to s(d)\gamma} = \frac{\mathrm{BR}(\bar{B} \to X_{s(d)}\gamma) - \mathrm{BR}(\bar{B} \to X_{\bar{s}(\bar{d})}\gamma)}{\mathrm{BR}(\bar{B} \to X_{s(d)}\gamma) + \mathrm{BR}(\bar{B} \to X_{\bar{s}(\bar{d})}\gamma)} \quad (81)$$

Their theoretical predictions depend on various Wilson coefficients and CKM matrix elements involving the CP-violating phase [115]. Lumping all coefficients together, where external parameters like the strong coupling constant, the b-quark mass, the photon infrared cut-off and the renormalization scale have to be fixed, gives the estimate [117]

$$A_\mathrm{CP}^{b\to s(d)\gamma} \approx 0.33\times\mathrm{Im}\left[\frac{V_{ub}V_{us(d)}^*}{V_{tb}V_{ts(d)}^*}\right]\ . \quad (82)$$

This yields asymmetries of $A_\mathrm{CP}^{b\to s\gamma} \approx 0.6\%$ and $A_\mathrm{CP}^{b\to d\gamma} \approx -16\%$ for some typical values of the CKM elements.

– $\boldsymbol{B \to \pi K}$ **:** Due to the work of many authors (see, *e.g.*, [118–122] - this list is far from being complete), it could be shown that the (ratios of) branching fractions of charmless B_d decays into π and K final states provide constraints on the UT angle γ. Most recent branching ratios read

$$\begin{aligned}&\mathrm{BR}(B_d^0 \to \pi^+\pi^-) + \mathrm{BR}(\bar{B}_d^0 \to \pi^+\pi^-)\\&\quad = (4.43\pm 0.89)\times 10^{-6}\ , \quad (83)\\&\mathrm{BR}(B_d^0 \to K^+\pi^-) + \mathrm{BR}(\bar{B}_d^0 \to K^-\pi^+)\\&\quad = (17.25\pm 1.55)\times 10^{-6}\ , \quad (84)\\&\mathrm{BR}(B^+ \to K^+\pi^0) + \mathrm{BR}(B^- \to K^-\pi^0)\\&\quad = (12.10\pm 1.70)\times 10^{-6}\ , \quad (85)\\&\mathrm{BR}(B^+ \to K^0\pi^+) + \mathrm{BR}(B^- \to \bar{K}^0\pi^-)\\&\quad = (17.19\pm 2.54)\times 10^{-6}\ , \quad (86)\\&\mathrm{BR}(B_d^0 \to K^0\pi^0) + \mathrm{BR}(\bar{B}_d^0 \to \bar{K}^0\pi^0)\\&\quad = (10.33\pm 2.53)\times 10^{-6}\ , \quad (87)\end{aligned}$$

where the values given are the weighted means of the *preliminary* results on charmless B decays presented by the BABAR, Belle and CLEO collaborations [123–125] (asymmetric errors have been averaged). The authors of [120,121] have obtained predictions of relative amplitudes and phases of the tree and penguin diagrams involved in the above decays. Very recently, a new theoretical analysis of two-body B decays to pions and kaons, based on non-leading Factorization Approximation, has been published [126]. The authors obtain an allowed region for $|V_{ub}|e^{-i\gamma}$ which is in agreement with the results found in this work.

A statistical discussion and formulae for the treatment of ratios of branching fractions or, more precisely, constraints from parameters with arbitrary absolute, but known relative normalization is given in Appendix C of this paper.

We have attempted in this section to recall some of the most striking prospects for future CKM constraints which, however, is far from being complete. A more quantitative

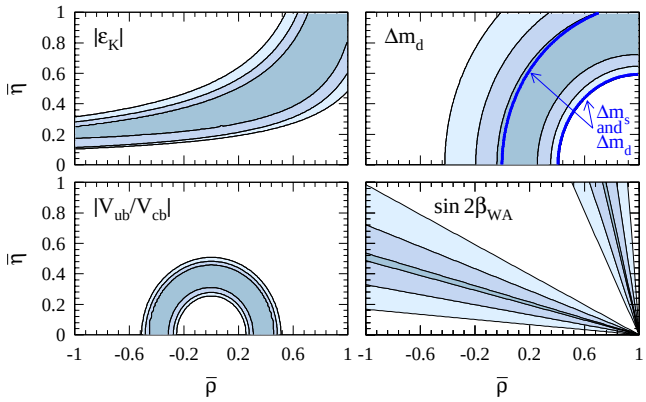

Fig. 7. Confidence levels in the $(\bar{\rho}, \bar{\eta})$ plane for the individual constraints. The upper right hand plot shows in addition to Δm_d the improved constraint from Δm_s *via* ξ on Δm_d. The shaded areas indicate the regions of $\geq 90\%$, $\geq 32\%$ and $\geq 5\%$ CLs, respectively

and broader selection of CKM sensitive quantities as well as extrapolations into the future can be found in [14]. If Δm_s is not much larger than suggested by the current SM constraints it will likely be measured during the forthcoming Tevatron II run. The precision of the combined Δm_d and Δm_s constraints on $\bar{\rho}$, $\bar{\eta}$ will then be dominated by the QCD parameter ξ. An experimental determination of the decay constant ratio f_{D_s}/f_{D_d} at a τ/charm factory would be helpful to check lattice QCD calculations of ξ. Measurements of time-dependent CP-asymmetries at the $\Upsilon(4S)$ and at hadron machines, in particular at the forthcoming experiments LHCb and BTeV, aim to extract the UT angle α in non-strange, charmless two and three body decays. The remaining angle γ is expected to be determined in $B_d \to DK$ and $B_s \to D_s K$ decays, though these measurements require very large data samples of the corresponding $B_{d(s)}$ mesons. The dedicated experiments LHCb and BTeV will also measure the most promising channel $B_s \to \psi\phi$.

7 Constrained fits within the SM

After the discussions in the preceding sections, we are prepared to perform the constrained fits of the CKM parameters and related quantities. We place ourselves in the framework of the *R*fit scheme (see Sect. 3.1.2 for an introduction and Sect. 4.1 for a summary of its main features) and hence define the theoretical likelihoods of (32) to be one within the allowed ranges and zero outside[24]. As a consequence, no hierarchy is introduced for any permitted set of theoretical parameters, *i.e.*, the χ^2 which is minimized in the fit receives no contribution from theoretical systematics, but theoretical paramaters cannot exit their allowed ranges. When relevant, statistical and theoretical uncertainties are combined beforehand, as presented in Sect. 5.1 (*i.e.*, applying (39)). Floating theoretical parameters are labelled by an asterisk in the "Float." column of Table 1. For parameters with small uncertainties, errors are propagated through the theoretical predictions, and added in quadrature to the experimental error of the corresponding measurements[25]: they are labelled by an asterisk in the "Prop." column of Table 1.

7.1 Two dimensional parameter spaces

It is customary to present the constraints on the CP-violating phase in the two-dimensional $(\bar{\rho}, \bar{\eta})$ plane of the Wolfenstein parameterization. Other representations involving the UT angles α, β and γ are also considered in the analysis. For the two-dimensional graphical displays we define the a parameter space by the coordinates $a = \{x, y\}$ (*e.g.*, $a = \{\bar{\rho}, \bar{\eta}\}$) and the μ space by the other CKM parameters λ and A, as well as the $y_{\rm QCD}$ parameters.

7.1.1 Metrology in the $(\bar{\rho}, \bar{\eta})$ plane

The individual constraints, sensitive to $\bar{\rho}$ and $\bar{\eta}$, are drawn in Fig. 7. Shown are the CLs of (40) which, according to the frequentist approach adopted in *R*fit, have to be interpreted as *upper bounds for the optimal set of theoretical parameters at a given point in the* $(\bar{\rho}, \bar{\eta})$ *plane* (this is implicit in the following when invoking the term CL). Obviously, CLs should not be interpreted as PDFs, *i.e.*, inferring equal relative probability density from equal shades. Instead, a CL value expresses a probability which is defined for a given coordinate $\{\bar{\rho}, \bar{\eta}\}$: it is the probability that the agreement between data and the most favorable realization of the SM at that point be worse than the one observed. However, although the CLs have a well defined statistical meaning, one must be aware of their strong dependence on the, to some extent, arbitrary $[y_{\rm QCD}]$ ranges. (*c.f.*, Sect. 3.2).

The results of the global fit in the $(\bar{\rho}, \bar{\eta})$ plane are shown in Fig. 8 not including (upper plot) and including (lower plot) in the fit the world average of $\sin 2\beta$ (see Table 1). The dark, medium and light shaded areas correspond to $\geq 90\%$, $\geq 32\%$ and $\geq 5\%$ CL, respectively. The outer regions with lower probabilities are outside the $y^{\rm opt}_{\rm mod}$ domain where an adjustment of the μ parameters can maintain maximal agreement (*i.e.*, can reproduce the $\chi^2_{{\rm min};y_{\rm mod}}$ value). Also shown are the 5% CL contours of the individual constraints as well as the $\geq 32\%$ and $\geq 5\%$ CL regions corresponding to $\sin 2\beta_{\rm WA}$ (hatched areas). As described in Sect. 3.2, the CLs obtained belong to the metrological phase of the analysis and, by construction, do not constitute a test of goodness of the theory. A probe of the SM is obtained from the numerical value of $\chi^2_{{\rm min};y_{\rm mod}}$ as discussed in Sect. 3.3 and used in Sect. 7.3.

[24] In other words, we use $\kappa = 0$ and $\zeta = 1$ for the Hat function $\mathcal{L}_{\rm syst}(x_0)$ of (61)

[25] This procedure neglects the correlations occurring when such parameters are used in more than one theoretical prediction

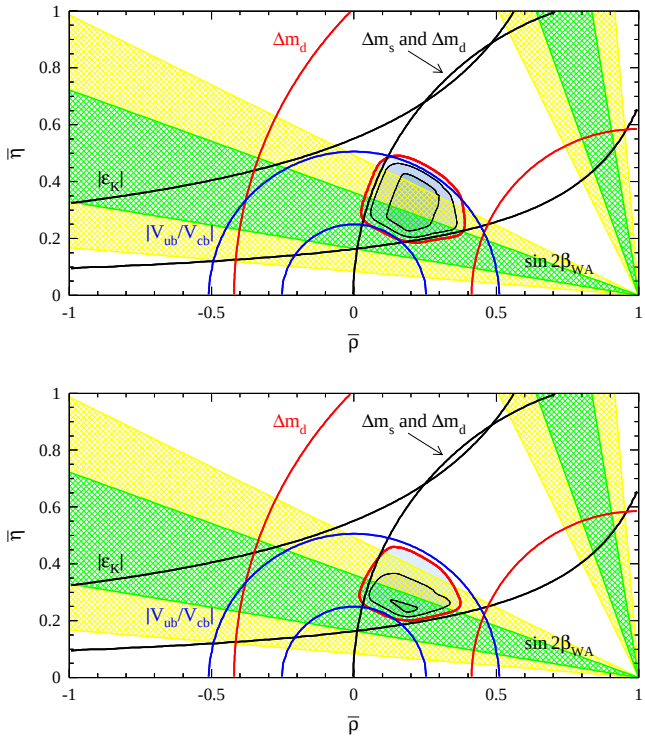

Fig. 8. *Upper plot:* confidence levels in the $(\bar{\rho}, \bar{\eta})$ plane for the global CKM fit. The shaded areas indicate the regions of $\geq 90\%$, $\geq 32\%$ and $\geq 5\%$ CLs, respectively. Also shown are the 5% CL contours of the individual constraints. The $\geq 32\%$ and $\geq 5\%$ CL constraints from the world average of the $\sin 2\beta$ measurements, not entering the combined fit, are depicted by the dashed areas. See Table 1 for a compendium of the fit input values. *Lower plot:* confidence levels obtained when including the world average of $\sin 2\beta$ in the combined fit

7.1.2 Other two dimensional parameter spaces

Except for possible multi-valuedness problems[26], it is straightforward to replace the $(\bar{\rho}, \bar{\eta})$ plane by any other one, two or higher dimensional parameter constellation. Figures 9 and 10 show the results from the global fits and for the individual constraints in the planes $(\sin 2\alpha, \sin 2\beta)$, $(\sin 2\alpha, \gamma)$ and $(\sin 2\beta, \gamma)$, respectively. The constraint from $\sin 2\beta$ does not enter the fits. As aforementioned, the individual constraints are given as 5% CL contours, and the shaded areas depict $\geq 90\%$, $\geq 32\%$ and $\geq 5\%$ CL areas. In general, for the $(\bar{\rho}, \bar{\eta})$ plane as well as for any parameter spaces, the individual inputs are less constraining than what they yield in the combined fit: *i.e.*, a combination of input variables can lead to a suppression of solutions thus enhancing the individual constraints. An example for this is drawn in the lower plot of Fig. 10: the

[26] For example, when exploring the $(\sin 2\alpha, \sin 2\beta)$ plane, care should be taken to account for multiple solutions. A given value of $\sin 2\omega$ ($\omega = \alpha$ or β) can be obtained with four values of ω ($\omega_1 = \frac{1}{2}\arcsin(\sin 2\omega)$, $\omega_2 = \frac{\pi}{2} - \omega_1$, $\omega_3 = \pi + \omega_1$, $\omega_4 = \frac{3\pi}{2} - \omega_1$) and corresponds to a pair of curves in the $(\bar{\rho}, \bar{\eta})$ plane intersecting on the $\bar{\eta} = 0$ axis. Each intersection of one of the $\sin 2\beta$ curves with one of the $\sin 2\alpha$ curves should be considered

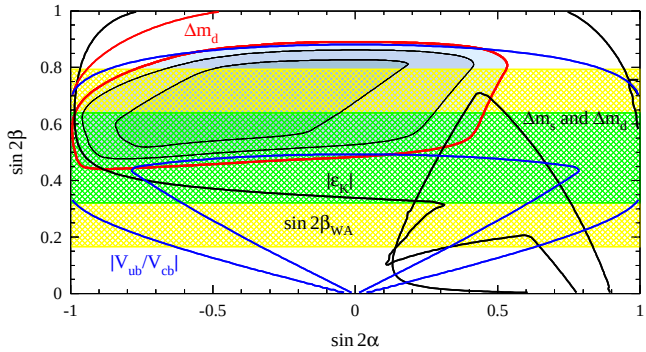

Fig. 9. Confidence levels in the $(\sin 2\alpha, \sin 2\beta)$ plane for the global CKM fit. The shaded areas indicate the regions of $\geq 90\%$, $\geq 32\%$ and $\geq 5\%$ CLs, respectively. Also shown are the 5% CL contours of the individual constraints. The $\geq 32\%$ and $\geq 5\%$ CL constraints from the world average of the $\sin 2\beta$ measurements, not entering the global fit, are given by the dashed areas

individual contribution of $|\epsilon_K|$ is stronger in combination with $|V_{ub}/V_{cb}|$ (indicated by the arrows). The lower plot of Fig. 10 visualises the complementarity between $|V_{ub}/V_{cb}|$ and $|\epsilon_K|$, constraining $\sin 2\beta$, on one hand, and $\Delta m_s/\Delta m_d$ and $|\epsilon_K|$, constraining γ, on the other hand.

7.2 One dimensional parameter spaces

Following the line of the preceding sections we can derive one-dimensional constraints for all parameters involved, such as the various CKM parameters, the moduli of the CKM matrix elements, branching ratios of rare K and B meson decays as well as theoretical parameters. Consequently, we define the parameter we are interested in to be a and all others to be μ (*c.f.*, Sect. 3.2), and scan a. Numerical and graphical results are obtained for CKM fits not including (including) the world average value of $\sin 2\beta$ (see Table 1 for the input parameters).

As an example, Fig. 11 shows the CLs obtained for the Wolfenstein parameters, the UT angles and the Jarlskog parameter, without (solid line, gray area) and with (dotted line) including the world average $\sin 2\beta_{\rm WA}$ in the global fit. As in the two-dimensional case, the CLs shown correspond to the most compatible theory for a given point in a. Since the parameter λ is not significantly constrained by the other inputs, its CL corresponds to the error function for one degree of freedom. In contrast, parameters such as, *e.g.*, $\bar{\rho}$, $\bar{\eta}$, are constrained by observables whose SM predictions are dominated by systematic theoretical errors. The positions of the flanks of the CL functions are determined by the $[y_{\rm QCD}]$ ranges, whereas their sharp rises are determined by statistical errors. Therefore, one should not attribute an absolute meaning to the precise locations of the flanks: they are due to the *assumptions* made to define the $[y_{\rm QCD}]$ ranges. In particular, when the world average $\sin 2\beta_{\rm WA}$ is used, the CL function obtained (*c.f.*, Fig. 11) exhibits a triangular shape: whereas the fall off on the right hand side of the function is well defined,

Table 2. Fit results for the various CKM parameters, the CKM matrix elements, branching ratios of some rare K and B meson decays and theoretical quantities. Ranges are quoted for the quantities that are limited by systematic theoretical errors. The last three lines give the ranges obtained for chosen theoretical parameters when removing their respective bounds in the fit

Parameter	$\geq 32\%$ CL	half width	$\geq 5\%$ CL	half width
λ	0.2221 ± 0.0021		0.2221 ± 0.0041	
A	0.770 - 0.888	0.059	0.754 - 0.906	0.076
$\bar{\rho}$	0.08 - 0.35	0.14	0.04 - 0.38	0.17
$\bar{\eta}$	0.22 - 0.46	0.12	0.21 - 0.49	0.14
J (10^{-5})	2.0 - 3.5	0.8	1.9 - 3.7	0.9
$\sin 2\alpha$	-0.91 - 0.34	0.63	-0.96 - 0.49	0.73
$\sin 2\beta$	0.50 - 0.86	0.18	0.47 - 0.89	0.21
α	80° - 123°	22°	75° - 127°	26°
β	15.0° - 29.7°	7.4°	14.0° - 31.4°	8.7°
$\gamma = \delta$	37° - 75°	19°	34° - 82°	24°
$\sin\theta_{12}$	0.2221 ± 0.0021		0.2221 ± 0.0041	
$\sin\theta_{13}$ (10^{-3})	2.70 - 4.31	0.81	2.49 - 4.55	1.03
$\sin\theta_{23}$ (10^{-3})	38.4 - 43.2	2.4	37.9 - 43.6	2.8
$\lvert V_{ud}\rvert$	0.97504 ± 0.00049		0.97504 ± 0.00094	
$\lvert V_{us}\rvert$	0.2221 ± 0.0021		0.2221 ± 0.0042	
$\lvert V_{ub}\rvert$ (10^{-3})	2.70 - 4.31	0.81	2.49 - 4.55	1.03
$\lvert V_{cd}\rvert$	0.2220 ± 0.0021		0.2220 ± 0.0042	
$\lvert V_{cs}\rvert$	0.97422 ± 0.00056		0.97422 ± 0.00102	
$\lvert V_{cb}\rvert$ (10^{-3})	38.4 - 43.2	2.4	37.9 - 43.6	2.8
$\lvert V_{td}\rvert$ (10^{-3})	6.6 - 9.2	1.3	6.3 - 9.6	1.6
$\lvert V_{ts}\rvert$ (10^{-3})	37.7 - 42.8	2.6	37.3 - 43.2	3.0
$\lvert V_{tb}\rvert$	0.99907 - 0.99927	10×10^{-5}	0.99905 - 0.99929	12×10^{-5}
Δm_s (ps^{-1})	15.5 - 33.7	9.1	15.0 - 42.0	13.5
BR($K_L^0 \to \pi^0\nu\bar{\nu}$) ($10^{-11}$)	1.3 - 4.0	1.4	1.2 - 4.4	1.6
BR($K^+ \to \pi^+\nu\bar{\nu}$) ($10^{-11}$)	5.1 - 9.6	2.3	4.8 - 10.5	2.9
BR($B^+ \to \tau^+\nu_\tau$) (10^{-5})	4.6 - 20.0	7.7	3.6 - 23.6	10.0
BR($B^+ \to \mu^+\nu_\mu$) (10^{-7})	1.8 - 7.9	3.1	1.5 - 9.3	3.9
$f_{B_d}\sqrt{B_d}$ (MeV)	193 - 271	39	184 - 284	50
B_K	> 0.55		> 0.50	
m_t (GeV)	106 - 406	150	93 - 565	236

the location of the flank on the left hand side is somewhat arbitrary and hence arguable.

The results for all relevant parameters considered in this work are listed in Table 2, without $\sin 2\beta$ in the fit, and Table 3 when including $\sin 2\beta_{\rm WA}$. Given are the ranges for $\geq 32\%$ and $\geq 5\%$ CLs in the case of theoretically limited quantities and the corresponding Gaussian errors in the case of experimentally limited quantities. The 95% CL allowed ranges for the CKM matrix elements are similar to the ones quoted by the PDG. Numerical results involving $B \to \pi\pi/K\pi$ decays are not presented here[27].

[27] It proved however straightforward to implement them into the *R*fit scheme of the *CkmFitter* package (see also the discussion in Sect. 6.3)

7.2.1 Indirect evidence for CP violation

It is interesting to investigate the possibility of an indirect evidence for CP violation, from the measurements of non CP-violating observables. The dashed curves in Fig. 11 give the CLs which are obtained when using neither $\sin 2\beta$, nor $|\epsilon_K|$ in the fits. For $\bar{\eta} = 0$ (hence no CP violation) *R*fit yields CL $\simeq 50\%$. Therefore, we find that CP conservation cannot be excluded without $\sin 2\beta$ or $|\epsilon_K|$: a better knowledge of $|V_{ub}/V_{cb}|$ and Δm_d (Δm_s) is needed to draw any further conclusions. The large value of CL($\bar{\eta} = 0$) stems from the fact that the quoted CLs are upper bounds. There exist realizations of the SM, with $\bar{\eta} = 0$ and with all $y_{\rm QCD}$ parameters within their allowed [$y_{\rm QCD}$] ranges, which provide a perfectly acceptable description of data (without $|\epsilon_K|$ and $\sin 2\beta$). The realizations of the SM which yield the best agreement are chosen to compute CL($\bar{\eta} = 0$).

Table 3. Fit results including the world average on $\sin 2\beta_{\rm WA}$. As in Table 2, ranges are given for the quantities that are limited by systematic theoretical errors. The two right columns give the relative improvements (in percent) of the $\geq 32\%$ CL and $\geq 5\%$ CL half widths with respect to the fit results without $\sin 2\beta$ given in Table 2. The last three lines give the ranges obtained for chosen theoretical parameters when removing their respective bounds in the fit

Parameter	$\geq 32\%$ CL	half width	$\geq 5\%$ CL	half width	δ_{32}	δ_5
λ	0.2221 ± 0.0021		0.2221 ± 0.0041		0	0
A	0.782 - 0.888	0.053	0.758 - 0.906	0.074	10	3
$\bar{\rho}$	0.09 - 0.29	0.10	0.04 - 0.37	0.16	29	6
$\bar{\eta}$	0.22 - 0.32	0.05	0.21 - 0.42	0.11	58	21
J (10^{-5})	2.0 - 2.9	0.5	1.9 - 3.5	0.8	38	11
$\sin 2\alpha$	-0.88 - 0.04	0.46	-0.95 - 0.33	0.64	27	12
$\sin 2\beta$	0.50 - 0.67	0.09	0.47 - 0.81	0.17	50	19
α	89° - 121°	16°	80° - 126°	23°	27	12
β	15.0° - 21.0°	3.0°	14.0° - 27.0°	6.5°	59	25
$\gamma = \delta$	42° - 74°	16°	34° - 82°	24°	16	0
$\sin\theta_{12}$	0.2221 ± 0.0021		0.2221 ± 0.0041		0	0
$\sin\theta_{13}$ (10^{-3})	2.70 - 4.03	0.67	2.49 - 4.38	0.95	17	8
$\sin\theta_{23}$ (10^{-3})	38.4 - 43.2	2.4	38.0 - 43.6	2.8	0	0
$\lvert V_{ud}\rvert$	0.97504 ± 0.00049		0.97504 ± 0.00094		0	0
$\lvert V_{us}\rvert$	0.2221 ± 0.0021		0.2221 ± 0.0042		0	0
$\lvert V_{ub}\rvert$ (10^{-3})	2.70 - 3.71	0.51	2.45 - 4.38	0.96	37	7
$\lvert V_{cd}\rvert$	0.2220 ± 0.0021		0.2220 ± 0.0042		0	0
$\lvert V_{cs}\rvert$	0.97414 ± 0.00049		0.97414 ± 0.00098		13	4
$\lvert V_{cb}\rvert$ (10^{-3})	38.7 - 43.2	2.3	38.1 - 43.6	2.8	4	0
$\lvert V_{td}\rvert$ (10^{-3})	7.2 - 9.2	1.0	6.6 - 9.6	1.5	23	6
$\lvert V_{ts}\rvert$ (10^{-3})	38.0 - 42.7	2.4	37.4 - 43.1	2.9	8	3
$\lvert V_{tb}\rvert$	0.99907 - 0.99926	9×10^{-5}	0.99905 - 0.99928	11×10^{-5}	10	8
Δm_s (ps^{-1})	15.5 - 33.7	9.1	15.0 - 41.3	13.1	0	3
BR($K^0_{\rm L} \to \pi^0\nu\bar{\nu}$) ($10^{-11}$)	1.2 - 2.6	0.7	1.1 - 3.8	1.4	50	13
BR($K^+ \to \pi^+\nu\bar{\nu}$) ($10^{-11}$)	6.6 - 9.5	1.5	5.4 - 10.4	2.5	35	14
BR($B^+ \to \tau^+\nu_\tau$) (10^{-5})	4.6 - 12.4	3.9	3.6 - 21.0	8.7	49	13
BR($B^+ \to \mu^+\nu_\mu$) (10^{-7})	1.8 - 4.9	1.6	1.4 - 8.3	3.5	48	10
$f_{B_d}\sqrt{B_d}$ (MeV)	194 - 246	26	185 - 272	44	33	12
B_K	> 0.72		> 0.55		31	10
m_t (GeV)	124 - 406	141	102 - 550	224	6	5

This result is not in qualitative agreement with the one obtained in [6]: this illustrates how widely different conclusions can be reached depending on the choice made for the statistical treatment. The Bayesian approach, while computing CL($\bar{\eta} = 0$), is incorporating in passing, through the use of PDFs and (56), the "volume" of the domain in the $y_{\rm QCD}$ space (weighted by the theoretical PDFs) where realizations of the SM are in agreement with data. The "volume" of this $y_{\rm QCD}$ domain is small, as a result CL($\bar{\eta} = 0$) is small. The frequentist *R*fit scheme does not consider PDFs for $y_{\rm QCD}$ parameters as a valid concept, it thus cannot define "volume" of a domain in this space deprived of metric: only the best realizations are retained to define the CL.

An expanded view of the *R*fit CL($\bar{\eta}$) function is shown on the left hand side of Fig. 12 in the range $[-1, +1]$. Since no observable sensitive to CP violation is incorporated in the fit, CL($\bar{\eta}$) is an even function. The Bayesian PDF of $\bar{\eta}$ (*c.f.*, (56)) is shown on the right hand side of the same figure. The solid (resp. dashed) line is obtained using Gaussian (resp. uniform) PDFs for the $y_{\rm QCD}$ parameters. One observes that, independently of the definition used to derive a CL from the PDF, both Bayesian CLs will be low (at the percent level if one uses (57)) and most notably the one obtained from the uniform PDFs, although the inputs to the fit are identical to the ones used by *R*fit.

7.2.2 Impact of the new $\sin 2\beta$ measurements

The measurement of $\sin 2\beta$ provides the UT angle β up to a four-fold ambiguity. To illustrate this, we have enlarged

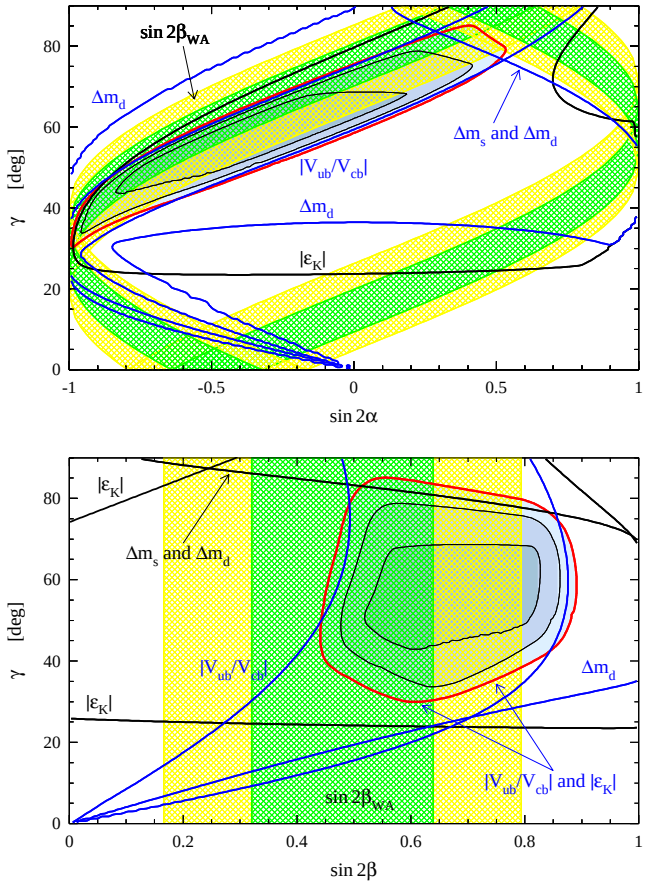

Fig. 10. Confidence levels in the (sin2α, γ) (upper plot) and the (sin2β, γ) plane (lower plot), obtained from the global CKM fit. The shaded areas indicate the regions of $\geq$ 90%, $\geq$ 32% and $\geq$ 5% CLs, respectively. Shown in addition are the 5% CL contours of the individual constraints. The $\geq$ 32% and $\geq$ 5% CL constraints from the world average of the sin2β measurements are given by the dashed areas. It does not enter the global fit

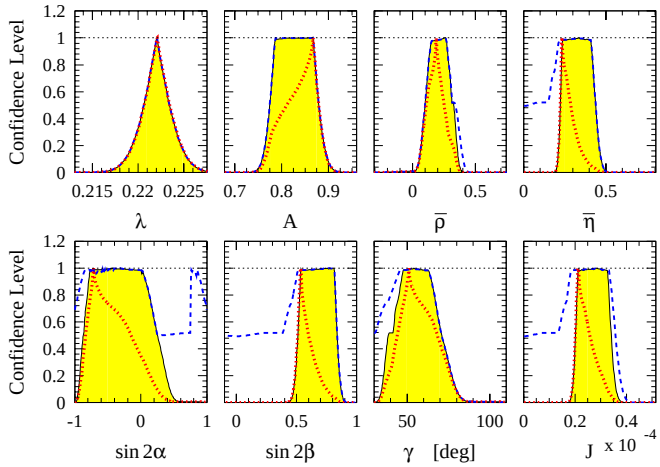

Fig. 11. Confidence levels for one-dimensional parameter fits of the CKM parameters, UT angles and the Jarlskog parameter, not including sin2β in the fits. The solid (dashed) lines give the results with (without) the CP-violating $|\epsilon_K|$ as fit input. The fits corresponding to the dotted lines include sin2$\beta_{\rm WA}$ and $|\epsilon_K|$

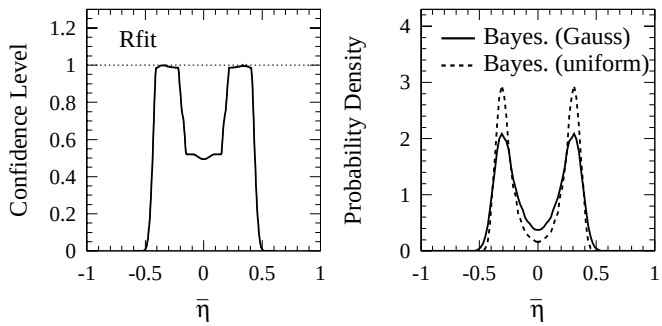

Fig. 12. The *R*fit CL($\bar{\eta}$) function (left hand side) and the Bayesian PDF of $\bar{\eta}$ (right hand side) obtained using the same inputs. The solid (resp. dashed) line is obtained using Gaussian (resp. uniform) PDFs for the $y_{\rm QCD}$ parameters

the borders of the ($\bar{\rho}$, $\bar{\eta}$) plane in Fig. 13. Shown are the individual constraints and the result from the global fit corresponding to Fig. 8, as well as the four solutions from the world average sin2$\beta_{\rm WA}$. It is a non-trivial outcome of the SM fit that it leads to an exclusion of three out of the four ambiguities.

The confidence levels for the sin2β measurements of BABAR[96] and Belle [97] together with the world average[28] and the result of *R*fit (without sin2β) are shown[29] in Fig. 14. Given in addition are the results of the integrated PDFs obtained in the Bayesian analysis, when ascribing Gaussian or uniform PDFs to the systematic theoretical errors. While they significantly differ in the precision they claim for, the *R*fit and Bayesian indirect determinations of sin2β are both compatible with the world average.

The last two columns of Table 3 give the relative improvements (in percent) of the parameter constraints gained by including sin2$\beta_{\rm WA}$ in the CKM fit: the two quoted numbers δ_{32} and δ_5 refer respectively to the ranges allowed at 32% and 5% CLs. All quantities sensitive to CP violation benefit from significantly smaller 32% allowed ranges, with a relative reduction of up to 50%. This reduction however gets suppressed when going to 5% CLs. This reduced improvement is explained by the fact that a

[28] As stated before, the measurements are assumed to be Gaussian distributed. Therefore, the CLs given are direct confidence levels and not upper bounds, as one obtains when theoretical systematics contribute significantly to the uncertainty of a quantity

[29] The plateau of the CL function obtained from *R*fit corresponds to sin2β values belonging to the $y_{\rm mod}^{\rm opt}$ domain. The CL on the plateau is not exactly equal to unity. The slight slope which is observed is due to the $|V_{cd}|$ input: being a function of $\bar{\rho}$ and $\bar{\eta}$, and having a statistically dominated uncertainty, $|V_{cd}|$ lifts by a very slight amount the degeneracy discussed in Sect. 3.2

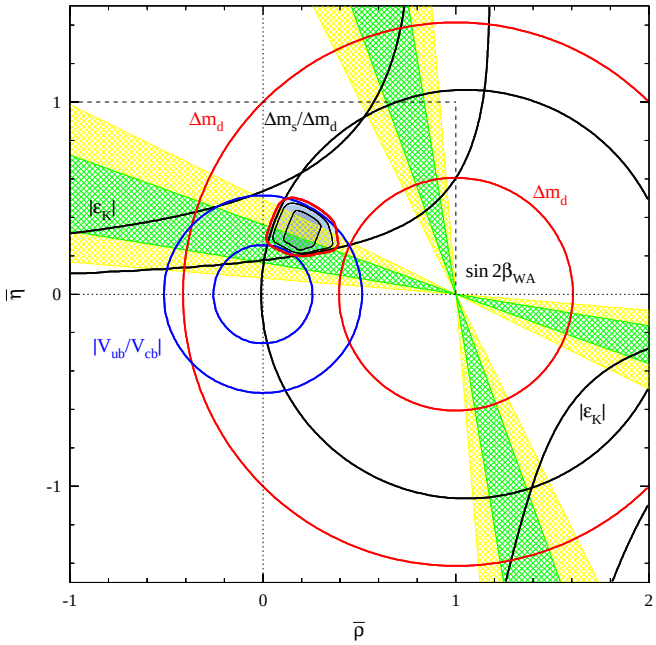

Fig. 13. Confidence levels in the enlarged $(\bar{\rho}, \bar{\eta})$ plane for the global CKM fit. See Fig. 8 for a description of the curves shown. The $\geq 32\%$ and $\geq 5\%$ CL constraints from the world average of the $\sin 2\beta$ measurements, not entering the combined fit, are depicted by the dashed areas. All four ambiguities are drawn, three of which are excluded by the Standard Model

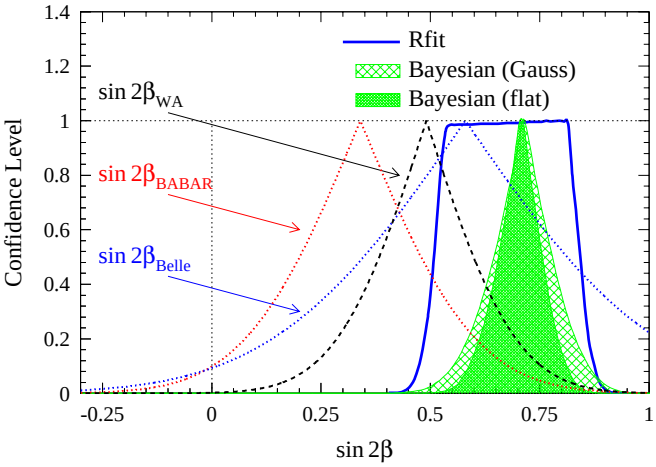

Fig. 14. Confidence levels of the recent $\sin 2\beta$ measurements of *BABAR*[96] and Belle [97] together with the world average. These CLs are compared to the indirect SM constraints obtained from *R*fit in this work (solid line). Also shown are the results from the integrated PDFs obtained in the Bayesian analysis using Gaussian and uniform PDFs for the systematic theoretical errors

significant fraction of 5% CL limits remain determined by the border of the $y_{\rm mod}^{\rm opt}$ domain, as can be seen on Fig. 11. Although the Gaussian shape of the additional constraint from $\sin 2\beta_{\rm WA}$ leads to significant structures within the allowed ranges of CLs, the $[y_{\rm QCD}]$ ranges still determine part of the CL function tails.

7.2.3 Numerical comparison with Bayesian results

As discussed previously, the Bayesian treatment identifies experimental and theoretical likelihoods as probability density functions which are folded according to (56) (*c.f.*, Sect. 4.4, Appendix A, and also [6,7]). In practice, the convolution integrals are solved within the *CkmFitter* package using Monte Carlo techniques generating some 10^8 test samples. All input quantities fluctuate according to Gaussian distributions, using their statistical experimental uncertainties $\sigma_{\rm exp}$, and to the PDFs ascribed to systematics, characterized by their systematical uncertainties $\sigma_{\rm syst}$. For the sake of simplicity, two choices are discussed here for the latter PDFs: all are taken to be Gaussians with standard deviations identified to $\sigma_{\rm syst}$, or all are taken to be uniform distributions with half-widths identified to $\sigma_{\rm syst}$. In the latter case, one does not identify the half-widths to $\sqrt{3}\sigma_{\rm syst}$ in order to use PDFs which (naively) would lead the Bayesian approach to yield results the closest to the one of the *R*fit scheme. Because of that, in the uniform case the RMS is smaller than in the Gaussian case: one thus expects to claim for significantly smaller uncertainties for the former choice than for the latter choice. A comparison of the results for both choices is given in Table 4. The central values quoted correspond to the mean values of the resulting PDFs, while the errors are their RMS (*i.e.*, when present, asymmetric errors have been averaged). The corresponding ranges provide a good approximation of the 68% confidence intervals which can be defined from an explicitly asymmetrical integration of the PDFs, since most of them closely resemble Gaussian distributions. As expected, the uncertainties are larger for the Gaussian choice. We observe a factor of about 2.2 (resp. 2.8) for the ratio between the $\geq 32\%$ CL intervals of *R*fit (Table 2) and the Bayesian ranges, for the Gaussian choice (resp. the uniform choice). For the Gaussian choice, this ratio reduces to about 1.3 for the $\geq 5\%$ CL.

Figure 15 provides a graphical comparison between the *R*fit result (the broad solid curve) and the Bayesian results (the dashed-dotted and dotted curves) on $\sin 2\beta$. Shown in addition is the result obtained when asymmetrically integrating the output PDF obtained from the Bayesian analysis of [6] (the narrow solid curve) where mostly uniform PDFs were chosen for the dominant theoretical uncertainties. The plot visualizes the tendency observed in Table 4 for the tails of lower CLs from *R*fit and the Bayesian approach to evolve towards comparable uncertainty ranges. However, the curves do not converge the ones to the others: the various treatments do not provide identical results, even for very low CLs.

7.3 Probing the standard model

We have seen in the introduction that the metrological phase is intrinsically unable to detect a failure of the SM to describe the data. The interpretation of the test statistics $\chi^2_{{\rm min};y_{\rm mod}}$ is performed by means of a toy Monte Carlo simulation as described in Sect. 3.3. The fits in the previ-

Table 4. Results for the CKM parameters, the UT angles and theoretical parameters, using Bayesian statistics with Gaussian (second column) or uniform (third column) distributed probability density functions for the systematic theoretical part of the input parameters. Note that asymmetric errors have been averaged. The constraint from $\sin 2\beta$ is not used. The fourth and fifth columns give the ratios of the half widths of the $\geq 32\%$ and $\geq 5\%$ CL *R*fit error intervals of Table 2, to the Bayesian errors given for the Gaussian case in the second column

Parameter	Gaussian	Uniform	ΔFreq./ΔBayes(Gauss) $\geq 32\%$ CL	$\geq 5\%$ CL
λ	0.2219 ± 0.0021	0.2219 ± 0.0021	1.0	1.0
A	0.832 ± 0.040	0.830 ± 0.028	1.5	1.0
$\bar{\rho}$	0.217 ± 0.063	0.203 ± 0.048	2.2	1.3
$\bar{\eta}$	0.331 ± 0.056	0.330 ± 0.039	2.1	1.3
J	$(2.70 \pm 0.36) \times 10^{-5}$	$(2.70 \pm 0.25) \times 10^{-5}$	2.2	1.3
$\sin 2\alpha$	-0.32 ± 0.30	-0.30 ± 0.24	2.1	1.2
$\sin 2\beta$	0.710 ± 0.093	0.705 ± 0.065	1.9	1.1
γ	$57.0^\circ \pm 8.7^\circ$	$58.5^\circ \pm 7.0^\circ$	2.2	1.4
$f_{B_d}\sqrt{B_d}$	(230 ± 27) MeV	(227 ± 13) MeV	1.4	0.9
B_K	0.91 ± 0.12	0.89 ± 0.08	-	-

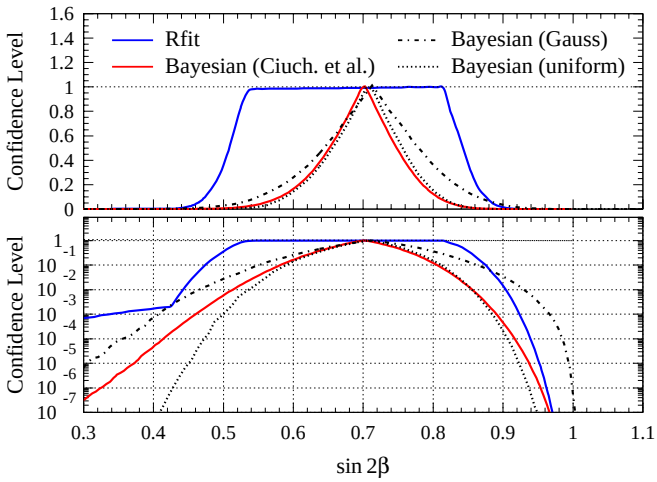

Fig. 15. Comparison between *R*fit (broad solid curve) and Bayesian fits for the indirect CKM constraint on $\sin 2\beta$. The lower plot displays the identical curves as in the upper plot but in logarithmic scale. For the Bayesian fits: Gaussian (uniform) systematic theoretical PDFs are depicted as dashed-dotted (dotted) curves. Shown in addition is the (integrated) result obtained in the Bayesian analysis of [6] (narrow solid curve) for which mostly uniform PDFs were chosen for the dominant theoretical uncertainties

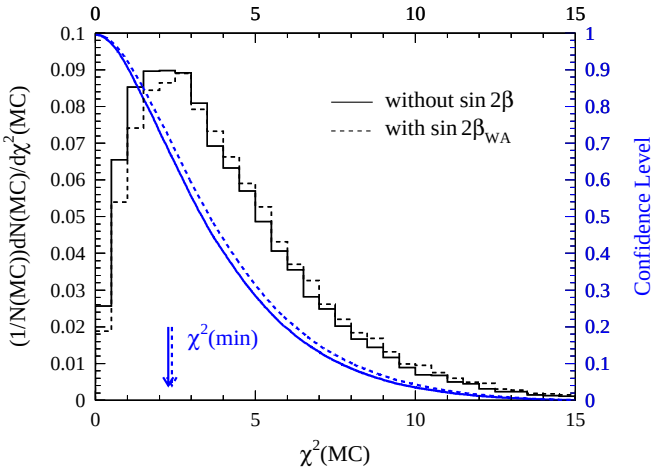

Fig. 16. Simulated $\mathcal{F}(\chi^2)$ distributions and corresponding CLs not including (solid lines) and including $\sin 2\beta_{\rm WA}$ (dashed lines) in the fit. Indicated by the arrows are the corresponding minimal $\chi^2_{{\rm min};y_{\rm mod}}$ found in the analyses

ous section yield for the point of best compatibility

$$\chi^2_{{\rm min};y_{\rm mod}} = 2.3\ (2.4)\ , \tag{88}$$

for the data set without (with) $\sin 2\beta_{\rm WA}$. We now generate the distribution $\mathcal{F}(\chi^2)$ of $\chi^2_{{\rm min};y_{\rm mod}}$ by fluctuating the measurements and parameters according to their non-theoretical errors around the theoretical values obtained using the parameter set $y^{\rm opt}_{\rm mod}$ for which is obtained $\chi^2_{{\rm min};y_{\rm mod}}$. The resulting toy distributions are shown by the solid (with $\sin 2\beta_{\rm WA}$ in the fit) and dashed (no $\sin 2\beta$) histograms in Fig. 16. Integrating the distributions according to (50) yields the corresponding CL (smooth curves in Fig.16). We find

$$\mathcal{P}({\rm SM}) \leq {\rm CL}(\chi^2_{{\rm min};y_{\rm mod}}) = 69\%\ (71\%)\ , \tag{89}$$

for the validity of the SM without (with) $\sin 2\beta_{\rm WA}$. Repeating the study using the $\sin 2\beta$ measurement of BABAR (Belle) instead of the world average, gives confidence levels of ${\rm CL}(\chi^2_{{\rm min};y_{\rm mod}}) = 59\%\ (77\%)$ for the validity of the SM.

8 Supersymmetric extensions of the standard model

Having considered both metrology and probing the SM with the present data set, one is now led to attempt an example analysis within an extended theoretical framework. This section aims to illustrate the search for specific new physics within a simple, predictive supersymmetric (SUSY) extension of the SM.

There exist a considerable number of SUSY models in which new phases appear in the coupling between supersymmetric and SM fields. However, these models remain unpredictive as long as the additional phases are unconstrained. We therefore cannot forecast how the shape of the UT is affected by the new fields. As a starting point, one can use restrictive assumptions which lead to more predictive models. In particular, one may only retain models which do not involve additional CP violating phases, so that flavour-changing processes are described by the same quark flavor mixing matrix V as in the SM. Supersymmetric contributions to the transitions between the down-type quarks ($b \to s$, $b \to d$, $s \to d$) are then proportional to the SM CKM matrix elements. This restriction defines the category of the so-called Minimal Flavour Violation (MFV) models which comprise some variants of the Minimal Supersymmetric Standard Model (MSSM), as well as the Two Higgs Doublet Models.

The MSSM has been extensively studied in the literature, and next-to-leading order (NLO) corrections to the SM have been calculated [127,128]. In this framework, the SUSY correction to neutral K and B meson mixing can be described by a single parameter which scales with the Inami-Lim function (66) of the top-quark loops in the box diagrams [129]:

$$S(x_t) \to S(x_t)(1+f) \ , \tag{90}$$

leading to the following modified expressions

$$\Delta m_d(\mathrm{MSSM}) = \Delta m_d(\mathrm{SM})\,[S(x_t) \to S(x_t)(1+f_d)] \tag{91}$$
$$\Delta m_s(\mathrm{MSSM}) = \Delta m_s(\mathrm{SM})\,[S(x_t) \to S(x_t)(1+f_s)] \tag{92}$$
$$|\epsilon_K|(\mathrm{MSSM}) = |\epsilon_K|(\mathrm{SM})\,[S(x_t) \to S(x_t)(1+f_\epsilon)] \ . \tag{93}$$

As pointed out in [131,129], the parameters f_d, f_s and f_ϵ belong to the same subprocesses so that the equality

$$f \equiv f_d = f_s = f_\epsilon \ , \tag{94}$$

holds in general. The numerical value for f is assumed to vary within the range $0 \le f \le 0.75$ [129], while other authors find broader ranges $0 \le f \le 1.13$ [130],depending on whether or not Supergravity constraints (*i.e.*, additional relations between masses and other terms of the MSSM Lagrangian) are applied to the MSSM.

The constraint from $|\epsilon_K|$ in the $\bar\rho - \bar\eta$ plane follows the form of a hyperbola [131]

$$\bar\eta \propto \frac{1}{(1-\bar\rho)S(x_t) + P_c} \ , \tag{95}$$

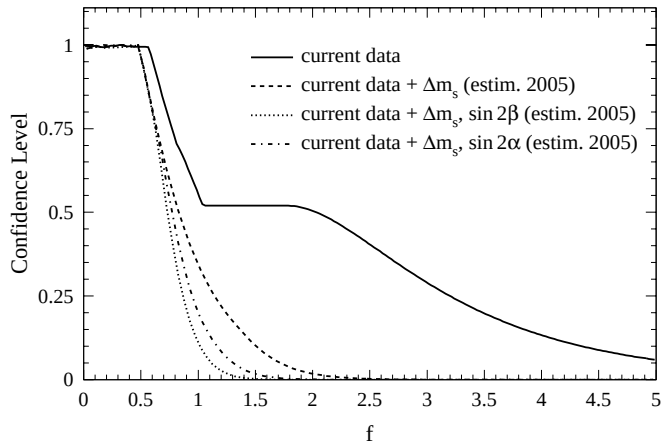

Fig. 17. Confidence level bounds of the global fit versus the MSSM offset f (solid line). The ledge at $f \approx 1$ starts at the upper frequency limit to which Δm_s amplitudes have been measured. It continues as a plateau due to the theoretically limited Δm_d information and eventually decreases statistically. Also shown are the estimated constraints from additional measurements of Δm_s (CDF, D0) (dashed line), sin2β (dotted line) and sin2α (dashed-dotted line) (both *BABAR*, Belle) for the year 2005 [14]

where $P_c = \eta_{ct} S(x_c, x_t) + \eta_{cc} S(x_c)$ stems from the charm loop contribution for which the SUSY contribution is expected to be small. The neutral B mass difference Δm_d measures the side R_t of the UT (18), where:

$$R_t \propto \frac{1}{\sqrt{S(x_t)}} \ , \tag{96}$$

so that SUSY will reduce R_t in case of positive f. Using the above formulae, one readily derives the dependence of sin2β on the SUSY parameter f [131]

$$\sin 2\beta = \frac{2\bar\eta(1-\bar\rho)}{R_t^2} \propto (\mathrm{const} - \bar\eta P_c) \ . \tag{97}$$

The first, constant term dominates by a factor of two to three the second term, while SUSY modifies the second term *via* the parameter $\bar\eta$ only. Note also that if the charm contribution to $K^0 - \bar K^0$ mixing were negligeable, SUSY effects would be totally absent in sin2β. It follows from this that γ constitutes the most sensitive UT angle to the SUSY contribution f.

8.1 Supersymmetric fits

The above SUSY parameterization has been included in *CkmFitter* and constrained fits are performed by setting $a = \{f\}$ and μ to all other parameters y_{mod}. The resulting confidence level bounds for the global CKM fit is shown in Fig. 17 from which one obtains for the present data set the upper limit

$$f \le 5.2 \quad (95\%\ \mathrm{CL}) \ . \tag{98}$$

This is in agreement with the findings of [131]. Also shown in Fig. 17 are the improved constraints from future measurements assumed to yield $\Delta m_s = (17.0 \pm 0.9)$ ps^{-1} (CDF, D0), $\sin 2\beta = 0.77 \pm 0.03$ and $\sin 2\alpha = -0.32 \pm 0.20$ (BABAR, Belle) (see also the more detailed discussion about future precision measurements in [14]).

9 Conclusions

We present a new approach to a global fit of the CKM matrix. It is denoted *R*fit and is based on frequentist statistics. We emphasize the thorough statistical definition of the method and discuss differences from Bayesian statistics [6,7] and from the 95% CL Scan method [5]. The choice for the fit input parameters, their values and errors are discussed to some detail; in cases of doubts we favor the more conservative estimates. The CKM analysis is formally subdivided into three distinct phases: a metrological phase in which the Standard Model is assumed to be valid and confidence levels for the parameters are computed; a probing phase addressing the issue of the validity of the Standard Model description of data; a probing phase for new physics relying on predictive parameterizations. For the first phase of the analysis, graphical results are displayed in several one and two dimensional representations and numerical results are given for relevant CKM parameterizations, CKM matrix elements, Standard Model predictions of rare K and B meson decays, and selected theoretical parameters. For the parameters related to the CP violating phase of the CKM matrix, we find for the different parameterizations (the fit includes the present world average of $\sin 2\beta$ measurements)

$$\begin{aligned} J &= (1.9 \ - \ 3.5) \times 10^{-5} \ , \\ \bar{\rho} &= 0.04 \ - \ 0.37 \ , \\ \bar{\eta} &= 0.21 \ - \ 0.42 \ , \\ \sin 2\alpha &= -0.95 \ - \ 0.33 \ , \\ \sin 2\beta &= 0.47 \ - \ 0.81 \ , \\ \alpha &= 80^\circ \ - \ 126^\circ \ , \\ \beta &= 14^\circ \ - \ 27^\circ \ , \\ \gamma = \delta &= 34^\circ \ - \ 82^\circ \ , \end{aligned}$$

where the $\geq 5\%$ confidence level ranges are quoted. The second phase of the analysis provides an upper bound for the validity of the Standard Model,

$$\mathcal{P}(\mathrm{SM}) \leq 71\% \ .$$

A simple predictive supersymmetric extension of the Standard Model has been studied in the third analysis phase.

Acknowledgements. We gratefully acknowledge the most interesting and helpful discussions with our colleagues from other active CKM analysis groups: F. Parodi, S. Plaszczynski, P. Roudeau, M.H. Schune and A. Stocchi, who have led pioneering analyses on this subject. We are indebted to the advice of D. Abbaneo, M. Artuso, C. Bernard, I. Bigi, G. Boix, A. Falk, A. El Khadra, A. Kronfeld, Z. Ligeti, D. London, G. Martinelli, M. Neubert, J. Ocariz, H. Quinn, A.I. Sanda, H. Wittig and many others. We thank our BABAR collaborators for the many discussions on this subject and especially acknowledge the very fruitful conversations with G. Dubois-Felsmann, G. Hamel de Monchenault, H.L. Lynch and K. Schubert. Special thanks to H. Lynch and K. Schubert for the careful reading of this manuscript and their thoughtful and most constructive comments. The work of HL was supported by the Fifth Framework Programme of the European Commission Community Research under the grant No. HPMF-CT-1999-00032.

Appendix

A Critical issues of the Bayesian approach

The Bayesian approach injects in the analysis pieces of information in the form of probability density distributions for the y_{QCD} parameters. The *R*fit scheme proposed in the present paper advocates a non-Bayesian approach because most theoretical uncertainties on the y_{QCD} parameters do not stem from statistical fluctuations. The y_{QCD} parameters are not random variables following probability density functions: there are poorly known, but fixed, parameters. The following examples serve the purpose to illustrate in a simplified framework the impact of using the Bayesian approach.

Let x_i denote $\mathcal{N}$ y_{QCD} parameters taking their values in identical allowed ranges $[x_i] = [-\Delta, +\Delta]$. These $\mathcal{N}$ y_{QCD} parameters are assumed to combine to form $T_{\mathrm{P}}^{(\mathcal{N})}$, the theoretical prediction for an observable, as follows[30]

$$T_{\mathrm{P}}^{(\mathcal{N})} = \prod_{i=1}^{\mathcal{N}} x_i \ . \tag{99}$$

The theoretical prediction $T_{\mathrm{P}}^{(\mathcal{N})}$ enters in the analysis as an unique y_{QCD} parameter, which

- within the *R*fit scheme, is characterized by its allowed range
$$[T_{\mathrm{P}}^{(\mathcal{N})}] = [-\Delta^{\mathcal{N}}, +\Delta^{\mathcal{N}}] \ , \tag{100}$$
- within the Bayesian approach, is characterized by its PDF
$$\rho(T) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{i=1}^{\mathcal{N}} \mathrm{d}x_i G(x_i) \delta(T - T_{\mathrm{P}}^{(\mathcal{N})}) \ , \tag{101}$$
where we assumed that identical PDFs, $G(x_i)$, were attributed to the x_i quantities.

[30] This is not an academic exercize: products of y_{QCD} terms are not rare in the theoretical predictions. For example, the Δm_d expression (*c.f.*, (70) of Sect. 6.2) involves the product of η_B, $f_{B_d}^2$ and B_d. However, the choice of $[x_i]$ ranges containing the origin is not met in practice. This choice is made here to highlight the difference between the *R*fit scheme and a PDF-based Bayesian approach. However, the singularity discussed below is present when Gaussian PDFs are used, as a matter of principle, whether or not $[x_i]$ ranges contains the origin

Independently of the details of the shape of $G(x)$, if this PDF is non-zero at the origin, $\rho(T)$ will exhibit a singularity at $T = 0$ with leading term $(-\ln|T|)^{(\mathcal{N}-1)}$. Hence, the larger $\mathcal{N}$ is, the more pronounced the peak at the origin of $\rho(T)$ is. If one chooses a uniform distribution $G(x) = 1/(2\Delta)$, that is to say the PDF the "closest" to the *R*fit assumptions (albeit being a fundamentally different object), one obtains a striking result. For $\mathcal{N} = 1$, the Bayesian approach and the *R*fit scheme are equivalent (though not for the associated CLs), insofar no other variable is involved: both state that the value of $T_{\mathrm{P}}^{(1)}$ is simply within the allowed range $[-\Delta, +\Delta]$. However, though originally no x_i values were favored, the Bayesian approach departs drastically from *R*fit as soon as $\mathcal{N} \geq 2$: it states that T_{P} is most likely close to zero. *In effect, when the number of* y_{QCD} *parameters entering the computation of the theoretical prediction increases, and hence when our knowledge of the corresponding observable decreases, the Bayesian approach claims the converse.*

This is less a consequence of the initial *ad hoc* choice for the PDF $G(x)$, than a consequence of the inescapable properties of (101), when applied to a product of terms. For instance, the peaking effect remains present for a sum of terms, but it is far less pronounced. The above generalities are illustrated, for $\mathcal{N} = 3$ and $\Delta = \sqrt{3}$, in Fig. 18. The choice $\mathcal{N} = 3$ is made because the resulting allowed ranges for the product $T_{\mathrm{P}}^{(3)} = x_1 x_2 x_3$ and for the sum $T_{\mathrm{S}}^{(3)} = x_1 + x_2 + x_3$ are identical when using *R*fit, namely $[-3\Delta, 3\Delta]$. The figure shows the Bayesian PDF of $T_{\mathrm{S}}^{(3)}$ (left hand side) and $T_{\mathrm{P}}^{(3)}$ (right hand side) assuming for $G(x)$ either a Gaussian distribution of standard deviation $\sigma = 1$ (the solid lines) or a uniform distribution in the range $[-\sqrt{3}, +\sqrt{3}]$ (the dashed lines) for the three parameters $x = x_i$, $i = 1, 2, 3$. The latter case is the closest to the *R*fit scheme for which the allowed range is $[T] = [-3\sqrt{3}, +3\sqrt{3}]$, indicated by the arrows in Fig. 18. For both, the sum T_{S} and the product T_{P}, the uniform and the Gaussian PDFs yield similar $\rho(T)$ distributions, because their RMS are chosen identical. The still noticeable difference between the two $G(x)$ PDFs is damped away by (101), even for such a moderate number like $\mathcal{N} = 3$. As a consequence, when following the Bayesian approach, it is not a particularly conservative choice to adopt a uniform PDF instead of a Gaussian PDF.

It is a remarkable difference between the two treatments that, whereas within the *R*fit scheme the two theoretical predictions T_{P} and T_{S} are genuinely indistinguishable, they yield sharply different PDFs in the Bayesian approach. This shows that, even more than the *ad hoc* choice made to ascribe PDF to the y_{QCD} parameters, it is the functional dependence of the theoretical predictions in these parameters, and the interplay between the various theoretical predictions within the complete CKM analysis, which play the central role. The attractiveness of the Bayesian approach, offering a straightforward procedure for analyses, is fallacious. The deep implication of ascribing PDFs to non-random variables is hidden inside an apparently innocuous convolution, the outcome of which reflects, more than anything else, the mathematical structure of the problem at hand.

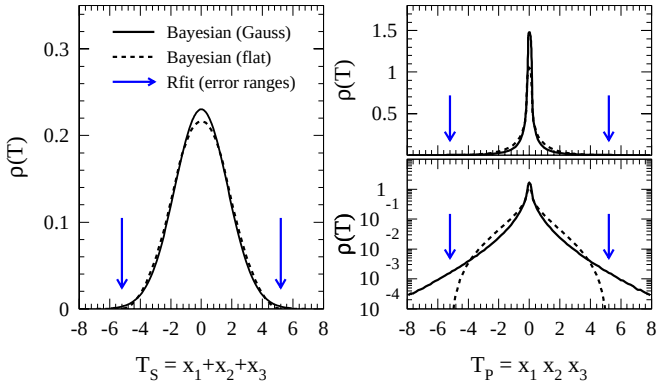

Fig. 18. Convolution of the sum $T_{\mathrm{S}}^{(3)} = x_1 + x_2 + x_3$ (left hand plot) and the product $T_{\mathrm{P}}^{(3)} = x_1 x_2 x_3$ (right hand plot) of $\mathcal{N} = 3$ y_{QCD} parameters as defined in the text. Plotted is the PDF $\rho(T)$ obtained from Bayesian statistics using for the $G(x)$ PDF a uniform (solid lines, $\Delta = \sqrt{3}$) or a Gaussian (dashed lines, $\sigma = 1$) distribution. Both PDFs $\rho(T)$ of $T_{\mathrm{P}}^{(3)}$ present a singularity at the origin which is not shown. The *R*fit ranges of $T_{\mathrm{S}}^{(3)}$ and $T_{\mathrm{P}}^{(3)}$ are indicated by the arrows located in both instances at $\pm 3\sqrt{3}$

B Critical issues of the 95% CL scan method

The 95% CL Scan method 4.2 presents several unwelcome features which are reviewed below.

Drawing features: Its final graphical result depends strongly on the choices made for $\mathrm{CL_{cut}}$, the selection threshold used to retain *models*, and for $\mathrm{CL_{cont}}$, the constant defining the contours to be drawn. It is not possible to infer how a given drawing is modified by changing these choices. Regions of weak statistical confidence, just barely passing the selection threshold, and thus retained, bear the same graphical weight than regions of fair CL. Similarly, regions of lower CL, barely not passing the selection threshold, and thus not retained, are ignored.

Contour features: The relevance of drawing the contours is not obvious. Whereas a given contour has a clear-cut meaning, the message it carries is to show how precise the determination of a would be, in the academic situation where all theoretical uncertainties would supposedly be resolved. These contours are not ellipses in general. For instance, in the case of the measurement of $\sin 2\beta$ alone, they are built from straight lines (*c.f.*, Sect. 3.2.3). They can take complicate shapes, depending on the choice made for the a variables, and depending on the presence of sizeable secondary minima. In simple situations, using the 95% CL Scan method yields awkward results. For instance, in the first example given in Sect. 3.2.3, where a standard situation is considered, the method would conclude by a set of short intervals in a, each obtained for a fixed y_{QCD} value instead of providing the overall allowed interval, *i.e.*, applying the *R*fit scheme.

Envelope features: In the a space no information can be carried about possibly more or less favored regions: a point is either within the envelope of the countours, and thus acceptable, or outside, and thus not acceptable. The envelope of the contours can be unstable with respect to change of $\mathrm{CL_{cut}}$. Indeed, the envelope can be a discontinuous function of this parameter: if lowering $\mathrm{CL_{cut}}$ allows an outsider *model* to be selected outside of the envelope, this outsider *model* surfaces in the a space with its full contour and thus lead to an abrupt change of the envelope. The envelope provided by the 95% CL Scan method tends to be over-conservative. For the example of Sect. 3.2.3, the method yields for the envelope of the 95% CL intervals:

$$a_\pm = a^0 \pm 1.96\,\frac{\sigma[a]}{\sqrt{1-c^2}}\left(c+\sqrt{1-c^2}\right)\ , \tag{102}$$

which is always larger than the correct result given by (45) (the two limiting cases $c \to 0$ and $c \to 1$ excepted).

C Comments on statistics of normal ratios

Ratios of branching fractions of rare $b \to u$ transitions have attracted much attention in recent theoretical and experimental analyses, by virtue of their potential to constrain the unitarity angle γ (see Sect. 6.3). It is shown in this section that the extraction of physical observables out of ratios of normally (Gaussian) distributed quantities, *e.g.*, branching fractions, requires some precaution.

C.1 A Numerical Example

The statistical discussion will be accompanied by a numerical example for rare charmless B decays. It is assumed that branching ratios are measured *via* the relation $\mathrm{BR}_z = N_z/(\epsilon_z\,\sigma\mathcal{L})$, where the reconstruction efficiencies of the two considered final states, $z \equiv x, y$, shall be $\epsilon_x = \epsilon_y = 13\%$, the collected integrated luminosity $\mathcal{L} = 20\ \mathrm{fb}^{-1}$, and the production cross section $\sigma = 1.1$ nb. The "measured" branching fractions

$$\begin{aligned}\mathrm{BR}_x &= (16.0 \pm 2.4)\times 10^{-6}\ ,\\ \mathrm{BR}_y &= (8.0 \pm 1.7)\times 10^{-6}\ ,\end{aligned} \tag{103}$$

correspond thus to 46 and 23 detected signal events in the channels x and y, respectively. The number of events is sufficient to escape from Poissonian to normal PDFs for the branching fractions[31]. To derive constraints from the above measurements one needs a predictive theory which, in our example, shall be given by the expressions

$$\begin{aligned}\mathrm{BR}_{x,\,\mathrm{theo}}(\gamma) &= |F(0)A_x(\gamma)|^2\ ,\\ \mathrm{BR}_{y,\,\mathrm{theo}}(\gamma) &= |F(0)A_y(\gamma)|^2\ ,\end{aligned} \tag{104}$$

with the "form factor" $F(0)$, being identical for both final states, and where the "amplitudes", which are functions of the "angle" γ, shall read

$$\begin{aligned}A_x(\gamma) &= 1 + e^{i\gamma}\ ,\\ A_y(\gamma) &= 1 - e^{i\gamma}\ .\end{aligned} \tag{105}$$

Our theory depends on the external parameters $F(0)$ and γ. The latter provides an example of a quantity we are interested in ($a = \{\gamma\}$), while the modulus-squared of the first is an example of a y_{QCD} parameter. In the framework of Rfit, $|F(0)|^2$ can be eliminated according to Sect. 3.2.2, using (39), as illustrated in Sect. 5.2. It can also be eliminated by taking the ratio $R_{\mathrm{theo}}(\gamma) \equiv \mathrm{BR}_{x,\,\mathrm{theo}}(\gamma)/\mathrm{BR}_{y,\,\mathrm{theo}}(\gamma) = |A_x(\gamma)|^2/|A_y(\gamma)|^2$. Although it appears more straightforward to use the ratio than the Rfit treatment, it is shown below that the converse is true: using the ratio leads to cumbersome formulae.

C.2 Probability Density Functions

We define a set of two statistically independent measurements

$$x = x_0 \pm \sigma_x\ , \tag{106}$$

$$y = y_0 \pm \sigma_y\ , \tag{107}$$

obeying normal distributions

$$G(z, z_0, \sigma_z) = \frac{1}{\sqrt{2\pi}\,\sigma_z}\exp\left(-\frac{(z-z_0)^2}{2\,{\sigma_z}^2}\right)\ , \tag{108}$$

with $z \equiv x, y$. For the ratio

$$R = \frac{x}{y}\ , \tag{109}$$

the marginal PDF obtained from error propagation

$$\rho_1(R) \approx G\left(R,\, R_0,\, R_0\sqrt{\frac{\sigma_x^2}{x^2}+\frac{\sigma_y^2}{y^2}}\right)\ , \tag{110}$$

where $R_0 = x_0/y_0$, is only a coarse approximation of the exact solution

$$\begin{aligned}\rho_0(R) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta\left(R-\frac{x}{y}\right)G(x,x_0,\sigma_x)\\ &\quad\times G(y,y_0,\sigma_y)\,dx\,dy\ ,\\ &= G(0,x_0,\sigma_x)G(0,y_0,\sigma_y)\frac{2}{\eta(R)}\\ &\quad\times\left[1+\sqrt{\pi\xi(R)}\,e^{\xi(R)}\,\mathrm{Erf}\left(\sqrt{\xi(R)}\right)\right]\ .\end{aligned} \tag{111}$$

Here, the functions ξ and η are defined as

$$\begin{aligned}\xi(R) &= \frac{1}{2\eta(R)}\left(R\frac{x_0}{\sigma_x^2}+\frac{y_0}{\sigma_y^2}\right)^2\ ,\\ \eta(R) &= \frac{1}{\sigma_x^2}R^2+\frac{1}{\sigma_y^2}\ .\end{aligned}$$

[31] As ratios of branching fractions are the subject of discussion here, the difference between normal and Poissonian statistics is much reduced in the resulting PDF, so that the results obtained in this section remain approximately valid also for a low number of signal events

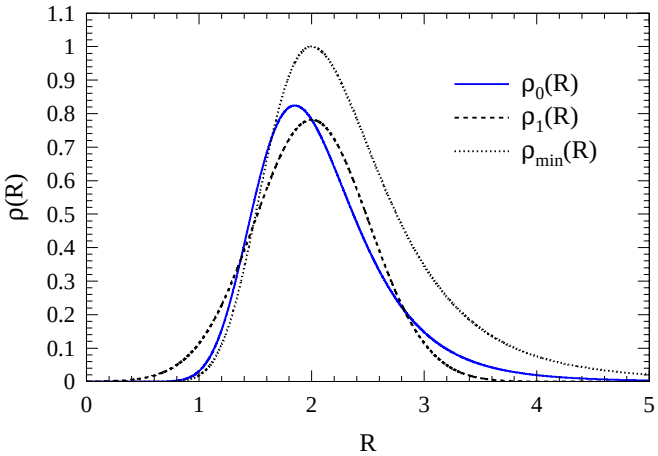

Fig. 19. Probability density functions $\rho_0(R)$ (solid line), $\rho_1(R)$ (dashed line), and the likelihood $\rho_{\min}(R)$ (dotted line), where $R = x/y$. The numerical values are those chosen for the example (see text). The distribution of $\rho_1(R)$ is symmetric by construction. The likelihood $\rho_{\min}$ is normalized in such a way that its maximal value (not its integral) is unity (see (120))

Equations (110) and (111) satisfy

$$\int_{-\infty}^{\infty} \rho_{0,1}(R)\, dR \;=\; 2 \cdot \int_{-\infty}^{R_0} \rho_{0,1}(R)\, dR \;=\; 1\ . \qquad (112)$$

The densities $\rho_0(R)$ and $\rho_1(R)$ are plotted in Fig. 19 for the example of the previous section. The maximum of $\rho_0(R)$ is shifted from the naive expectation, $\mathrm{Max}_R\{\rho_1(R)\} = \rho_1(R_0)$, to lower values, while its mean value is larger than the naive mean: $\langle\rho_0\rangle = 2.10 > \langle\rho_1\rangle = 2$. The root mean square (RMS) of the correct solution exceeds the one of the naive approximation: $\mathrm{RMS}(\rho_0) = 0.62 > \mathrm{RMS}(\rho_1) = 0.51$, so that the use of (110) will tend to over-optimistic results.

C.3 Confidence Levels

In our example, the analysis of the branching fractions, or of their ratio, aims at constraining the physical quantity γ. Figure 20 shows in its upper plot the theoretical ratio $R_{\mathrm{theo}}(\gamma)$ versus γ, together with the asymmetric, one standard deviation error band (using (116)) of the "experimental" value. Without loss of generality, we may assume in the following that, apart from γ, the theoretical predictions (104) and (105) suffer only from the theoretical uncertainty on $F(0)$, which shall be maximally unknown: *i.e.*, we do not assume that $|F(0)|^2$ takes its values within an *a priori* finite range. An approach to obtain confidence levels for γ, that is often met in the literature, is to eliminate the factor $F(0)$ by using the ratio (109) and defining the χ^2 as

$$\chi_1^2(\gamma) = \left(\frac{R_0 - R_{\mathrm{theo}}(\gamma)}{\sigma_R}\right)^2 , \qquad (113)$$

where σ_R is the RMS of ρ_1 (see (110)). The associated CL is then given by the cumulative distribution of a normal

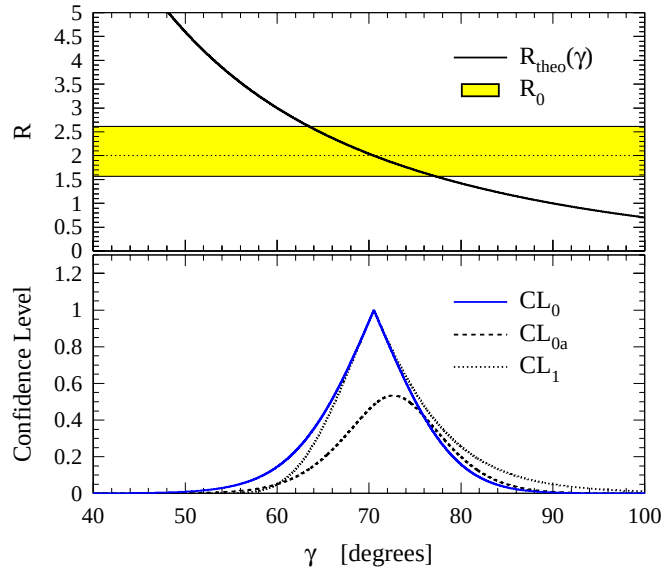

Fig. 20. *Upper plot:* "measured" ratio of branching fractions within its asymmetric, one standard deviation error band (grey zone), corresponding to the example (103), and the "theoretical prediction" (104) (solid curve) versus γ. *Lower plot:* confidence levels obtained from (113) (dotted line), (115) (dashed line) (both using (114)), and (116) (solid line), the latter being identical to (119)

PDF

$$\mathrm{CL}_1(\gamma) = \mathrm{Erfc}\left(\sqrt{\chi_1^2(\gamma)/2}\right) , \qquad (114)$$

shown versus γ by the dotted line in the lower plot of Fig. 20.

Yet, we have seen that the PDF ρ_1 is only an approximation of the correct PDF ρ_0 and hence, one is tempted to build a more accurate χ^2 by means of

$$\chi_{0a}^2(\gamma) = -2\ln\rho_0(R_{\mathrm{theo}}(\gamma)) , \qquad (115)$$

where one may wish to build the corresponding CL, $\mathrm{CL}_{0a}(\gamma)$, *via* (114). However, this again constitutes an approximation since the error function assumes a normal PDF. The dotted curve in the lower plot of Fig. 20 shows the CL corresponding to (115) and (114) as a function of γ. Indeed, the correct CL is obtained *via* an asymmetric integration of the PDF ρ_0:

$$\mathrm{CL}_0(\gamma) = \begin{cases} 2\int_{-\infty}^{R(\gamma)} \rho_0(R')\, dR' \ , \forall R(\gamma) \le R_0 \\ 2\int_{R(\gamma)}^{\infty} \rho_0(R')\, dR' \ , \forall R(\gamma) > R_0 \end{cases} \qquad (116)$$

plotted versus γ as solid line in the lower plot of Fig. 20.

The complication of (116) can be readily circumvented when not explicitly using the ratio, but keeping the original branching fractions in the definition of the χ^2:

$$\chi^2(\gamma, y_{\mathrm{QCD}}) = \left(\frac{x - y_{\mathrm{QCD}} \cdot |A_x(\gamma)|^2}{\sigma_x}\right)^2$$

$$+\left(\frac{y - y_{\mathrm{QCD}} \cdot |A_y(\gamma)|^2}{\sigma_y}\right)^2 . \quad (117)$$

Applying (39), hence eliminating y_{QCD}, yields the minimum

$$\chi^2_{\mathrm{min};y_{\mathrm{QCD}}}(\gamma) = \frac{(x - yR_{\mathrm{theo}}(\gamma))^2}{\sigma_x^2 + \sigma_y^2 R^2_{\mathrm{theo}}(\gamma)} , \quad (118)$$

which, by construction, only depends on $R_{\mathrm{theo}}(\gamma)$. The CL obtained using (114)

$$\mathrm{CL}_{\mathrm{min}}(\gamma) = \mathrm{Erfc}\left(\sqrt{\chi^2_{\mathrm{min};y_{\mathrm{QCD}}}(\gamma)/2}\right) , \quad (119)$$

is identical to the one obtained using (116). The likelihood corresponding to (118)

$$\rho_{\mathrm{min}}(R_{\mathrm{theo}}(\gamma)) = e^{-\frac{1}{2}\chi^2_{\mathrm{min};y_{\mathrm{QCD}}}(\gamma)} , \quad (120)$$

is shown as the dotted curve in Fig. 19. The likelihood is equal to unity for $R_{\mathrm{theo}}(\gamma) = x/y$. It is worth emphasizing that $\rho_{\mathrm{min}}(R)$, which yields the correct CL, when using (114), is not identical to $\rho_0(R)$, although the latter is the correct PDF, which yields the correct CL, when *not* using (114), but (116) instead.

References

1. N. Cabibbo, Phys. Rev. Lett. **10**, 351 (1963)
2. M. Kobayashi, T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973)
3. L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983)
4. A.J. Buras, M.E. Lautenbacher, G. Ostermaier, Phys. Rev. **D50**, 3433 (1994)
5. M.H. Schune, S. Plaszczynski, "Overall Determination of the CKM Matrix", LAL-99-67, hep-ph/9911280, 1999; BABAR Collaboration, "The BABAR Physics Book", Chapter 14, SLAC-R-504, 1998
6. M. Ciuchini et al., "2000 CKM triangle analysis: a critical review with updated experimental inputs and theoritical parameters", LAL-00-77, ROME1-1307-00, RM3-TH-00-16, hep-ph/0012308, 2000
7. F. Caravaglios, F. Parodi, P. Roudeau, A. Stocchi, "Determination of the CKM unitarity parameters by end 1999", LAL-00-04, Feb 2000; hep-ph/0002171, 2000 F. Caravaglios, F. Parodi, P. Roudeau, A. Stocchi, "Constraints on the V(CKM) matrix by end 1998", LAL-99-03, DELPHI-99-27, Nuovo Cim. **A112**, 833 (1999); F. Parodi, P. Roudeau, A. Stocchi, "Constraints on the V(CKM) matrix by end 1997", LAL-98-49, 1998, hep-ph/9802289,
8. H. Abele, S. Baessler, D. Dubbers, J. Reich, Nucl. Phys **A663**, 947 (2000); A. Ali, Acta Phys. Polon. **B27**, 3529 (1996); A. Ali, B. Kayser, hep-ph/9806230, 1998; A. Ali, D. London, Eur. Phys. J. **C18**, 665 (2001) Phys. Rept. **320**, 79 (1999) Eur. Phys. J. **C9**, 687 (1999) Nucl. Phys. **54A**, 297 (1997) Nuovo Cim. **A109**, 957 (1996) Z. Phys. **C65**, 431 (1995) hep-ph/9409399, 1994 hep-ph/9405283, 1994; D. Atwood, A. Soni, AMES-HET-01-02, hep-ph/0103197 (2001); R. Barbieri, L.J. Hall, S. Raby, A. Romanino Nucl. Phys. **B493**, 3 (1997); G.C. Branco, F. Cagarrinho, F. Kruger, Phys. Lett. **B459**, 224 (1999); G. Buchalla, A.J. Buras, M.E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996) A. Buras et al., Phys. Lett. **B500**, 161 (2001); A.J. Buras, hep-ph/9711217, 1997 hep-ph/9704376, 1997; R.N. Cahn, M.P. Worah, Phys. Rev. **D60**, 076006 (1999); M. Ciuchini, E. Franco, G. Martinelli, L. Reina, L. Silvestrini, Z. Phys. **C68**, 239 (1995); M. Ciuchini, E. Franco, L. Giusti, V.Lubicz, G. Martinelli, Nucl.Phys. **B573**, 201 (2000); T. DeGrand, COLO-HEP-447, hep-ph/0008234; B.R. Desai, A.R. Vaucher, UCRHEP-T282, hep-ph/0007233; P. Faccioli, hep-ph/0011269; A.F. Falk, JHU-TIPAC-200005, hep-ph/0007339; F.J. Gilman, hep-ph/0102345; F.J. Gilman, K. Kleinknecht, B. Renk, Eur. Phys. J. **C3**, 103 (1998); M. Gronau, TECHNION-PH-2000-30, hep-ph/0011392; S. Herrlich, U. Nierste, Phys. Rev. **D52**, 6505 (1995); M. Randhawa, M. Gupta, "Constructing 'Reference' Triangle through Unitarity of CKM Matrix" hep-ph/0102274 (2001); D.E. Jaffe, S. Youssef, Comput. Phys. Commun. **101** (1997) 206 M. Lusignoli, L. Maiani, G. Martinelli, L. Reina, Nucl. Phys. **B369**, 139 (1992); S. Mele, Phys.Rev. **D59**, 113011 (1999); M. Randhawa, M. Gupta, hep-ph/0102274; J.L. Rosner, EFI-2000-42, hep-ph/0011184; M. Schmidtler, K. R. Schubert, Z. Phys. **C53**, 347 (1992); S. Stone hep-ph/9910417 (1999); D. Wyler, ZU-01-01, hep-ph/0101259; see also [18] and other references of this bibliography
9. L.L. Chau, W.Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984)
10. Particle Data Group, D.E. Groom et al., Eur. Phys. J. **C15**, 1 (2000)
11. C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985)
12. I. Bigi, A.I. Sanda, "CP Violation", Cambridge University Press, UK, 2000
13. G.C. Branco, L. Lavoura, J.P. Silva, "CP Violation", The Intern. Series of Monographs on Physics - 103, Oxford Science Publications, Oxford, UK, 1999
14. H. Lacker et al., CKM Constraints as a Function of Luminosity (Time)", Talk given at the International Workshop on B Physics and CP Violation, BCP4, Ise-Shima, Japan, February 2001, LAL 01-14, 2001
15. G. D'Agostini, "Sceptical combination of experimental results: General considerations and application to ϵ'/ϵ", CERN-EP-99-139, hep-ex/9910036, 1999
16. A. Höcker, H. Lacker, S. Laplace, F. Le Diberder, "Combining Inconsistent Measurements", LAL 01-12 (2001)
17. I.S. Towner, J.C. Hardy, "The current status of $|V_{ud}|$", nucl-th/9809087, 1998
18. M. Bargiotti et al., Riv. Nuovo Cim. **23** N3 (2000) 1
19. K. Saito, A.W. Thomas, Phys. Lett. **B363**, 157 (1995)
20. H. Abele, Nucl. Instr. & Meth. **A440**, 499 (2000)
21. S. Arzumanov, Nucl. Instr. & Meth. **A440**, 511 (2000)
22. J.F. Donoghue, B.R. Holstein, S.W. Klimt, Phys. Rev. **D35**, 934 (1987)
23. R. Flores-Mendieta, A. Garcia, G. Sanchez-Colon, Phys. Rev. **D54**, 6855 (1996)
24. H. Leutwyler, M. Roos, Z. Phys. **C25**, 91 (1984)
25. E.A. Paschos, U. Türke, Phys. Rep. **178**, 145 (1989)
26. W. Jaus, Phys. Rev. **D44**, 2851 (1991)
27. W.J. Marciano, A. Sirlin, Phys. Rev. Lett. **56**, 22 (1986)
28. H.H. Williams et al., Phys. Rev. Lett. **33**, 240 (1974); T. Becherrawy et al., Phys. Rev. **D1**, 1452 (1970)
29. ALEPH Collaboration (R. Barate et al.), Eur. Phys. J. **C6**, 555 (1999)

30. L3 Collaboration (M. Acciarri et al.), Phys. Lett. **B436**, 174 (1998)
31. DELPHI Collaboration (P. Abreu et al.), Phys. Lett. **B478**, 14 (2000)
32. CLEO Collaboration (J.P. Alexander et al.), Phys. Rev. Lett. **77**, 5000 (1996); CLEO Collaboration (B.H. Behrens et al.), Phys. Rev. **D61**, 052001 (2000)
33. ARGUS Collaboration (H. Albrecht et al.), Phys. Lett. **B255**, 297 (1991)
34. CLEO Collaboration (R. Fulton et al.), Phys. Rev. Lett. **64**, 16 (1990)
35. CLEO Collaboration (J. Bartelt et al.), Phys. Rev. Lett. **71**, 4111 (1993)
36. G. Altarelli, N. Cabibbo, G. Corbò, L. Maiani, G. Martinelli, Nucl. Phys. **B208**, 365 (1982)
37. N. Isgur, D. Scora, B. Grinstein, M.B. Wise, Phys. Rev. **D39**, 799 (1989)
38. J.G. Körner, G.A. Schuler, Z. Phys. **C38**, 511, (1988) Erratum-ibid. **C41**, 690 (1989)
39. M. Wirbel, B. Stech, M. Bauer, Z. Phys. **C29**, 637 (1985)
40. D. Abbaneo et al., "Combined results on b-hadron production rates, lifetimes and semileptonic decays", SLAC-PUB-8492, CERN-EP-2000-096, March 2000, hep-ex/0009052, 2000
41. I.I. Bigi, "Heavy Flavor Physics: On its more than 50 years of history, its future and the Rio Manifesto", Summary Talk given at HQ2K Heavy Quarks at Fixed Target 2000, Rio de Janeiro, Brazil, Oct. 9 -12, UND-HEP-00-BIG 11, hep-ph/0012161, 2000
42. CDHS Collaboration (H. Abramowicz et al.), Z. Phys. **C15**, 19 (1982)
43. CCFR Collaboration (A.O. Bazarko et al.), Z. Phys. **C65**, 189 (1995)
44. CHARM II Collaboration (P. Vilain et al.), Eur. Phys. J. **C11**, 19 (1999)
45. H. Bolton et al., "Determining the CKM parameter $|V_{cd}|$ from νN charm production", KSU-HEP-97-04, hep-ex/9708014, 1997
46. Fermilab E531 Collaboration (N. Ushida et al.), Phys. Lett. **B206**, 375 (1988)
47. CLEO Collaboration (Y. Kubota et al.), Phys. Rev. **D54**, 2994 (1996)
48. L. Montanet et al., Phys. Rev. **D50**, 1173 (1994)
49. ALEPH Collaboration (R. Barate et al.), Phys. Lett. **B453**, 107 (1999); ALEPH Collaboration (R. Barate et al.), ALEPH CONF 99-038, July 1999; ALEPH Collaboration (R. Barate et al.), Phys. Lett. **B465**, 349 (1999)
50. DELPHI Collaboration (P. Abreu et al.), Phys. Lett. **B439**, 209 (1998)
51. L3 Collaboration, HEP'99, L3 Note 2376, March 1999
52. OPAL Collaboration, OPAL Physics Note PN378, March 1999; OPAL Physics Note PN402, July 1999
53. OPAL Collaboration, (G. Abbiendi et al.), Phys. Lett. **B490**, 71-86 (2000)
54. N. Isgur, M. Wise, Phys. Lett. **B232**, 113 (1989); **237**, 527 (1990); H. Georgi, Phys. Lett. **B238**, 395 (1990)
55. I.I. Bigi, M. Shifman, N.G. Uraltsev, Annu. Rev. Nucl. Part. Sci. **47**, 591 (1997)
56. J. Simone et al., Nucl. Phys. Proc. Supp. **83**, 33 (2000)
57. LEP $|V_{cb}|$ working group, "LEP $|V_{cb}|$ average", http://lepvcb.web.cern.ch/LEPVCB/Osaka00.html
58. CLEO Collaboration (J.P. Alexander et al.), "Determination of the $B \to D^* \ell \nu$ decay width and $|V_{cb}|$", Contribution to XXXth International Conference on High Energy Physics, Osaka, Japan, CLEO CONF 00-03, ICHEP-00-770, hep-ex/0007052, 2000
59. I.I. Bigi, "Memo on extracting $|V_{cb}|$ and $|V_{ub}/V_{cb}|$ from semileptonic B decays", Contributed to the Informal Workshop on the Derivation of $|V_{cb}|$ and $|V_{ub}|$, CERN, May 28 - June 2, 1999, hep-ph/9907270
60. CLEO Collaboration (B. Barish et al.), Phys. Rev. Lett. **76**, 1570 (1996)
61. M. Artuso, private communications, Sep. 2000
62. CDF Collaboration, (T. Affolder et al.) "First measurement of the ratio $B(t \to Wb)/B(t \to Wq)$ and associated limit on the CKM element $|V_{tb}|$", hep-ex/0012029, 2000
63. We use the average quoted in [18]
64. ALEPH Collaboration (R. Barate et al.), Phys. Lett. **B429**, 169 (1998)
65. CLEO Collaboration (M.S. Alam et al.), Phys. Rev. Lett. **74**, 2885 (1995); CLEO Collaboration (S. Glenn et al.), "Improved measurement of $B(B \to s\gamma)$", ICHEP'98 #1011, CLEO CONF 98-17, 1998
66. C. Geweniger et al., Phys. Lett. **B48**, 487 (1974)
67. A. Apostolakis et al., Phys. Lett. **B458**, 545 (1999)
68. A. Angelopoulos et al., Phys. Lett. **B420**, 191 (1998)
69. J. H. Christenson et al., Phys. Rev. Lett. **43**, 1209 (1979)
70. T. Inami, C.S. Lim, Prog. Theor. Phys. **65**, 297 (1981); Erratum-ibid. **65**, 1772 (1981)
71. G. Buchalla, A.J. Buras, M.E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996)
72. S. Herrlich, U. Nierste, Nucl. Phys. **B419**, 292 (1994)
73. A.J. Buras, M. Jamin, P.H. Weisz, Nucl. Phys. **B347**, 491 (1990)
74. S. Herrlich, U. Nierste, Phys. Rev. **D52**, 6505 (1995) Nucl. Phys. **B476**, 27 (1996)
75. L. Lellouch, Nucl. Phys. Proc. Suppl. **94**, 142 (2001)
76. NA31 Collaboration (G.D. Barr et al.), Phys. Lett. **B317**, 233 (1993)
77. E731 Collaboration (L.K. Gibbons et al.), Phys. Rev. **D55**, 6625 (1997)
78. KTeV Collaboration (A. Alavi-Harati et al.), Phys. Rev. Lett. **83**, 22 (1999)
79. NA48 Collaboration, (G. Unal et al.) Nucl. Phys. Proc. Suppl. **93**, 241 (2001)
80. M. Ciuchini, G. Martinelli, Nucl. Phys. Proc. Suppl. **99**, 27 (2001)
81. C. Bozzi for the BABAR Collaboration, "B mixing and lifetime measurements with the BABAR detector.", SLAC-PUB-8791, BABAR-PROC-01-08, hep-ex/0103046, 2001
82. N. Hastings for the Belle Collaboration, "Measurements of B lifetimes and B^0-$\overline{B}^0$ mixing.", Talk given at the International Workshop on B Physics and CP Violation, BCP4, Ise-Shima, Japan, February 2001
83. C. Bernard, Nucl. Phys. Proc. Suppl. **94**, 159 (2001)
84. C. Bernard, private communications, Sep. 2000
85. CP-PACS Collaboration(A. Ali Khan et al.), "Decay constants of B and D mesons from improved relativistic lattice QCD with two flavors of sea quarks", UTCCP-P-68, hep-lat/0010009, 2000
86. C. Bernard et al., Nucl. Phys. Proc. Suppl. **94**, 346 (2001)
87. ALEPH Collaboration (R. Barate et al.), Eur. Phys. J. **C4**, 367 (1998); Eur. Phys. J. **C7**, 553 (1999)

88. DELPHI Collaboration, Conference Paper #236 ICHEP'98, Vancouver, 1998; DELPHI 99-109 CONF 296, EPS-HEP'99 Tampere, 1999
89. OPAL Collaboration (K. Ackerstaff et al.), Z. Phys. **C76**, 401 (1997); Z. Phys. **C76**, 417 (1997);
90. SLD Collaboration, SLAC-PUB-8225, Lepton-Photon'99, Stanford 1999
91. CDF Collaboration (F. Abe et al.), Phys. Rev. Lett. **82**, 3576 (1999)
92. H.G. Moser, A. Roussarie, Nucl. Inst. Meth. **A384**, 491 (1997)
93. D. Abbaneo, G. Boix, J. of HEP **08**, 4 (1999)
94. LEP working group on B oscillations, "Results for the BEAUTY 2000 conference", http://lepbosc.web.cern.ch/LEPBOSC/combined_results/sept_2000
95. P. Checchia, E. Piotto, F. Simonetto, "The hypothesis of a real Cabibbo Kobayashi Maskawa matrix", DFPD-99-EP-27, hep-ph/9907300, 1999
96. BABAR Collaboration, B. Aubert et al., Phys.Rev.Lett. **86**, 2515 (2001)
97. Belle Collaboration, A. Abashian et al., Phys. Rev. Lett. **86**, 2509 (2001)
98. BABAR Collaboration, T. Affolder et al., Phys. Rev. **D61**, 072005 (2000)
99. R. Barate et al., Phys. Lett. **B492**, 259 (2000)
100. K. Ackerstaff et al., Eur. Phys. J. **C5**, 379 (1998)
101. KTeV Collaboration (A. Alavi-Harati et al.), Phys. Rev. **D61**, 072006 (2000)
102. L. Littenberg, Phys. Rev. **D39**, 3322 (1989); L. Littenberg, "Rare Kaon Decays", BNL-67772 (2000), hep-ex/0010048, 2000
103. G. Buchalla, A. Buras, Nucl. Phys. **B548**, 309 (1999); G. Buchalla, A. Buras, Phys. Rev. **D54**, 6782 (1996); G. Buchalla, A. Buras, Nucl. Phys. **B400**, 225 (1993)
104. W. Marciano, Z. Parsa, Phys. Rev. **D53**, 1 (1996)
105. G. Buchalla, G. Isidori, Phys. Lett. **B440**, 170 (1998)
106. KOPIO Collaboration, http://sitka.triumf.ca/e926/pub/nsf/index.html
107. KAMI Collaboration, http://kpasa.fnal.gov:8080/public/kami/kami.html
108. G. Buchalla, A. Buras, Nucl. Phys. **B412**, 106 (1994)
109. E787 Collaboration (S. Adler et al.), Phys. Rev. Lett. **84**, 3768 (2000)
110. E949 Collaboration, http://www.phy.bnl.gov/e949/
111. CKM Collaboration, http://www.fnal.gov/projects/ckm/Welcome.html
112. J. Atwood, G. Eilam, A. Soni, Mod. Phys. Lett. **A11**, 1061 (1996)
113. J. Soares, Nucl. Phys. **B367**, 575 (1991)
114. K. Chetyrkin, M. Misiak, M. Münz, Phys. Lett. **B400**, 206 (1997); erratum-ibid. **B425**, 414 (1998)
115. A.L. Kagan, M. Neubert, Phys. Rev. **D58**, 094012 (1998)
116. A.L. Kagan, M. Neubert, Eur. Phys. J. **C7**, 5 (1999)
117. K. Kiers, A. Soni, G.H. Wu, Phys. Rev. **D62**, 116004 (2000)
118. R. Fleischer, Phys. Lett. **B365**, 399 (1996); R. Fleischer, T. Mannel, Phys. Rev. **D57**, 2752 (1998)
119. R. Fleischer, A. Buras, Eur. Phys. J. **C16**, 97 (2000)
120. M. Beneke, G. Buchalla, M. Neubert, C.T. Sachrajda, Phys. Rev. Lett. **83**, 1914 (1999); M. Beneke, G. Buchalla, M. Neubert, C.T. Sachrajda, Nucl. Phys. **B591**, 313 (2000); M. Beneke, G. Buchalla, M. Neubert, C.T. Sachrajda, "QCD Factorization for $B \to \pi K$ Decays", PITHA 00/13, hep-ph/0007256, 2000
121. Y.Y. Keum, H.N. Li, C. Kung, A.I. Sanda, Phys. Rev. **D63**, 054008 (2001) Y.Y. Keum, H.N. Li, A.I. Sanda, "Fat Penguins and Imaginary Penguins in Perturbative QCD", KEK-TH-642, NCKU-HEP-00-01, KIAS-P00014, DPNU-00-13, hep-ph/0004004, 2000
122. Y. Grossman, Y. Nir, S. Plaszczynski, M.H. Schune, Nucl. Phys. **B511**, 69 (1998)
123. G. Cavoto, "Charmless hadronic B decays at BABAR", Talk given at the XXXVIth Rencontres de Moriond Electroweak Interactions and Unified Theories, Les Arcs, France, March 2001
124. T. Iijima, "B meson rare decays", Talk given at the International Workshop on B Physics and CP Violation, BCP4, Ise-Shima, Japan, February 2001
125. E. Thorndike, "Recent Results from CLEO", Talk given at the International Workshop on B Physics and CP Violation, BCP4, Ise-Shima, Japan, February 2001
126. M. Beneke, G. Buchalla, M. Neubert and C.T. Sachrajda, "QCD Factorization in $B \to \pi K$, $\pi\pi$ Decays and Extraction of Wolfenstein Parameters", CERN-TH/2001-107, CLNS 01/1728, PITHA 01/01, SHEP 01/11, hep-ph/0104110, 2001
127. F. Krauss, G. Soff, "NLO QCD corrections to $\bar{B}B$ mixing and ϵ_K within the MSSM", hep-ph/9807238, 1998
128. M. Ciuchini, G. Degrassi, P. Gambino, G.F. Giudice, Nucl. Phys. **B534**, 3 (1998)
129. A. Ali, D. London, Eur. Phys. J. **C9**, 687 (1999)
130. A.J. Buras, P. Gambino, M. Gorbahn, S. Jager, L. Silvestrini, Nucl. Phys. **B592**, 55 (2001)
131. A. Buras, R. Buras, Phys.Lett. **B501**, 223 (2001)