# Global analysis of inclusive $B$ decays

Christian W. Bauer,[1] Zoltan Ligeti,[2] Michael Luke,[3] Aneesh V. Manohar,[4] and Michael Trott[3]

[1] *California Institute of Technology, Pasadena, CA 91125*
[2] *Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720*
[3] *Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7*
[4] *Department of Physics, University of California at San Diego, La Jolla, CA 92093*

In light of the large amount of new experimental data, we revisit the determination of $|V_{cb}|$ and $m_b$ from inclusive semileptonic and radiative $B$ decays. We study shape variables to order $\Lambda_{\rm QCD}^3/m_b^3$ and $\alpha_s^2\beta_0$, and include the order $\alpha_s\,\Lambda_{\rm QCD}/m_b$ correction to the hadron mass spectrum in semileptonic decay, which improves the agreement with the data. We focus on the $1S$ and kinetic mass schemes for the $b$ quark, with and without expanding $m_b - m_c$ in HQET. We perform fits to all available data from BABAR, BELLE, CDF, CLEO, and DELPHI, discuss the theoretical uncertainties, and compare with earlier results. We find $|V_{cb}| = (41.9 \pm 0.6 \pm 0.1_{\tau_B}) \times 10^{-3}$ and $m_b^{1S} = 4.68 \pm 0.04\,\text{GeV}$, including our estimate of the theoretical uncertainty in the fit.

## I. INTRODUCTION

In the last few years there has been intense theoretical and experimental activity directed towards a precise determination of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{cb}|$ from combined fits to inclusive semileptonic $B$ decay distributions [1–4]. The idea is that using the operator product expansion (OPE), sufficiently inclusive observables can be predicted in terms of $|V_{cb}|$, the $b$ quark mass, $m_b$, and a few nonperturbative matrix elements that enter at order $\Lambda_{\rm QCD}^2/m_b^2$ and higher orders. One then extracts these parameters and $|V_{cb}|$ from shapes of $B$ decay spectra and the semileptonic $B$ decay rate. This program also tests the consistency of the theory and the accuracy of the theoretical predictions for inclusive decay rates. This is important also for the determination of $|V_{ub}|$, whose error is a major uncertainty in the overall constraints on the unitarity triangle.

The OPE shows that in the $m_b \gg \Lambda_{\rm QCD}$ limit inclusive $B$ decay rates are equal to the $b$ quark decay rates [5, 6], and the corrections are suppressed by powers of $\alpha_s$ and $\Lambda_{\rm QCD}/m_b$. High-precision comparison of theory and experiment requires a precise determination of the heavy quark masses, as well as the non-perturbative matrix elements that enter the expansion. These are $\lambda_{1,2}$, which parameterize the nonperturbative corrections to inclusive observables at $\mathcal{O}(\Lambda_{\rm QCD}^2/m_b^2)$. At order $\Lambda_{\rm QCD}^3/m_b^3$, six new matrix elements occur, usually denoted by $\rho_{1,2}$ and $\mathcal{T}_{1,2,3,4}$.

In this paper, we will perform a global fit to the available inclusive decay observables from BABAR, BELLE, CDF, CLEO, and DELPHI, including theoretical expressions computed to order $\alpha_s^2\beta_0$, $\alpha_s\Lambda_{\rm QCD}/m_b$ and $\Lambda_{\rm QCD}^3/m_b^3$. A potential source of uncertainty in the OPE predictions is the size of possible violations of quark-hadron duality [7]. Studying $B$ decay distributions is the best way to constrain these effects experimentally, since it should influence the relationship between shape variables of different spectra. We find that at the current level of precision, there is excellent agreement between theory and experiment, with no evidence for violations of duality in inclusive $b \to c$ decays.

A previous analysis of the experimental data was presented in 2002 [1]. There has been considerable new data since then, which has been included in the present analysis, and reduces the errors on $|V_{cb}|$ and $m_b$. In addition, the $\alpha_s\Lambda_{\rm QCD}/m_b$ corrections to the hadronic invariant mass spectrum as a function of the lepton energy cut have now been computed [8], and are included in the present analysis. This reduces the theoretical uncertainty on the hadronic mass moments. We also compare our results with other recent analyses [2, 4, 9].

## II. POSSIBLE SCHEMES

The inclusive $B$ decay spectra depend on the masses of the $b$ and $c$ quarks, which can be treated in many different ways. The $b$ quark is treated as heavy, and theoretical computations for $B^{(*)}$ decays are done as an expansion in powers of $\Lambda_{\rm QCD}/m_b$. The use of the $1/m_b$ expansion is common to all methods.

The decay rates for $B \to X_c$ decay depend on the mass of the $c$ quark, for example, through its effect on the decay phase space. One can treat the $c$ quark as a heavy quark. This allows one to compute the $D^{(*)}$ meson masses as an expansion in powers of $\Lambda_{\rm QCD}/m_c$. The observed $D^{(*)}$ masses can be used to determine $m_c$. Since the computations are performed to $\Lambda_{\rm QCD}^3/m_c^3$, this introduces errors of fractional order $\Lambda_{\rm QCD}^4/m_c^4$ in $m_c$, which gives fractional errors of order $\Lambda_{\rm QCD}^4/(m_b^2 m_c^2)$ in the in-

clusive $B$ decay rates, since charm mass effects first enter at order $m_c^2/m_b^2$. In this method, one starts with the parameters $|V_{cb}|$, $m_b$, $m_c$, $\lambda_{1,2}$, $\rho_{1,2}$ and $\mathcal{T}_{1-4}$. The $B^*-B$, $D^*-D$ and $B-D$ mass differences can be used to eliminate $m_b-m_c$, $\lambda_2$ and $\rho_2$. Only mass differences are used to avoid introducing the parameter $\bar{\Lambda}$ of order $\Lambda_{\rm QCD}$; thus we do not use the $B$ meson mass to eliminate $m_b$. Three linear combinations of the four $\mathcal{T}_i$'s occur in inclusive $B$ decays, and the remaining linear combination would be needed for inclusive $B^*$ decays. In summary the parameters used are (i) $|V_{cb}|$; (ii) one parameter of order the quark mass: $m_b$; (iii) one parameter of order $\Lambda^2_{\rm QCD}$: $\lambda_1$; and (iv) four parameters of order $\Lambda^3_{\rm QCD}$: $\rho_1$, $\mathcal{T}_1-3\mathcal{T}_4$, $\mathcal{T}_2+\mathcal{T}_4$, $\mathcal{T}_3+3\mathcal{T}_4$. These seven parameters are determined by a global fit to moments of the $B$ decay distributions, and the semileptonic branching ratio. This is the procedure used in Ref. [1].

An alternative approach is to avoid using the $1/m_c$ expansion for the charm quark [9], since it introduces $\Lambda_{\rm QCD}/m_c$ corrections, which are larger than the $\Lambda_{\rm QCD}/m_b$ corrections of the $1/m_b$ expansion. In this case heavy quark effective theory (HQET) can no longer be used for the $c$ quark system, and there are no constraints on $m_c$ from the $D$ and $D^*$ meson masses. At the same time, it is not necessary to expand heavy meson states in an expansion in $1/m_{b,c}$, so that the time-ordered products $\mathcal{T}_{1-4}$ can be dropped. With this procedure, one has in addition to (i) $|V_{cb}|$; (ii) two parameters of order the quark mass: $m_{b,c}$; (iii) two parameters of order $\Lambda^2_{\rm QCD}$: $\lambda_{1,2}$; and (iv) two parameters of order $\Lambda^3_{\rm QCD}$: $\rho_{1,2}$. The number of parameters is the same whether or not one expands in $1/m_c$. If one does not expand, two parameters of order $\Lambda^3_{\rm QCD}$ are replaced by two lower order parameters, one of order the quark mass, and one of order $\Lambda^2_{\rm QCD}$. The expansion parameters, such as $\lambda_{1,2}$ are not the same in the two approaches. The values of $\lambda_{1,2}$ not expanding the states in $1/m_Q$ are the values of $\lambda_{1,2}$ plus various time-ordered products $\mathcal{T}_1-\mathcal{T}_4$ when one expands the states in powers of $1/m_Q$.

In addition to the choice of expanding or not expanding in $1/m_c$, one also has a choice of possible quark mass schemes. It has long been known that a "threshold mass" definition for $m_b$ is preferred over both the pole and the $\overline{\rm MS}$ schemes, and it was shown in Ref. [1] that the expansions are indeed better behaved in the $1S$ [10, 11] or PS [12] schemes for $m_b$. If one expands in $1/m_c$, then $m_c$ is eliminated through use of the meson masses, and does not enter the final results. If one does not expand in $1/m_c$, then $m_c$ is a fit parameter. In this method, $m_c$ is treated as much lighter than $m_b$, so the charm quark mass is chosen to be $\overline{m}_c(\mu)$, the $\overline{\rm MS}$ mass renormalized at a scale $\mu\sim m_b$. This is similar to how strange quark mass effects could be included in $B\to X_s\gamma$ decay. In our computation, we will choose the scale $\mu=m_b$.

In addition to the $1S$, PS, pole and $\overline{\rm MS}$ schemes, we have also used the kinetic scheme mass for the $b$-quark, $m_b^{\rm kin}(\mu)$, renormalized at a low scale $\mu\sim 1\,{\rm GeV}$. The scale $\mu$ enters the definition of the kinetic mass, and should not be confused with the scale parameter in dimensional regularization. The relation between the pole and kinetic masses is computed as a perturbative expansion in powers of $\alpha_s(\mu)$, so one cannot make $\mu$ too small. In the kinetic scheme [9] the definitions $\mu_\pi^2 = -\lambda_1+\mathcal{O}(\alpha_s)$, $\mu_G^2=3\lambda_2$, $\rho_D^3=\rho_1+\mathcal{O}(\alpha_s)$, and $\rho_{\rm LS}^3=3\rho_2$ are used.

One cannot decide which scheme is best by counting parameters, or by assuming that not expanding in $1/m_c$ is better than expanding in $1/m_c$. Ultimately, what matters is the accuracy to which experimentally measured quantities can be reliably computed with currently available techniques. For example, full QCD has two parameters, $m_b$ and $m_c$, which can be fixed using the $B$ and $D$ meson masses. [Unfortunately, this is not possible to less than 1% precision at the present time.] Then one can predict all inclusive $B$ decays, as well as the $B^*$ and $D^*$ masses with no parameters. This would be the "best" method to use — unfortunately, we cannot accurately compute the desired quantities reliably in QCD. At present, it is better to use the HQET expansion in $1/m_b$ and $1/m_c$, with 6 parameters, and compute to order $1/m_Q^3$. In the (distant) future, it could well be that using full QCD, with no parameters, is the best method to use.

We have done a fit to the experimental data using 11 schemes: the $1S$, PS, pole, $\overline{\rm MS}$ and kinetic schemes expanding in $1/m_c$, not expanding in $1/m_c$ and using $\overline{m}_c(m_b)$, and finally, not expanding in $1/m_c$ and using the kinetic scheme for both $m_b$ and $m_c$. In addition, the PS and kinetic schemes introduce a scale $\mu$, which is sometimes called the factorization scale. We have also examined the factorization scale dependence which is present in these two schemes. We confirm the conclusions of Ref. [1], that the pole and $\overline{\rm MS}$ schemes are significantly worse than the threshold mass schemes, as expected theoretically. This holds regardless of whether or not one expands in $1/m_c$. We recommend that these schemes be avoided for high precision fits to inclusive $B$ decays. We also find that the PS scheme gives results comparable to those of the $1S$ scheme (both expanding and not expanding in $1/m_c$), and that the PS scheme results do not significantly depend on the choice of factorization scale. We compared the PS scheme with the 1S scheme in Ref. [1], and do not repeat the results here.

Based on the above discussion, we present our results in five different mass schemes, using:

1. $m_b^{1S}$ and expand $m_b-m_c$ in terms of HQET matrix elements [Scheme $1S_{\rm EXP}$],
2. $m_b^{1S}$ and $\overline{m}_c(m_b)$ and no time ordered products [Scheme $1S_{\rm NO}$],
3. $m_b^{\rm kin}(\mu=1\,{\rm GeV})$ and expand $m_b-m_c$ [Scheme ${\rm kin}_{\rm EXP}$],
4. $m_b^{\rm kin}(\mu=1\,{\rm GeV})$ and $\overline{m}_c(m_b)$ [Scheme ${\rm kin}_{\rm NO}$],
5. $m_b^{\rm kin}(\mu=1\,{\rm GeV})$ and $m_c^{\rm kin}(\mu=1\,{\rm GeV})$ [Scheme ${\rm kin}_{\rm UG}$]. (1)

Schemes $1S_{\rm EXP}$ and $\rm kin_{EXP}$ contain time ordered products at order $\Lambda_{\rm QCD}^3/m_b^3$, while they are absent from $1S_{\rm NO}$, $\rm kin_{NO}$, and $\rm kin_{UG}$. As discussed, the latter three schemes have the charm quark mass as an additional parameter at leading order in $\Lambda_{\rm QCD}/m_b$. Scheme $1S_{\rm EXP}$ is that used in Ref. [1], while scheme $\rm kin_{UG}$ is that used in Ref. [9].

## III. SHAPE VARIABLES AND THE DATA

We study three different distributions, the charged lepton energy spectrum [13–16] and the hadronic invariant mass spectrum [8, 15, 17, 18] in semileptonic $B \to X_c \ell \bar{\nu}$ decays, and the photon spectrum in $B \to X_s \gamma$ [19–22]. The theoretical predictions for these (as well as for the semileptonic $B \to X_c \ell \bar{\nu}$ rate [23]) are known to order $\alpha_s^2 \beta_0$ and $\Lambda_{\rm QCD}^3/m_b^3$, where $\beta_0 = 11 - 2n_f/3$ is the coefficient of the first term in the QCD $\beta$-function. For the $B \to X \ell \bar{\nu}$ branching rate, we use the average of the $B^\pm$ and $B^0$ data as quoted in the PDG [24],[1]

$$\mathcal{B}(B \to X \ell \bar{\nu}) \,=\, 10.73 \pm 0.28\,\%. \tag{2}$$

We apply a relative $-2\%$ correction to $\mathcal{B}(B \to X \ell \bar{\nu})$ to account for the $B \to X_u \ell \bar{\nu}$ fraction, and so use

$$\mathcal{B}(B \to X_c \ell \bar{\nu}) \,=\, 0.98\; \mathcal{B}(B \to X \ell \bar{\nu}). \tag{3}$$

The uncertainty of $|V_{ub}|$ is not a dominant error in $\mathcal{B}(B \to X_c \ell \bar{\nu})$. The fit result for $|V_{cb}|$ depends not only on $\mathcal{B}(B \to X_c \ell \bar{\nu})$, but also on the partial semileptonic branching ratios measured by the BABAR Collaboration [25], which have smaller errors than Eq. (2). The published BABAR results have already been corrected for $B \to X_u \ell \bar{\nu}$ contamination.

### A. Lifetime

The value of $|V_{cb}|$ depends on the $B$ meson lifetimes. The ratio of $B^+$ and $B^0$ lifetimes is $\tau_+/\tau_0 \,=\, 1.086 \pm 0.017$ [24]. Isospin violation in the $B$ meson semileptonic width is expected to be smaller than both $\tau_+/\tau_0$ and the uncertainties in the current analysis. The $\approx 8\%$ isospin violation in the lifetimes is probably due to the non-leptonic decay channels.

An additional source of isospin violation in the experimental measurements is through the production rates of $B^+$ and $B^0$ mesons, which is expected to be of the order of a few percent to perhaps as large as 10% [26]. Let $f_+$ and $f_0$ be the fraction of $B^+$ and $B^0$ mesons produced in $\Upsilon(4S)$ decay, with $f_+ + f_0 = 1$. Then the measured semileptonic branching ratios are

$$\mathcal{B} \,=\, f_+ \frac{\Gamma_{\rm sl}}{\Gamma_{\rm sl} + \Gamma_{{\rm nl},+}} + f_0 \frac{\Gamma_{\rm sl}}{\Gamma_{\rm sl} + \Gamma_{{\rm nl},0}} \tag{4}$$

where $\mathcal{B}$ and $\Gamma_{\rm sl}$ are computed with the same lepton energy cut. In writing Eq. (4), we have used the fact that isospin violation in the semileptonic rates is small, so that the same value for $\Gamma_{\rm sl}$ is used for both $B^+$ and $B^0$.

The measured semileptonic branching ratios can thus be written as

$$\mathcal{B} \,=\, \tau_{\rm eff}\; \Gamma_{\rm sl} \tag{5}$$

in terms of the effective lifetime

$$\tau_{\rm eff} \,=\, f_+ \tau_+ + f_0 \tau_0. \tag{6}$$

One can rewrite this as

$$\tau_{\rm eff} \,=\, \frac{\tau_+ + \tau_0}{2} + \frac{(f_+ - f_0)(\tau_+ - \tau_0)}{2}. \tag{7}$$

Using the PDG 2004 lifetime values, and the measured $f_+/f_0$ ratio [27] gives

$$\tau_{\rm eff} \,=\, 1.60 \pm 0.01 \text{ ps} \tag{8}$$

where the contribution from the second term in Eq. (7) is negligible to both the value and the error.

### B. Lepton Moments

For the charged lepton energy spectrum we define the integrals

$$R_n(E_{\rm cut}, \mu) = \int_{E_{\rm cut}} (E_\ell - \mu)^n \; \frac{{\rm d}\Gamma}{{\rm d}E_\ell} \, {\rm d}E_\ell \,, \tag{9}$$

where ${\rm d}\Gamma/{\rm d}E_\ell$ is the spectrum in the $B$ rest frame and $E_{\rm cut}$ is a lower cut on the lepton energy. Moments of the lepton energy spectrum with a lepton energy cut $E_{\rm cut}$ are given by

$$\langle E_\ell^n \rangle_{E_{\rm cut}} = \frac{R_n(E_{\rm cut}, 0)}{R_0(E_{\rm cut}, 0)} \,, \tag{10}$$

and central moments by

$$\langle (E_\ell - \langle E_\ell \rangle)^n \rangle_{E_{\rm cut}} = \frac{R_n(E_{\rm cut}, \langle E_\ell \rangle_{E_{\rm cut}})}{R_0(E_{\rm cut}, 0)} \,, \tag{11}$$

which can be determined as a linear combination of the non-central moments.

The BABAR Collaboration [25] measures the partial branching fraction $\tau_B R_0(E_{\rm cut}, 0)$, the mean lepton energy $\langle E_\ell \rangle_{E_{\rm cut}}$, and the second and third central moments $\langle (E_\ell - \langle E_\ell \rangle)^n \rangle_{E_{\rm cut}}$ for $n = 2, 3$, each for lepton energy cuts of $E_{\rm cut} = 0.6,\ 0.8,\ 1.0,\ 1.2,\ 1.5$ GeV.

[1] It would be inconsistent to use the average $b$ hadron semileptonic rate (including $B_s$ and $\Lambda_b$ states), since hadronic matrix elements have different values in the $B/B^*$ system, and in the $B_s/B_s^*$ or $\Lambda_b$.

The CLEO Collaboration [3, 28] measures the mean lepton energy and second central moment (variance) for $E_{\rm cut} = 0.7 - 1.6$ GeV in steps of 0.1 GeV.

The DELPHI Collaboration [29] measures the mean lepton energy, and the $n = 2, 3$ central moments, all with no energy cut.

In total, we have 43 experimental quantities from the lepton moments, 20 from BABAR, 20 from CLEO, and 3 from DELPHI.

### C. Hadron Moments

For the $B \to X_c \ell \bar{\nu}$ hadronic invariant mass spectrum, we define

$$\langle m_X^{2n} \rangle_{E_{\rm cut}} = \frac{\displaystyle\int_{E_{\rm cut}} (m_X^2)^n \, \frac{{\rm d}\Gamma}{{\rm d}m_X^2} \, {\rm d}m_X^2}{\displaystyle\int_{E_{\rm cut}} \frac{{\rm d}\Gamma}{{\rm d}m_X^2} \, {\rm d}m_X^2} \,, \qquad (12)$$

where $E_{\rm cut}$ is again the cut on the lepton energy. Sometimes $\overline{m}_D^2 \equiv [(m_D + 3m_{D^*})/4]^2$ is subtracted out in the definitions, $\langle (m_X^2 - \overline{m}_D^2)^n \rangle$, or the measurements of the normal moments are quoted, $\langle (m_X^2 - \langle m_X^2 \rangle)^n \rangle$, but these can easily be computed from $\langle m_X^{2n} \rangle$.

The BABAR Collaboration [30] has measured the mean values of $m_X$, $m_X^2$, $m_X^3$ and $m_X^4$ (i.e., $n = 1/2,\ 1,\ 3/2,\ 2$ moments) for lepton energy cuts $E_{\rm cut} = 0.9 - 1.6$ GeV in steps of 0.1 GeV.

The CDF Collaboration [31] measures the mean value of $m_X^2$ and its variance, with no lepton energy cut.

The CLEO Collaboration [32] has measured the mean value of $m_X^2 - \overline{m}_D^2$ and the variance of $m_X^2$ for lepton energy cuts of 1.0 and 1.5 GeV.

The DELPHI Collaboration [29] has measured the mean value of $m_X^2 - \overline{m}_D^2$, $(m_X^2 - \overline{m}_D)^2$, the variance of $m_X^2$, and the third central moment of $m_X^2$, all with no energy cut.

Recently half-integer moments of the $m_X^2$ spectrum [8, 9] have received some attention. While non-integer moments of the lepton energy spectrum have been computed in a power series in $1/m_b$ [33], this is not true for fractional moments of the $m_X^2$ spectrum. In [8, 9] expressions for the half-integer moments were proposed which involve expansions that were claimed to be in powers of $\Lambda_{\rm QCD}/m_c$. However, in the limit $m_c \ll m_b$ (i.e., $m_c$ of order $\sqrt{m_b \Lambda_{\rm QCD}}$ or less), the higher order terms in these expansion scale with powers of $m_b \Lambda_{\rm QCD}/m_c^2$, which in this limit is of order unity or larger. On the other hand, in the small velocity limit, $m_b \sim m_c \gg m_b - m_c \gg \Lambda_{\rm QCD}$, the expansion of $\langle m_X^{2n+1} \rangle$ is well-behaved. Thus, the calculations of the half-integer moments as presented in [8, 9] do not correspond to a power series in $1/m_{b,c}$ in the $m_c \ll m_b$ limit and omitted terms are only power suppressed in the small velocity limit. In addition, the BLM corrections to these moments are currently unknown, because they require the BLM contribution from the virtual terms, which have not been computed. For these reasons we will not use these half-integer moments in the fit, but will compare the fit results with the measured values. Omitting the half-integer moments, there are 16 data points from BABAR, 2 from CDF, 4 from CLEO, and 4 from DELPHI, for a total of 26 measurements.

### D. Photon Spectrum

For $B \to X_s \gamma$, we define

$$\langle E_\gamma^n \rangle_{E_{\rm cut}} = \frac{\displaystyle\int_{E_{\rm cut}} E_\gamma^n \, \frac{{\rm d}\Gamma}{{\rm d}E_\gamma} \, {\rm d}E_\gamma}{\displaystyle\int_{E_{\rm cut}} \frac{{\rm d}\Gamma}{{\rm d}E_\gamma} \, {\rm d}E_\gamma} \,, \qquad (13)$$

where ${\rm d}\Gamma/{\rm d}E_\gamma$ is the photon spectrum in the $B$ rest frame, and $E_{\rm cut}$ is the photon energy cut. In this case the variance, $\langle (E_\gamma - \langle E_\gamma \rangle)^2 \rangle = \langle E_\gamma^2 \rangle - \langle E_\gamma \rangle^2$, is often used instead of the second moment, and higher moments are not used as they are very sensitive to the boost of the $B$ meson in the $\Upsilon$ rest frame (though this is absent if ${\rm d}\Gamma/{\rm d}E_\gamma$ is reconstructed from a measurement of ${\rm d}\Gamma/{\rm d}E_{m_{X_s}}$) and to the details of the shape function. $\langle E_\gamma^n \rangle$ are known to order $\alpha_s^2 \beta_0$ [21] and $\Lambda_{\rm QCD}^3/m_b^3$ [22]. These moments are expected to be described by the OPE once $m_B/2 - E_{\rm cut} \gg \Lambda_{\rm QCD}$. Precisely how low $E_{\rm cut}$ has to be to trust the results can only be decided by studying the data as a function of $E_{\rm cut}$; one may expect that $E_{\rm cut} = 1.8\,{\rm GeV}$ available at present [34] is sufficient. Note that the perturbative corrections included are sensitive to the $m_c$-dependence of the $b \to c\bar{c}s$ four-quark operator ($O_2$) contribution. This is a particularly large effect in the $O_2 - O_7$ interference [21], but its relative influence on the moments of the spectrum is less severe than that on the total decay rate.

We use the BELLE [34], and CLEO [35] measurements of the mean photon energy and variance, with photon energy cuts of 1.8 and 2.0 GeV, respectively, and the BABAR measurement [36] of the mean photon energy with a cut of 2.094 GeV for a total of 5 measurements.

## IV. FIT PROCEDURE

As discussed in Sec. II, there are many ways to treat the quark masses and hadronic matrix elements that occur in the OPE results for the spectra. In the schemes where $m_b - m_c$ is expanded in HQET (such as $1S_{\rm EXP}$ and $\rm kin_{EXP}$), the theoretical expressions for the shape variables defined in Eqs. (9), (12), and (13) include 17

terms

$$\begin{aligned} X_{E_{\rm cut}} &= X^{(1)} + X^{(2)}\,\Lambda + X^{(3)}\,\Lambda^2 + X^{(4)}\,\Lambda^3 + X^{(5)}\,\lambda_1 \\ &+ X^{(6)}\,\lambda_2 + X^{(7)}\,\lambda_1\Lambda + X^{(8)}\,\lambda_2\Lambda + X^{(9)}\,\rho_1 \\ &+ X^{(10)}\,\rho_2 + X^{(11)}\,\mathcal{T}_1 + X^{(12)}\,\mathcal{T}_2 + X^{(13)}\,\mathcal{T}_3 \\ &+ X^{(14)}\,\mathcal{T}_4 + X^{(15)}\,\epsilon + X^{(16)}\,\epsilon^2_{\rm BLM} + X^{(17)}\,\epsilon\,\Lambda\,, \end{aligned} \tag{14}$$

while in the schemes when $m_c$ is treated as an independent free parameter (such as $1S_{\rm NO}$, $\rm kin_{NO}$, and $\rm kin_{UG}$), we have 22 terms

$$\begin{aligned} Y_{E_{\rm cut}} &= Y^{(1)} + Y^{(2)}\,\Lambda + Y^{(3)}\,\Lambda_c + Y^{(4)}\,\Lambda^2 + Y^{(5)}\,\Lambda\Lambda_c \\ &+ Y^{(6)}\,\Lambda_c^2 + Y^{(7)}\,\Lambda^3 + Y^{(8)}\,\Lambda^2\Lambda_c + Y^{(9)}\,\Lambda\Lambda_c^2 \\ &+ Y^{(10)}\,\Lambda_c^3 + Y^{(11)}\,\lambda_1 + Y^{(12)}\,\lambda_2 + Y^{(13)}\,\lambda_1\Lambda \\ &+ Y^{(14)}\,\lambda_2\Lambda + Y^{(15)}\,\lambda_1\Lambda_c + Y^{(16)}\,\lambda_2\Lambda_c + Y^{(17)}\,\rho_1 \\ &+ Y^{(18)}\,\rho_2 + Y^{(19)}\,\epsilon + Y^{(20)}\,\epsilon^2_{\rm BLM} + Y^{(21)}\,\epsilon\,\Lambda \\ &+ Y^{(22)}\,\epsilon\,\Lambda_c\,. \end{aligned} \tag{15}$$

In Eqs. (14) and (15) $\Lambda$ and $\Lambda_c$ are respectively the differences between the $b$ and $c$ quark masses and their reference values about which we expand. The coefficients $X^{(k)}$ and $Y^{(k)}$ are functions of $E_{\rm cut}$, and $(X_{E_{\rm cut}}, Y_{E_{\rm cut}})$ are any of the experimental observables discussed earlier. The parameter $\epsilon \equiv 1$ counts powers of $\alpha_s$. We have used $\alpha_s(m_b) = 0.22$. The strong coupling constant is not a free parameter, but is determined from other measurements such as the hadronic width of the $Z$. The hadron and lepton moments are integrals of the same triple differential decay rate with different weighting factors. The use of different values of $\alpha_s$ for the hadron and lepton moments, as done in Ref. [9], is an ad hoc choice.

## V. THE FIT

We use the program MINUIT to perform a global fit to all observables introduced in Sec. III in each of the 11 schemes mentioned in Sec. II. There are a total of 74 lepton, hadron, and photon moments, plus the semileptonic width, to be fit using 7 parameters, so the fit has $\nu = 68$ degrees of freedom.

To evaluate the $\chi^2$ required for the fit, we include both experimental and theoretical uncertainties. For the experimental uncertainties we use the full correlation matrix for the observables from a given differential spectrum as published by the experimental collaborations. In addition to these experimental uncertainties there are theoretical uncertainties, which correspond to how well we expect to be able to compute each observable theoretically. For a given observable, our treatment of theoretical uncertainties is similar to that in Ref. [1].

It is important to include theoretical uncertainties in the fit, since not all quantities can be computed with the same precision. We have treated theoretical errors as though they have a normal distribution with zero mean, and standard deviation equal to the error estimate.[2] Strictly speaking, the theoretical formula has some definite higher order correction, which is at present unknown. One can then view the normal distribution used for the theory value as the prior distribution in a Bayesian analysis. The way in which theoretical errors are included is a matter of choice, and there is no unique prescription.

We now discuss in detail the theoretical uncertainties included in the fit. Those who find this procedure abhorrent can skip the entire discussion, since we will also present results not including theory errors.

### A. Theory uncertainties

Theoretical uncertainties in inclusive observables as discussed here originate from four main sources. First, there are uncertainties due to uncalculated power corrections. For schemes $1S_{\rm EXP}$ and $\rm kin_{EXP}$, these are of order $\Lambda^4_{\rm QCD}/(m_b^2 m_c^2) \sim 0.001$, while for schemes $1S_{\rm NO}$ and $\rm kin_{NO}$ where no $1/m_c$ expansion is performed, these are of order $\Lambda^4_{\rm QCD}/m_b^4 \sim 0.0001$. Next, there are uncertainties due to uncalculated higher order perturbative terms. In particular, the full two loop result proportional to $\alpha_s^2/(4\pi)^2 \sim 0.0003$ is not available. An alternative way to estimate these perturbative uncertainties is by the size of the smallest term computed in the series, which is the term proportional to $\alpha_s^2\beta_0$. We choose here to use half of this last computed term as an estimate of the uncertainty. There are also uncalculated effects of order $(\alpha_s/4\pi)\Lambda^2_{\rm QCD}/m_b^2 \sim 0.0002$. Finally, there is an uncertainty originating from effects not included in the OPE in the first place. Such effects sometimes go under the name "duality violation," and are very hard to quantify. For this reason, we do not include an explicit contribution to the overall theoretical uncertainty from such effects. If duality violation would be larger than the other theoretical uncertainties they would give rise to a poor fit to the data. To determine the uncertainties for dimensionful quantities such as the moments considered here, we have to multiply these numbers by the appropriate dimensionful quantity. This number is obtained from dimensional analysis, and we use for the $n$'th hadronic moment $(m_B^2)^n f_n$, while we use $(m_B/2)^n f_n$ for the $n$'th leptonic moment. The factors $f_n$ are chosen to be $f_0 = f_1 = 1$, $f_2 = 1/4$ and $f_3 = 1/(6\sqrt{3})$. The values for $f_2$ and $f_3$ are the maximum allowed values for the sec-

[2] This is the same procedure as that used by CODATA in doing a fit to the fundamental constants [37]. An example which makes clear why theoretical errors should be included is: The Hydrogen hyperfine splitting is measured to 14 digits, but has only ben computed to 7 digits. The Positronium hyperfine splitting is measured and computed to 8 digits. It would not be proper to give the H hyperfine splitting a weight $10^6$ larger than the Ps hyperfine splitting in a global fit to the fundamental constants.

ond and third central moments (variance and skewness) for a probability distribution on the interval $[0,1]$.

The complete BLM piece has not been computed for the non-integer hadronic moments. The perturbative uncertainty is therefore dominated by this contribution of order $\beta_0\alpha_s^2/(4\pi)^2 \sim 0.003$. We will use $A \sim 0.003$ for the non-integer hadronic moments when we compare experiment with theory.

For the hadronic mass and lepton energy moments, which depend on the value of the cut on the lepton energy, we have to decide how to treat the correlation of the theoretical uncertainties. In the global fit by the BABAR Collaboration [4], the theory errors for a given observable with different cuts on $E_\ell$ were treated as 100% correlated. This ignores the fact that the higher order terms omitted in the OPE depend on the lepton energy cut. In Ref. [1], only the two extreme values of the lepton energy cut were included in the fit, and the correlation of the theory uncertainties was neglected. Here we take the correlation of the theoretical uncertainties to be given by the correlation between the experimental measurements, which captures the correlations due to the fact that observables with different cuts share some common events.

For the photon energy moments an additional source of uncertainty is the fact that the presence of any experimentally sensible value for $E_{\rm cut}$ affects the mean photon energy $\langle E_\gamma\rangle$ such that the extracted value of $m_b$ is biased toward larger values because of shape function effects [22]. However, this shift cannot be calculated model independently. Rather than include a model dependence, we have multiplied the theory uncertainties for the $b\to s\gamma$ rates by the ratios of the energy difference from the endpoint, relative to that for BELLE with $E_{\rm cut} = 1.8$ GeV.[3]

To summarize, we define the combined experimental and theoretical error matrix for a given observable to be

$$\sigma_{ij}^2 = \sigma_i\,\sigma_j\,c_{ij}\,, \qquad \text{no sum on } i,j\,, \tag{16}$$

where $i$ and $j$ denote observables, $c_{ij}$ is the experimental correlation matrix, and

$$\begin{aligned}
\sigma_i &= \sqrt{(\sigma_i^{\rm exp})^2 + \left(A\,f_n m_B^{2n}\right)^2 + \left(B_i/2\right)^2} \\
&\quad \text{for the n'th hadron moment}\,, \\
\sigma_i &= \sqrt{(\sigma_i^{\rm exp})^2 + \left(A\,f_n (m_B/2)^n\right)^2 + \left(B_i/2\right)^2} \\
&\quad \text{for the n'th lepton moment}\,, \\
\sigma_i &= \sqrt{(\sigma_i^{\rm exp})^2 + \left(A\,f_n (m_B/2)^n\right)^2 + \left(B_i/2\right)^2} \\
&\quad \text{for the n'th photon moment}\,,
\end{aligned} \tag{17}$$

and $f_0 = f_1 = 1$, $f_2 = 1/4$, $f_3 = 1/(6\sqrt{3})$. Here $\sigma_i^{\rm exp}$ are the experimental errors, $B_i = X^{(16)}$ or $Y^{(20)}$ are the coefficients of the last computed terms in the perturbation series, and $A$ contains the errors discussed earlier. We take $A = 0.001$ for the data used in the fit, except for the CLEO and BABAR photon moments, where we multiply $A$ by 1.3 and 1.5, respectively, to account for the increase in shape function effects as one limits the allowed region of the photon spectrum.

We stress that there is no unique way to estimate theoretical uncertainties to a given expression. Thus, while we believe that our estimates are reasonable, it is certainly not the only possible way to estimate the theory uncertainties (e.g., taking the theory correlation to be identical to the experimental correlations is just an educated guess).

### B. Experimental correlations

Some of the experimental correlation matrices have negative eigenvalues. In some cases, these are at the level of round-off errors. To avoid these negative values, we have added 0.01 to the diagonal entries for the correlation matrices for the BABAR and CLEO lepton moments, and the DELPHI hadron moments.

The correlation matrix for the BABAR hadronic moments [4] contains negative eigenvalues which are much larger than any round-off uncertainties. This persists even if only every second value of the cut is used, as advocated in [4], so we are forced to add 0.05 to the diagonal entries of the correlation matrix for the BABAR hadron moments to make the eigenvalues positive. Note that the correlation matrix can have negative eigenvalues only if the probability distribution can take on negative values.

### C. Constraints on parameters

Even though there are many more observables than there are parameters, the fit does not provide strong constraints on the $1/m_b^3$ parameters. Thus it is useful to add additional information to ensure that the fit converges to physically sensible values of the nonperturbative parameters. Thus, as in Ref. [1] we add to $\chi^2$ the contribution

$$\chi^2_{\rm param}(m_\chi, M_\chi) = \begin{cases} 0\,, & |\langle\mathcal{O}\rangle| \le m_\chi^3\,, \\ \left[|\langle\mathcal{O}\rangle| - m_\chi^3\right]^2/M_\chi^6\,, & |\langle\mathcal{O}\rangle| > m_\chi^3\,, \end{cases} \tag{18}$$

where $(m_\chi, M_\chi)$ are both quantities of order $\Lambda_{\rm QCD}$, and $\langle\mathcal{O}\rangle$ are the matrix elements of any of the $1/m^3$ operators in the fit. This way we do not prejudice $\langle\mathcal{O}\rangle$ to have any particular value in the range $|\langle\mathcal{O}\rangle| \le m_\chi^3$. In the fit we take $M_\chi = m_\chi = 500\,$MeV. We have already checked in Ref. [1] that the results for $|V_{cb}|$ and $m_b$ are insensitive to varying $m_\chi$ between 500 MeV and 1 GeV. The data are sufficient to constrain the $1/m_b^3$ operators in the sense that they can be consistently fit with reasonable values, but they are not determined with any useful precision. The data can be fit without including $\chi^2_{\rm param}$, but then

[3] The $B\to X_s\gamma$ photon spectrum also receives contributions of order $m_s^2/m_b^2$, which are negligible corrections for our analysis.

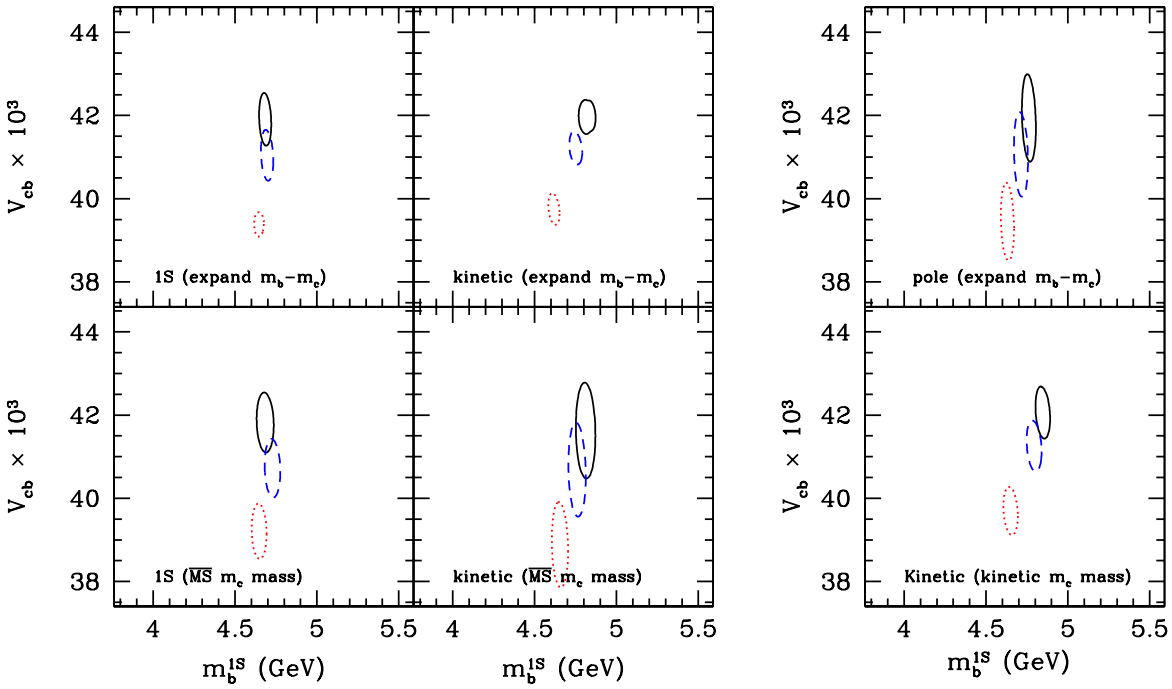


FIG. 1: Fit results for $|V_{cb}|$ and $m_b$ in the $1S_{\rm EXP}$, $1S_{\rm NO}$, $\rm kin_{EXP}$, $\rm kin_{NO}$, and $\rm kin_{UG}$ schemes defined in Eq. (1) and in the traditional pole scheme. The red doted, blue dashed, and black solid ellipses denote the results at tree level, order $\alpha_s$, and $\alpha_s^2\beta_0$, respectively, corresponding to $\Delta\chi^2 = 1$.

| Scheme | $\sigma^2_{\rm theory}$ | $\chi^2/\nu$ | $\lvert V_{cb}\rvert \times 10^3$ | $m_b^{1S}$ [GeV] | $m_b - m_c$ [GeV] | $\overline{m}_c(\overline{m}_c)$ [GeV] | $\lambda_1$ [GeV$^2$] |
|---|---|---|---|---|---|---|---|
| $1S_{\rm EXP}$ | yes | 41.7/68 | $41.9 \pm 0.6$ | $4.68 \pm 0.04$ | $3.40 \pm 0.01$ | $1.07 \pm 0.04$ | $-0.23 \pm 0.06$ |
| $\rm kin_{EXP}$ | yes | 41.9/68 | $41.7 \pm 0.6 \pm 0.1$ | $4.69 \pm 0.04 \pm 0.03$ | $3.39 \pm 0.01 \pm 0.01$ | $1.09 \pm 0.04 \pm 0.03$ | $-0.17 \pm 0.05 \pm 0.07$ |
| $1S_{\rm EXP}$ | no | 98.6/68 | $42.1 \pm 0.3$ | $4.69 \pm 0.02$ | $3.39 \pm 0.01$ | $1.09 \pm 0.02$ | $-0.27 \pm 0.04$ |
| $\rm kin_{EXP}$ | no | 126.5/68 | $42.1 \pm 0.3 \pm 0.2$ | $4.74 \pm 0.02 \pm 0.06$ | $3.36 \pm 0.01 \pm 0.04$ | $1.15 \pm 0.02 \pm 0.12$ | $-0.29 \pm 0.04 \pm 0.14$ |

TABLE I: Fit results for $|V_{cb}|$, $m_b$, $m_c$ and $\lambda_1$ in the $1S_{\rm EXP}$ and $\rm kin_{EXP}$ schemes. Our fits in the kinetic scheme use $\mu_\pi^2$, but the fit result has been converted to $\lambda_1$ to help comparison. The first two lines are the fit results including our estimates of the theoretical errors, the lower two lines are setting these to zero. The second error for the $\rm kin_{EXP}$ scheme is the shift due to changing $\mu$ from 1 to 1.5 GeV. The $|V_{cb}|$ value includes electromagnetic radiative corrections.

some of the $1/m_b^3$ parameters are not of natural size, with values of order 0.5 GeV$^3$. Including $\chi^2_{\rm param}$ gives a fit with reasonable values of the parameters, of order 0.1 GeV$^3$. The contribution of $\chi^2_{\rm param}$ is rather small, of order $0.1 - 0.2$, so that $\chi^2_{\rm param}$ does not drive the fit. This shows that there are some very flat directions in parameter space which are stabilized by including $\chi^2_{\rm param}$. We have shown our final results for $V_{cb}$ and $m_b$ with and without including $\chi^2_{\rm param}$ in the fit. The final results do not depend significantly on whether or not $\chi^2_{\rm param}$ is included.

Note that the fit performed by the BABAR Collaboration included the half-integer hadronic moments. We have checked that including these moments still leaves some $1/m_b^3$ parameters with values which are larger than natural size. We have chosen to not include these moments in the fit since they have large theoretical uncertainties.

## VI. FIT RESULTS AND DISCUSSION

The fit result for $|V_{cb}|$ and $m_b$ in the five mass schemes defined in Eq. (1) and in the traditional pole scheme are shown in Fig. 1. The fit results are shown at tree level, order $\alpha_s$, and order $\alpha_s^2\beta_0$. The kinetic scheme results are obtained using $\mu_\pi^2$ etc. in the fit, and then converting the results back to $\lambda_1$ etc. for easier comparison with the other schemes. One can see that the $1S_{\rm EXP}$ and $\rm kin_{EXP}$ schemes have better convergence than the pole scheme.

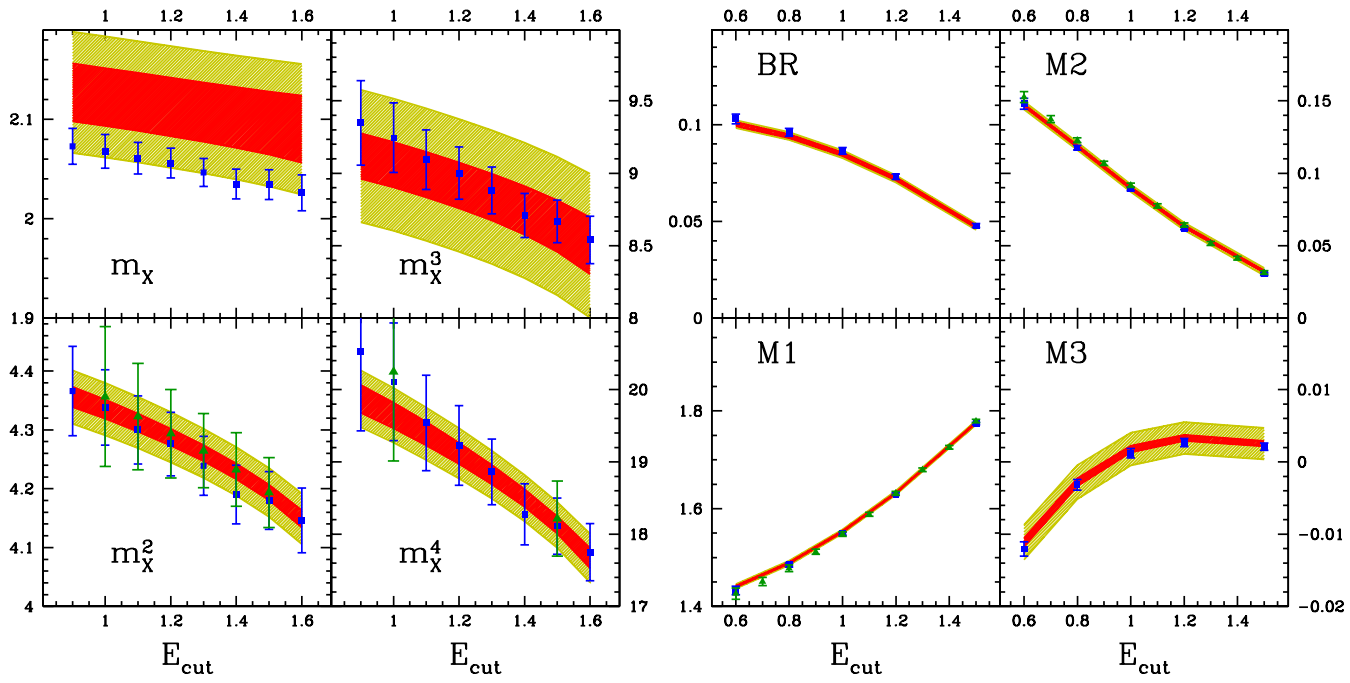


FIG. 2: Measurements (blue squares: BABAR [30], green triangles: CLEO [32]) and fit results for the hadron invariant mass (left) and the lepton energy moments (right) as functions of the lepton energy cut, $E_{\rm cut}$. For the hadron moments $m_X^n$ denotes $\langle m_X^n\rangle$, while for the lepton moments BR is branching ratio, M1 is first moment, and M2 and M3 are the second and third central moments, respectively. The red (dark) shaded regions show the fit error, while the yellow (light) shaded regions are our estimates of the theoretical uncertainties from the $A$ terms in Eq. (17). The $A$ term for $\langle m_X\rangle$ and $\langle m_X^3\rangle$ is three times larger than for $\langle m_X^2\rangle$ and $\langle m_X^4\rangle$, because of the worse expansion for the non-integer moments.

The main fit results in the $1S_{\rm EXP}$ and kin$_{\rm EXP}$ scheme are given in Table I. The remarkable agreement between the fit results shows that the main difference in the fits is not which short distance $b$ quark mass is used, but whether $m_b - m_c$ is or is not qqexpanded in terms of HQET matrix elements.

The uncertainties for the $1S_{\rm EXP}$ and kin$_{\rm EXP}$ schemes, which eliminate $m_c$, are smaller than for the $1S_{\rm NO}$ and kin$_{\rm NO}$ schemes, which use $\overline{m}_c(m_b)$. This is contrary to the claims made in [9], but is not unexpected, since the former schemes have only one parameter at leading order in $1/m_b$, while the latter schemes have two such parameters. While not expanding in $1/m_c$ gives slightly larger errors than expanding, the consistency of the central values between the two methods shows that one can use the $1/m_c$ expansion for inclusive $B$ decays.

One can clearly see that using the kinetic mass for $m_c$ (the kin$_{\rm UG}$ scheme) does not reduce the uncertainties compared to the $1S_{\rm EXP}$ and kin$_{\rm EXP}$ schemes. Also, as is now well known, the pole scheme does not work as well in inclusive calculations as the schemes which use a short distance mass. Thus, in the remainder of this work we will present results in the $1S_{\rm EXP}$ and in the kin$_{\rm EXP}$ schemes. We have carried out the fits in 6 additional schemes, including the PS and $\overline{\rm MS}$ schemes. All of the schemes give reasonable fits, but only the PS scheme with $m_b - m_c$ expanded in HQET gives rise to similarly small uncertainties as $1S_{\rm EXP}$ and kin$_{\rm EXP}$.

The charm quark mass enters into the computation, and we can extract the value of $m_c$ from our fit. The value of $m_b - m_c$, which is free of the order $\Lambda_{\rm QCD}$ renormalon ambiguity, is (in the $1S_{\rm EXP}$ scheme)

$$m_b - m_c = 3.40 \pm 0.01 \ {\rm GeV}\,. \qquad (19)$$

We can convert this result to the $\overline{\rm MS}$ mass of the charm quark,

$$\begin{aligned} \overline{m}_c(\overline{m}_c) &= 0.90 \pm 0.04 \ {\rm GeV}\,, \\ \overline{m}_c(\overline{m}_c) &= 1.07 \pm 0.04 \ {\rm GeV}\,, \end{aligned} \qquad (20)$$

where the two results depend on whether the perturbative conversion factor is reexpanded or not.[4] The reason for the large difference between the two results is that perturbative corrections are large at the scale $\overline{m}_c$. Taking the average of the two $m_c$ values, and adding half the

[4] I.e., the difference between dividing by $1 + a_1\alpha_s + a_2\alpha_s^2$ and multiplying by $1 - a_1\alpha_s + (a_1^2 - a_2)\alpha_s^2$. Only the larger value in Eq. (20) has been shown in Table I.

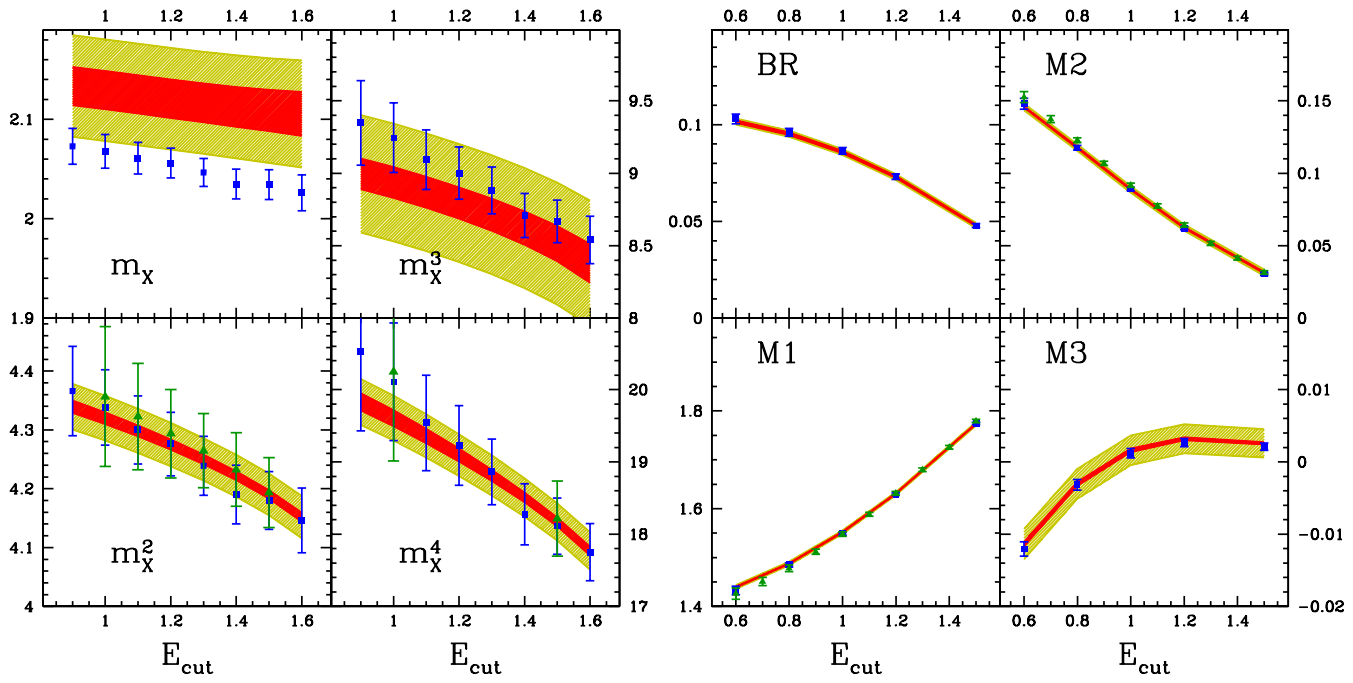


FIG. 3: Measurements and fit results for the hadron invariant mass and the lepton energy moments, setting all theory errors to zero in the fit. (See the caption for Fig. 2.) The yellow (light) shaded band gives the estimated theoretical uncertainty, as in Fig 2. It is not included in the fit, but it can help to decide the significance of any differences between theory and experiment.

difference between them as an additional error gives

$$\overline{m}_c(\overline{m}_c) = 0.99 \pm 0.1 \text{ GeV}\,. \tag{21}$$

The difference between the $m_c$ values is a nice illustration that one should avoid using perturbation theory at a low scale, if at all possible. The $\text{kin}_{\text{UG}}$ scheme uses perturbation theory at a scale below $m_c$, and suffers from the same problem.

Next, we compare how well the theory can reproduce the experimental measurements, focusing on the cut dependence of individual moments. The results for the hadronic moments and the leptonic moments are shown in Fig. 2 in the $1S_{\text{EXP}}$ scheme. (The DELPHI and CDF results are included in the fits, but are not shown, as they correspond to $E_{\text{cut}} = 0$.). The red (dark) shaded band is the uncertainty due to the errors on the fit parameter. The width of the yellow (light) shaded band is the theoretical uncertainty due to higher order non-perturbative effects not included in the computation [the $A$ term in Eq. (17)]. Within the uncertainties, the OPE predictions for all these moments agree well with the data. As we explained before, the moments $\langle m_X \rangle$ and $\langle m_X^3 \rangle$ were not included in the fit. The yellow bands shown for $\langle m_X \rangle$ and $\langle m_X^3 \rangle$ use $A = 0.003$ as an estimate of the uncertainty, a factor of three larger than for the integer moments, because of the worse theoretical expansion discussed in Sec. III C.

The agreement between the theory and experiment for the third lepton moment is better than our estimate of the theoretical uncertainty. This might be an indication that we overestimate the theoretical uncertainty for this moment.

The $\chi^2$ for the fit shows that the theory provides an excellent description of the data. In the $1S_{\text{EXP}}$ scheme, we get $\chi^2 = 41.7$ for $\nu = 68$ degrees of freedom, so $\chi^2/\nu = 0.61$. The standard deviation for $\chi^2/\nu$ is $\sqrt{2/\nu} = 0.17$, so that $\chi^2/\nu = 0.61$ is about two standard deviations below the mean value of 1.01. This is some evidence that the theoretical errors have been overestimated. To study the effect of the theoretical uncertainties, we also perform fits with all theoretical uncertainties set to zero. This fit gives $\chi^2 = 98.6$ for the $1S_{\text{EXP}}$ scheme, and $\chi^2 = 132.1$ for the $\text{kin}_{\text{EXP}}$ scheme. The resulting fits still agree well with the experimental data, as can be seen from Fig. 3. The fit results with no theory error are given in the lower half of Table I. The values $\chi^2/\nu = 1.45$ for the $1S_{\text{EXP}}$ scheme and $\chi^2/\nu = 1.86$ for the $\text{kin}_{\text{EXP}}$ scheme are significantly greater than one, which is some evidence that there are higher order theoretical effects which have not been included.

The calculations in the $\text{kin}_{\text{UG}}$ scheme [9] were used by the BABAR Collaboration [4], to perform a fit to its own data. While we agree with the results of Ref. [9] for the lepton energy moments, we are unable to reproduce their results for the hadronic invariant mass moments. One should also note that Ref. [9] (i) uses $\alpha_s = 0.22$ for the lepton moments, and $\alpha_s = 0.3$ for the hadron moments

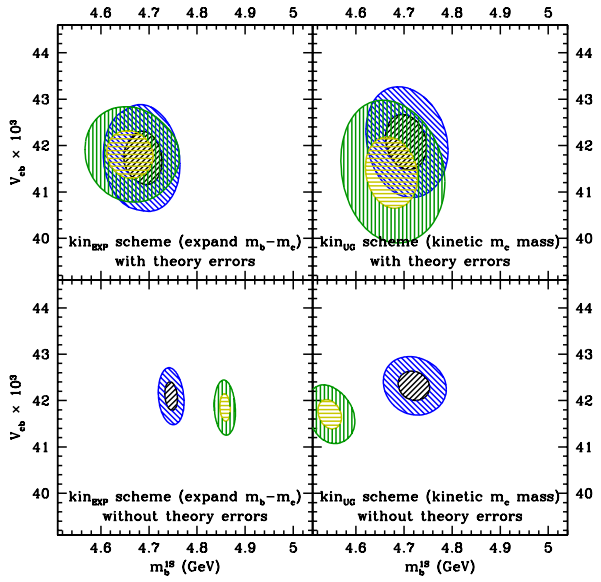


FIG. 4: Fit results for $|V_{cb}|$ and $m_b$ in the $\mathrm{kin}_{\mathrm{EXP}}$ and $\mathrm{kin}_{\mathrm{UG}}$ schemes using $\mu_b = 1$ GeV (blue and black) and using $\mu_b = 1.5$ GeV (green and yellow). $\mu_c$ for the $\mathrm{kin}_{\mathrm{UG}}$ scheme has been kept fixed at 1 GeV. The regions correspond to $\Delta\chi^2 = 1$ (black and yellow) and 4 (blue and green). The upper plots includes theory errors in the fit, and the lower plot does not.

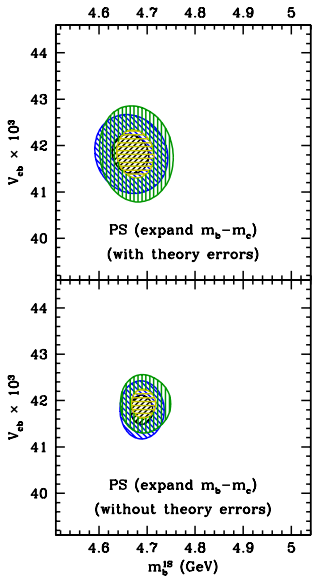


FIG. 5: Fit results for $|V_{cb}|$ and $m_b$ in the PS scheme using $\mu = 2$ GeV (blue and black) and using $\mu = 1.5$ GeV (green and yellow) The regions correspond to $\Delta\chi^2 = 1$ (black and yellow) and 4 (blue and green). The upper plots includes theory errors in the fit, and the lower plot does not.

(ii) includes the $\alpha_s^2\beta_0$ corrections (which are known for both the lepton and integer hadron moments) only in the lepton moments, but not in the hadron moments.

The $\mathrm{kin}_{\mathrm{EXP}}$ and $\mathrm{kin}_{\mathrm{NO}}$ schemes depend on a choice for $\mu_b$. In the $\mathrm{kin}_{\mathrm{UG}}$ scheme there is an additional dependence on $\mu_c$, and there is no reason why the theoretical predictions should be expanded using $m_b^{\mathrm{kin}}(\mu_b)$ and $m_c^{\mathrm{kin}}(\mu_c)$ defined at the same scale ($\mu_b = \mu_c$), since all that is required is that each $\mu$ should be small, so one has to choose both $\mu_b$ and $\mu_c$. To illustrate the sensitivity to the choice of $\mu_{b,c}$ we show a fit in Fig. 4 varying $\mu_b$ from 1 to 1.5 GeV keeping $\mu_c = 1$ GeV fixed. Clearly there is significant dependence in the kinetic schemes with respect to changes in $\mu$, and this should be included as an additional uncertainty for that scheme. We have included this scale uncertainty in Table I. The kinetic schemes use perturbation theory at a low scale $\mu_{b,c}$, and so are sensitive to precisely how these corrections are included, as was the case for $\overline{m}_c(\overline{m}_c)$. The PS scheme is much less sensitive to the value of $\mu$. In Fig. 5, we show the variation with $\mu$ in the PS scheme. Note that one advantage of the $1S$ scheme is that it does not depend on any factorization scale parameter $\mu$.

Reference [9] quotes smaller theoretical errors than the estimates used here — as much as a factor of ten smaller in some cases, as can be seen from the plots in Ref. [4]. We do not believe that this optimistic estimate of the theoretical uncertainty is justified.

Figure 6 shows the results for $|V_{cb}|$ and $m_b$ in the $1S_{\mathrm{EXP}}$ and $\mathrm{kin}_{\mathrm{EXP}}$ schemes with and without including our estimate of the theoretical uncertainties. This plot also shows for comparison $m_b = 4.69 \pm 0.03\,\mathrm{GeV}$ extracted by Hoang [38] from sum rules [39, 40] that fit to the $\bar{B}B$ system near threshold, and the PDG 2004 value [24] $|V_{cb}| = (42.0 \pm 1.1 \pm 1.8) \times 10^{-3}$ from exclusive decays. Hoang's determination of $m_b^{1S}$ is independent of the current determination, and the agreement is remarkable. The PDG 2004 value for $|V_{cb}|$ from exclusive decays is also independent of our determination from inclusive decays.

In summary, we find the following fit results:

$$\begin{aligned} |V_{cb}| &= (41.9 \pm 0.6 \pm 0.1_{\tau_B}) \times 10^{-3}, \\ m_b^{1S} &= (4.68 \pm 0.04)\,\mathrm{GeV}, \end{aligned} \tag{22}$$

from the $1S_{\mathrm{EXP}}$ fit including theory errors, where the first error is the uncertainty from the fit, and the second error (for $|V_{cb}|$) is due to the uncertainty in the average $B$ lifetime. From the $1S_{\mathrm{EXP}}$ fit with no theory errors, and using the PDG method of scaling the uncertainties so that $\chi^2/\nu$ is unity, we obtain

$$\begin{aligned} |V_{cb}| &= (42.1 \pm 0.4 \pm 0.1_{\tau_B}) \times 10^{-3}, \\ m_b^{1S} &= (4.69 \pm 0.03)\,\mathrm{GeV}. \end{aligned} \tag{23}$$

The increase in $|V_{cb}|$ compared to Ref. [1] is largely due to an increase in the experimental values for the semileptonic $B$ decay rate since two years ago.

The $1S_{\mathrm{EXP}}$ fit (including theoretical uncertainties) also gives

$$\frac{\Gamma(B \to X_c \ell \bar{\nu})}{|V_{cb}|^2} = (2.47 \pm 0.03) \times 10^{-11}\ \mathrm{GeV}. \tag{24}$$

| $E_{\rm cut}$ (GeV) | 0.6 | 0.8 | 1.0 | 1.2 | 1.4 |
|---|---|---|---|---|---|
| $\frac{\mathcal{B}(B\to X_c\ell\bar\nu)}{\mathcal{B}(B\to X_c\ell\bar\nu, E_{\rm cut})} = \frac{R_0(0,0)}{R_0(E_{\rm cut},0)}$ | $1.0521 \pm 0.0003$ | $1.1212 \pm 0.0009$ | $1.245 \pm 0.002$ | $1.464 \pm 0.004$ | $1.876 \pm 0.008$ |

TABLE II: The ratio of the semileptonic branching ratio $\mathcal{B}(B \to X_c\ell\bar\nu)$ to the semileptonic branching ratio with a lepton energy cut $E_{\rm cut}$ obtained using the $1S_{\rm EXP}$ scheme, including theoretical uncertainties.

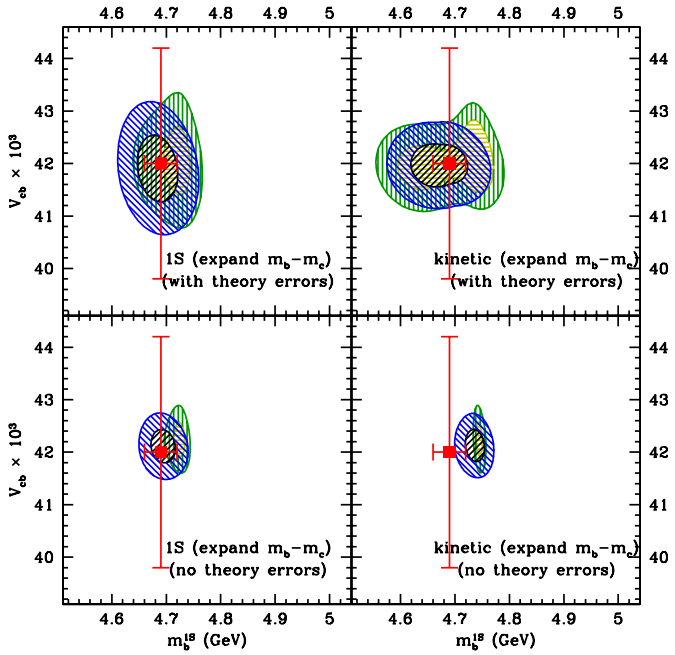


FIG. 6: Fit results for $|V_{cb}|$ and $m_b$ in the $1S_{\rm EXP}$ and $\text{kin}_{\rm EXP}$ schemes. The upper plots include our estimate of the theoretical errors, the lower ones set them to zero. The black and blue regions are the fit results ($\Delta\chi^2 = 1, 4$), and the yellow and green regions are the fit results ($\Delta\chi^2 = 1, 4$) omitting $\chi^2_{\rm param}$ in Eq. (18). We have also shown a red point given by combining Hoang's determination of $m_b^{1S}$ and the PDG 2004 value of $|V_{cb}|$ from exclusive decays.

The ratio of the semileptonic branching ratio with no energy cut to that with an energy cut is given Table II. The semileptonic branching ratio obtained from the fit (including theoretical uncertainties) is $\mathcal{B}(B \to X_c\ell\bar\nu) = 0.106 \pm 0.002$. Note that this number depends on the PDG 2004 value (corrected for $B \to X_u$ contamination, see Eq. (3)) of $0.105 \pm 0.003$, and the BABAR branching ratio measurements with an energy cut, which give a higher value of $0.107 \pm 0.002$, when converted to the branching ratio using Table II.

Another useful quantity is the $C$ parameter, needed for the $B \to X_s\gamma$ rate [41], which is defined to be

$$C = \frac{\Gamma(B \to X_c\ell\bar\nu)}{|V_{cb}|^2} \frac{|V_{ub}|^2}{\Gamma(B \to X_u\ell\bar\nu)} = 0.58 \pm 0.02\,. \quad (25)$$

The value of $C$ depends on the (unknown) matrix element of four-quark operators, which enter the $B \to X_u\ell\bar\nu$ rate at order $1/m_b^3$ (but not $B \to X_c\ell\bar\nu$). These absorb a logarithmic divergence in the $1/m_b^3$ corrections to the $B \to X_c\ell\bar\nu$ rate in the formal limit $m_c \to 0$. (For $m_c \gg \Lambda_{\rm QCD}$, four-quark operators only enter the $B \to X_c\ell\bar\nu$ rate at order $\alpha_s$.) The four-quark operator matrix element gives an uncertainty in addition to that in Eq. (25). An estimate of the four-quark operators' contribution is obtained by replacing the formally divergent term in the $B \to X_u\ell\bar\nu$ rate, $8(\rho_1/m_b^3)\ln(m_u^2/m_b^2)$ [15] by $8(\Lambda_{\rm QCD}/m_b)^3 \sim 0.01$ [42].

The above fits give a robust value for $|V_{cb}|$ and $m_b$. However, we recommend using the error estimate with caution. As we have pointed out, the fit seems to indicate that the unknown higher order corrections are smaller than our theoretical estimate of 0.1%, so that one can use Eq. (23). A theoretical uncertainty less than 0.1% is very small for a hadronic quantity at the relatively low scale of around 5 GeV. It is interesting that the current fit shows that the theoretical uncertainties in inclusive $B$ decay shape variables are so small. If this is confirmed by further comparisons between theory and experiment, the uncertainty in $V_{cb}$ can be reduced still further.

### Acknowledgments

We thank our friends at BABAR, CLEO and DELPHI for numerous discussions. We would also like to thank M. Misiak for pointing out the importance of computing $C$. This work was supported in part by the US Department of Energy under Contract DE-FG03-92ER40701 (CWB), DE-AC03-76SF00098 and by a DOE Outstanding Junior Investigator award (ZL), DE-FG03-97ER40546 (AVM), and by the Natural Sciences and Engineering Research Council of Canada (ML and MT).

[1] C. W. Bauer, Z. Ligeti, M. Luke and A. V. Manohar, Phys. Rev. D **67**, 054012 (2003).

[2] M. Battaglia *et al.*, Phys. Lett. B **556** (2003) 41.

[3] A. H. Mahmood *et al.* [CLEO Collaboration], Phys. Rev.

D **67**, 072001 (2003).
[4] B. Aubert *et al.* [BABAR Collaboration], hep-ex/0404017.
[5] J. Chay, H. Georgi and B. Grinstein, Phys. Lett. B247 (1990) 399; M.A. Shifman and M.B. Voloshin, Sov. J. Nucl. Phys. 41 (1985) 120; I.I. Bigi, N.G. Uraltsev and A.I. Vainshtein, Phys. Lett. B293 (1992) 430 [E. B297 (1992) 477]; I.I. Bigi, M.A. Shifman, N.G. Uraltsev and A.I. Vainshtein, Phys. Rev. Lett. 71 (1993) 496; A.V. Manohar and M.B. Wise, Phys. Rev. D49 (1994) 1310.
[6] A.V. Manohar and M.B. Wise, Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol. **10** (2000) 1.
[7] N. Isgur, Phys. Lett. B448 (1999) 111; Phys. Rev. D60 (1999) 074030.
[8] M. Trott, hep-ph/0402120.
[9] P. Gambino and N. Uraltsev, Eur. Phys. J. C **34**, 181 (2004).
[10] A. Hoang, Z. Ligeti and A. Manohar, Phys. Rev. Lett. 82 (1999) 277; Phys. Rev. D59 (1999) 074017.
[11] A.H. Hoang and T. Teubner, Phys. Rev. D60 (1999) 114027.
[12] M. Beneke, Phys. Lett. B434 (1998) 115.
[13] M.B. Voloshin, Phys. Rev. D51 (1995) 4934.
[14] M. Gremm, A. Kapustin, Z. Ligeti and M.B. Wise, Phys. Rev. Lett. 77 (1996) 20.
[15] M. Gremm and A. Kapustin, Phys. Rev. D55 (1997) 6924.
[16] M. Gremm and I. Stewart, Phys. Rev. D55 (1997) 1226.
[17] A.F. Falk, M. Luke, and M.J. Savage, Phys. Rev. D53 (1996) 2491; D53 (1996) 6316;
[18] A.F. Falk and M. Luke, Phys. Rev. D57 (1998) 424.
[19] A.F. Falk, M. Luke, and M.J. Savage, Phys. Rev. D49 (1994) 3367.
[20] A. Kapustin and Z. Ligeti, Phys. Lett. B355 (1995) 318.
[21] Z. Ligeti, M. Luke, A.V. Manohar and M.B. Wise, Phys. Rev. D60 (1999) 034019.
[22] C. Bauer, Phys. Rev. D57 (1998) 5611 [Erratum-ibid. D60 (1999) 099907].
[23] M. Luke, M.J. Savage, and M.B. Wise, Phys. Lett. B345 (1995) 301.
[24] S. Eidelman *et al.*, Particle Data Group, Phys. Lett. B **592**, 1 (2004).
[25] B. Aubert *et al.* [BABAR Collaboration], hep-ex/0403030.
[26] R. Kaiser, A. V. Manohar and T. Mehen, Phys. Rev. Lett. **90**, 142001 (2003).
[27] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. D **69**, 071101 (2004).
[28] A. H. Mahmood *et al.* [CLEO Collaboration], hep-ex/0403053.
[29] DELPHI Collaboration, DELPHI note: 2003-028-CONF-648, `http://delphiwww.cern.ch/pubxx/conferences/summer03/PapNo046.html`.
[30] B. Aubert *et al.* [BABAR Collaboration], hep-ex/0403031.
[31] CDF Collaboration, CDF note 6973, available at: `http://www-cdf.fnal.gov/physics/new/bottom/040428.blessed-bhadr-moments/`
[32] S. E. Csorna *et al.* [CLEO Collaboration], hep-ex/0403052.
[33] C.W. Bauer and M. Trott, Phys. Rev. D **67**, 014021 (2003).
[34] P. Koppenburg *et al.* [BELLE Collaboration], hep-ex/0403004.
[35] S. Chen *et al.* (CLEO Collaboration), Phys. Rev. Lett. 87 (2001) 251807.
[36] B. Aubert *et al.* [BABAR Collaboration], hep-ex/0207074.
[37] P. J. Mohr and B. N. Taylor, Rev. Mod. Phys. **72**, 351 (2000).
[38] A.H. Hoang, hep-ph/0008102.
[39] M. Beneke and A. Signer, Phys. Lett. B **471**, 233 (1999).
[40] M. B. Voloshin, Int. J. Mod. Phys. A **10**, 2865 (1995).
[41] P. Gambino and M. Misiak, Nucl. Phys. B **611**, 338 (2001).
[42] C. W. Bauer and C. N. Burrell, Phys. Lett. B **469**, 248 (1999).