Three loop estimate of the inclusive semileptonic $b \to c$ decay rate

M. R. Ahmady,* F. A. Chishtie,† V. Elias,† A. H. Fariborz,‡ D. G. C. McKeon,† T. N. Sherry,§ and T. G. Steele‖
Theory Group, KEK, Tsukuba, Ibaraki 305-0801, Japan
(Received 17 September 2001; published 11 February 2002)

The renormalization-scale (μ) dependence of the two-loop inclusive semileptonic $b \to c l^- \overline{\nu_l}$ decay rate is shown to be significant in the pole mass scheme, and the decay rate is shown to be poorly convergent in the $\overline{\mathrm{MS}}$ scheme. Three-loop contributions to the decay rate are estimated by developing Padé approximant techniques particularly suited to perturbative calculations in the pole mass scheme. An optimized Padé estimate of the three-loop contributions is obtained by comparison of the Padé estimates with the three-loop terms determined by renormalization-group invariance. The resulting three-loop estimate in the pole-mass scheme exhibits minimal sensitivity to the renormalization scale near $\mu = 1.0$ GeV, leading to an estimated decay rate of $192\pi^3 \Gamma(b \to c l^- \overline{\nu_l})/(G_F^2 |V_{cb}|^2) = 992 \pm 217\ \mathrm{GeV}^5$ inclusive of theoretical uncertainties and nonperturbative effects.

DOI: 10.1103/PhysRevD.65.054021 PACS number(s): 13.20.He

I. INTRODUCTION

The Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{cb}|$, which parametrizes one of the sides of the unitarity triangle, can be extracted from the inclusive semileptonic decay rate $\Gamma(B \to X_c l^- \overline{\nu_l})$. From a theoretical perspective, the inclusive process has the advantage that nonperturbative contributions are controllable; hence, an accurate perturbative determination of the $b \to c l^- \overline{\nu_l}$ decay rate is of value in obtaining $|V_{cb}|$ from data.

Complete two-loop calculations of semileptonic $b \to c$ decays exist at the end points of the lepton invariant mass spectrum (maximal and zero recoil) [1] and at an intermediate kinematic value [2]. From these explicit calculations, the total semileptonic decay rate at two-loop order has been estimated [2]. In this present work, we extend these results to generate an estimate of the three-loop contributions to the $b \to c l^- \overline{\nu_l}$ decay rate via renormalization-group and Padé-approximant methods.

In Sec. II we demonstrate that the pole mass scheme used in [2] has better perturbative behavior than the modified minimal subtraction ($\overline{\mathrm{MS}}$) scheme for the $b \to c$ semileptonic decay rate, a result which is somewhat surprising since the $\overline{\mathrm{MS}}$ scheme is better behaved within calculations of the $b \to u$ semileptonic decay rate [3]. The renormalization scale dependence of the two-loop $b \to c$ rate in the pole mass scheme is extracted using renormalization-group (RG) invariance, and the strong scale dependence that is found serves to motivate an estimate of next-order (three-loop) effects.

The scale dependence of the decay rate is an important component of the procedure for estimating three-loop corrections to the $b \to c$ semileptonic rate. Using information obtained from RG invariance, Padé approximation methods are developed in Sec. III that are appropriate for estimating next-order terms within pole-mass perturbative calculations. An optimized Padé estimate of the three-loop constant coefficient (i.e. the nonlogarithmic term) is obtained by finding the best agreement between Padé estimates and true values of the RG-accessible three-loop coefficients of logarithms. The RG determinations of these latter coefficients in conjunction with the optimized Padé estimate of the constant coefficient together constitute a scale-sensitive estimate of the full three-loop contribution to the perturbative rate.

This Padé estimate of aggregate three-loop effects allows the renormalization-scale dependence of the $b \to c l^- \overline{\nu_l}$ decay rate to be studied. In Sec. IV, a region of minimal scale sensitivity [4] is found in the resulting decay rate. This minimal-sensitivity scale is found to be close to the fastest-apparent-convergence renormalization scale at which the three-loop contributions vanish entirely [5]. The proximity of these two scales supports the validity of these scales for obtaining estimates of the three-loop perturbative $b \to c l^- \overline{\nu_l}$ decay rate. Theoretical uncertainties in this estimate are considered in Sec. V.

II. SCALE DEPENDENCE OF THE TWO-LOOP RATE

In Ref. [2], the inclusive semileptonic $b \to c l^- \overline{\nu_l}$ rate is estimated to two-loop order in terms of renormalization-group- (RG-) invariant pole masses m_b and m_c. The rate Γ can be expressed in the form

$$\Gamma = \frac{G_F^2 m_b^5 |V_{cb}|^2}{192\pi^3} F\left(\frac{m_c^2}{m_b^2}\right) S\left[\frac{\alpha_s(\mu)}{\pi}, \log\left(\frac{\mu^2}{m_b m_c}\right)\right] \qquad (1)$$

with the RG-invariant form-factor

*Permanent address: Department of Physics, Mount Allison University, Sackville, NB, Canada E4L 1E6.
†Permanent address: Department of Applied Mathematics, The University of Western Ontario, London, ON, Canada N6A 5B7.
‡Permanent address: Dept. of Mathematics/Science, State Univ. of New York Institute of Technology, Utica, NY 13504-3050.
§Permanent address: Department of Mathematical Physics, National University of Ireland, Galway, Ireland.
‖Permanent address: Department of Physics & Engineering Physics, University of Saskatchewan, Saskatoon, SK, Canada S7N 5E2.

$$F(r)=1-8r-12r^2\log(r)+8r^3-r^4 \tag{2}$$

preceding the perturbative series S whose two-loop contribution at $\mu^2=m_bm_c$ has been estimated in [2] by combining explicit results at the end and intermediate points [1,2] of the decay spectrum:[1]

$$S\left[\frac{\alpha_s(\sqrt{m_bm_c})}{\pi},0\right]=1-1.67\frac{\alpha_s(\sqrt{m_bm_c})}{\pi}-(8.9\pm0.3)\times\left[\frac{\alpha_s(\sqrt{m_bm_c})}{\pi}\right]^2. \tag{3}$$

If we assume central values $\alpha_s(m_\tau)=0.33$ [6,7] and $m_b=4.9$ GeV [8], and if we follow Ref. [2] in assuming that $m_c=0.3m_b$, we find that α_s evolves from $\mu=m_\tau$ through the four-loop, four-flavor[2] $(n_f=4)$ β-function [9] to $\alpha_s(\sqrt{m_bm_c})/\pi=0.087$, in which case $S[0.087,0]=1-0.145-(0.067\pm0.002)$. It is evident from Eq. (3) that truncation after two-loop order introduces theoretical uncertainty of order $8.5\%(=0.067/0.79)$.

There are, of course, other sources of theoretical uncertainty. Ambiguities concerning the definition of a pole mass in the presence of nonperturbative (confinement) effects have led the authors of Ref. [2] to reparametrize the rate (1) in terms of "low-scale" masses obtained through additional phenomenology—we will ultimately compare our results to the reparametrized rate in Sec. V. In light of more recent work [3,8,10,11] specifying accurate pole mass values by relating pole masses to $\overline{\text{MS}}$ and Υ-scheme masses with three-loop precision, we take a more empirical approach to the utilization of pole masses within inclusive semileptonic rates.

Renormalization-scale dependence provides an additional source of theoretical uncertainty to any rate calculated via the series (3). If S varies for different choices of μ, a value for Γ extracted from any particular choice of μ (e.g., $\mu^2=m_bm_c$) is compromised. The optimal value of μ has been argued to be the choice for which S has minimal sensitivity to the renormalization scale [4] i.e., the point at which $dS/d\mu=0$. For Eq. (3) to lead to a reliable estimate of the true semileptonic $b\to c$ rate, one would not necessarily need to establish that $\mu^2=m_bm_c$ is such a "principle of minimal sensitivity" (PMS) point, but rather that the rate calculated at this value of μ differs only inconsequentially from the rate calculated at the true PMS value of μ. Any discrepancy between rates calculated at $\mu^2=m_bm_c$ and at μ_{PMS} is a direct measure of the former rate's theoretical uncertainty arising from renormalization-scale ambiguities.

The two-loop order renormalization-scale dependence implicit in the perturbative series within Eq. (1) may be parametrized as follows:

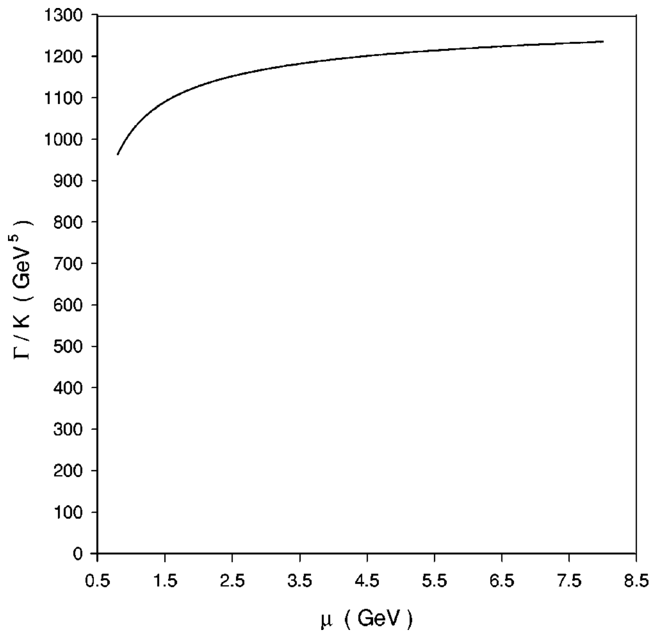

FIG. 1. Renormalization scale (μ) dependence of the two-loop reduced rate Γ/K in the pole mass scheme.

$$S[x(\mu),L(\mu)]=1+(a_0+a_1L)x+(b_0+b_1L+b_2L^2)x^2, \tag{4}$$

$$a_0=-1.67,\quad b_0=-8.9\pm0.3,$$

$$x(\mu)\equiv\frac{\alpha_s(\mu)}{\pi},\quad L(\mu)\equiv\log\left(\frac{\mu^2}{m_bm_c}\right). \tag{5}$$

The renormalization-scale invariance of the all orders rate implies that

$$\mu^2\frac{\mathrm{d}\Gamma}{\mathrm{d}\mu^2}=0 \tag{6}$$

hence that

$$0=\frac{\partial S[x,L]}{\partial L}-(\beta_0x^2+\beta_1x^3+\beta_2x^4+\cdots)\frac{\partial S[x,L]}{\partial x} \tag{7}$$

where the normalization of QCD β-function coefficients is explicitly defined by

$$\mu^2\frac{\mathrm{d}x}{\mathrm{d}\mu^2}=-(\beta_0x^2+\beta_1x^3+\beta_2x^4+\cdots) \tag{8}$$

with $n_f=4$ values $\beta_0=25/12$, $\beta_1=77/24$, and $\beta_2=21943/3456$ (in the $\overline{\text{MS}}$ scheme). It is easily seen from Eq. (7) that the logarithmic coefficients within Eq. (4) are

$$a_1=b_2=0,\quad b_1=a_0\beta_0=-3.479. \tag{9}$$

In Fig. 1, we have plotted the mass- and scale-dependent portion of the rate (1):

[1]The $\mathcal{O}(\alpha_s^2)$ coefficient quoted in Eq. (3) [A. Czarnecki (personal communication)] has been slightly corrected from the value -8.4 ± 0.4 appearing in Ref. [2].

[2]The use of four contributing flavors is necessary since α_s is referenced to four flavors in Eq. (3) [2].

$$\frac{\Gamma}{K}=m_b^5 F\left(\frac{m_c^2}{m_b^2}\right)S[x(\mu),L(\mu)], \quad \left(K\equiv\frac{192\pi^3}{G_F^2|V_{cb}|^2}\right), \tag{10}$$

as a function of μ. The Fig. 1 curve is obtained by assuming $m_b=4.9$ GeV [8], $m_c=0.3m_b$ [2], and the central value $b_0=-8.9$ from the estimate [2] given in Eq. (5). Substantial scale dependence is evident from the figure: the rate increases monotonically with μ, flattening out somewhat for larger values. Moreover the curve exhibits no PMS point (i.e., extremum), and at $\mu=\sqrt{m_c m_b}=2.68$ GeV yields a rate 10% smaller than the (still increasing) rate at $\mu=8$ GeV, the flattest portion of the curve shown in the figure. Thus, the two-loop pole-mass calculation of the $b\rightarrow c l^-\overline{\nu_l}$ rate exhibits at best only poorly-controllable dependence on the choice of renormalization scale. Such scale dependence may compromise any subsequent "low-scale" mass expression devolving from the two-loop $\mu=\sqrt{m_b m_c}$ pole-mass rate.

Similar scale dependence and even worse apparent nonconvergence characterize the two-loop order pole-mass calculation of the $b\rightarrow u l^-\overline{\nu_l}$ rate, thereby motivating a recasting of the calculation in terms of the $\overline{\rm MS}$ running b-quark mass [3]. For four contributing flavors, the two-loop relationship between pole and $\overline{\rm MS}$ running quark masses is [10]

$$\begin{aligned} m^{pole}=m(\mu)\Bigg[1&+\left(\frac{4}{3}+\log\left(\frac{\mu^2}{m^2(\mu)}\right)\right)x(\mu)\\ &+\left(10.3919+\frac{415}{72}\log\left(\frac{\mu^2}{m^2(\mu)}\right)\right.\\ &\left.+\frac{37}{24}\log^2\left(\frac{\mu^2}{m^2(\mu)}\right)\right)x^2(\mu)\Bigg]. \end{aligned} \tag{11}$$

By substituting this relation into Eqs. (1), (4), and (5), we find that the fully $\overline{\rm MS}$ version of the mass- and scale-dependent portion (10) of the $b\rightarrow c l^-\overline{\nu_l}$ decay rate (using the central value $b_0=-8.9$) is

$$\begin{aligned} \frac{\Gamma^{\overline{MS}}}{K}=m_b^5(\mu)F\left(\frac{m_c^2}{m_b^2}\right)\Bigg\{1&+\left(5.00+5\log\left(\frac{\mu^2}{m_b^2(\mu)}\right)\right)x(\mu)\\ &+\left[49.3-1.74\log\left(\frac{\mu^2}{m_c^2(\mu)}\right)+45.4\log\left(\frac{\mu^2}{m_b^2(\mu)}\right)\right.\\ &\left.+\frac{425}{24}\log^2\left(\frac{\mu^2}{m_b^2(\mu)}\right)\right]x^2(\mu)\Bigg\}. \end{aligned} \tag{12}$$

Note that $F(m_c^2/m_b^2)$ remains RG invariant if $m_c\rightarrow m_c(\mu)$ and $m_b\rightarrow m_b(\mu)$. If $\mu^2=m_b(\mu)m_c(\mu)$ and if we continue to assume that $m_c(\mu)=0.3m_b(\mu)$, then

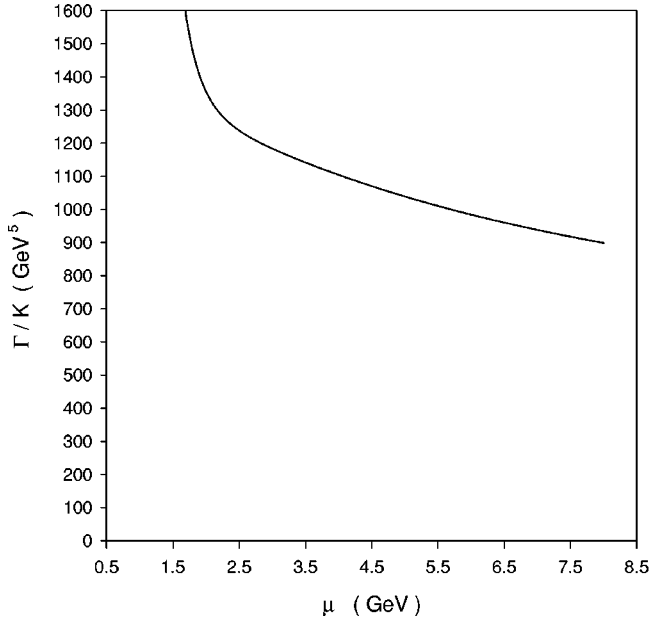

FIG. 2. Renormalization scale (μ) dependence of the two-loop reduced rate Γ/K in the $\overline{\rm MS}$ scheme. In the $\overline{\rm MS}$ scheme, $m_b(\mu)$ is obtained from the four-loop $n_f=4$ anomalous mass dimension [12] using $m_b(m_b)=4.2$ GeV [8] as a reference value.

$$\frac{\Gamma^{\overline{MS}}}{K}=m_b^5(\mu)F(0.09)[1-1.02x(\mu)+18.2x^2(\mu)],$$

$$\mu=\sqrt{0.3}m_b(\mu). \tag{13}$$

The convergence of this $\overline{\rm MS}$ perturbative series is even more ill-behaved than its pole-mass version (3). Moreover, the scale dependence of Eq. (12) is shown in Fig. 2 to be even more pronounced than that of the same rate in the pole-mass scheme (Fig. 1)—Fig. 2 displays a rate which decreases with μ with no apparent PMS point.

Thus the $\overline{\rm MS}$ approach, which substantially improves the perturbative series within the semileptonic $b\rightarrow u$ rate [3], fails to improve the pole-mass expressions (1)–(3) for the semileptonic $b\rightarrow c$ rate. If this latter rate is to be utilized to extract an estimate of $|V_{cb}|$ from the inclusive $B\rightarrow X_c l^-\overline{\nu_l}$ branching ratio, there is evident value in having an estimate of next-order corrections in order to obtain some control over renormalization-scale dependence. This three-loop order contribution to $S[x(\mu),L(\mu)]$ is necessarily of the form

$$\begin{aligned} \Delta S^{3L}[x(\mu),L(\mu)]=[c_0+c_1L(\mu)+c_2L^2(\mu)\\ +c_3L^3(\mu)]x^3(\mu). \end{aligned} \tag{14}$$

The RG equation (7) implies that

$$\begin{aligned} 0=a_1x&+(b_1-a_0\beta_0)x^2+(2b_2-\beta_0a_1)x^2L\\ &+[c_1-2\beta_0b_0-a_0\beta_1]x^3+[2c_2-2\beta_0b_1-\beta_1a_1]x^3L\\ &+[3c_3-2\beta_0b_2]x^3L^2+\mathcal{O}(x^4). \end{aligned} \tag{15}$$

The set of results (9) are evident from the requirement that $\mathcal{O}(x)$, $\mathcal{O}(x^2)$, and $\mathcal{O}(x^2L)$ terms in Eq. (15) separately vanish. The vanishing of subsequent terms in Eq. (15) implies that

$$c_1 = 2b_0\beta_0 + a_0\beta_1, \quad c_2 = a_0\beta_0^2, \quad c_3 = 0. \tag{16}$$

For a_0 and b_0 as given in Eq. (5) and three massless flavors (as appropriate to phenomenology devolving from $\alpha_s(m_\tau)$ [6]),

$$c_1 = -42.4 \pm 1.3, \quad c_2 = -7.25, \quad c_3 = 0. \tag{17}$$

The coefficient c_0, however, is RG-inaccessible to these orders of perturbation theory, and requires a direct three-loop calculation. In the absence of such a calculation, we estimate c_0 in the section which follows via asymptotic Padé approximant methodology, in much the same way as in a prior estimate [13] of the three-loop contribution to the $b \to u l^- \overline{\nu_l}$ decay rate.

III. RG-PADÉ ESTIMATE OF c_0

Consider a perturbative field-theoretical series with $N+M$ known terms:

$$S(x) = 1 + R_1 x + R_2 x^2 + \cdots + R_{N+M} x^{N+M} + \cdots. \tag{18}$$

The set of known coefficients $\{R_1, R_2, \ldots, R_{N+M}\}$ is sufficient to determine in full the $N+M$ coefficients characterizing an $[N|M]$ Padé approximant to the series S:

$$S_{[N|M]}(x) \equiv \frac{1 + a_1 x + a_2 x^2 + \cdots + a_N x^N}{1 + b_1 x + b_2 x^2 + \cdots + b_M x^M}. \tag{19}$$

The coefficients $\{a_1, \ldots, a_N, b_1, \ldots, b_M\}$ are obtained by the requirement that the power-series expansion of $S_{[N|M]}(x)$ recovers the $N+M$ known coefficients within Eq. (18), the series $S(x)$. The next $\mathcal{O}(x^{N+M+1})$ term in this power series is a Padé approximant prediction for the first unknown coefficient R_{N+M+1}. For example, if only the next-to-leading order coefficient R_1 is known, one can use this coefficient to construct a $[0|1]$ approximant to the series (18),

$$S_{[0|1]}(x) = \frac{1}{1 - R_1 x} = 1 + R_1 x + R_1^2 x^2 + \cdots, \tag{20}$$

for which R_1^2 is the *predicted* value of R_2:

$$R_2^{[0|1]} = R_1^2. \tag{21}$$

Similarly, if R_1 and R_2 are known (corresponding to two subleading orders of perturbation theory), one has enough information to construct a $[1|1]$ approximant to the series (18):

$$S_{[1|1]}(x) = \frac{1 + \left(\dfrac{R_1^2 - R_2}{R_1}\right) x}{1 - \dfrac{R_2}{R_1} x} = 1 + R_1 x + R_2 x^2 + \frac{R_2^2}{R_1} x^3 + \cdots. \tag{22}$$

The predicted value for the unknown coefficient R_3 is

$$R_3^{[1|1]} = R_2^2 / R_1. \tag{23}$$

In general, one can always use an $[N|M]$ approximant (19) to predict the first unknown series coefficient R_{N+M+1}. Such predictions have accuracy which increases as N and M increase. For perturbative field-theoretical series (characterized by asymptotic $R_N \sim N!$ behavior) the accuracy of such predictions has been argued by Ellis, Karliner and Samuel to satisfy the relative error formula [14]

$$\frac{\Delta R_{N+M+1}^{[N|M]}}{R_{N+M+1}^{true}} = \frac{R_{N+M+1}^{[N|M]} - R_{N+M+1}^{true}}{R_{N+M+1}^{true}} = -\frac{M! A^M}{(N + M + aM + b)^M} \tag{24}$$

where A, a, and b are constants to be determined. Of particular interest are the relative errors obtained from Eq. (24) for the predictions (21) and (23)

$$\frac{R_1^2 - R_2}{R_2} = -\frac{A}{1 + (a+b)}, \tag{25}$$

$$\frac{\dfrac{R_2^2}{R_1} - R_3}{R_3} = -\frac{A}{2 + (a+b)}. \tag{26}$$

Denoting the constant $a+b=k$, we can eliminate the other constant A within Eqs. (25),(26) and solve for R_3 algebraically to obtain the improved estimate

$$R_3 = \frac{(2+k) R_2^3}{(1+k) R_1^3 + R_1 R_2}. \tag{27}$$

The assumption $k=0$ has been utilized in prior applications to predict successfully the third subleading contribution to the QCD $\overline{\text{MS}}$ β function [14], the dimensional reduction (DRED) SQCD β function [15], the $\overline{\text{MS}}$ β function for massive scalar field theory [16], as well as to obtain estimates of such contributions for a number of processes calculated in the $\overline{\text{MS}}$ scheme: SM [17] and minimal supersymmetric standard model (MSSM) [18] Higgs boson $\to 2$ gluon decay rates, $b \to u l^- \overline{\nu_l}$ [13], $W^+ W^- \to ZZ$ [16], and the QCD static potential function [19].

The choice $k=0$, however, is ill suited to pole-mass calculations. To see this, consider the series (4) augmented by (14) within Eq. (1) characterizing the pole-mass expression for the $b \to c l^- \overline{\nu_l}$ decay rate

$$S[x,L]=1+(a_0+a_1L)x+(b_0+b_1L+b_2L^2)x^2 +(c_0+c_1L+c_2L^2+c_3L^3)x^3 \tag{28}$$

$$x=\frac{\alpha_s(\mu)}{\pi}, \quad L=\log\left(\frac{\mu^2}{m_b m_c}\right).$$

We have already seen that $a_0=-1.67, a_1=0, b_0\cong-8.9\pm0.3, b_1=a_0\beta_0=-3.48, b_2=0$. The above series is in the form of Eq. (18) with R_1 and R_2 respectively identified with a_0+a_1L and $b_0+b_1L+b_2L^2$. If $k=0$, we see from Eq. (27) that

$$R_3=\frac{2(b_0+b_1L)^3}{a_0(a_0^2+b_0+b_1L)} =\frac{2b_1^2}{a_0}L^2+\frac{2b_1}{a_0}(2b_0-a_0^2)L+\mathcal{O}(L^0). \tag{29}$$

Comparing this expression to the form (14) anticipated for the third subleading order of $S[x,L]$, we necessarily obtain the following predictions for the RG accessible coefficients c_2, c_1:

$$c_2=2\frac{b_1^2}{a_0}=2a_0\beta_0^2, \tag{30}$$

$$c_1=2(2b_0-a_0^2)\frac{b_1}{a_0}=4\beta_0 b_0-2\beta_0 a_0^2. \tag{31}$$

It is evident from Eqs. (16),(17) that these predictions are quite poor; Eq. (30) is double the true value for c_2, as obtained in Eq. (17), and c_1 is also badly overestimated [for values (9) and the central value $b_0=-8.9$, Eq. (31) implies that $c_1=-85.8$, in contrast to the correct (RG) value $c_1=-42.4$].

For a given choice of k, estimates of the coefficients c_i characterizing the third subleading order have been obtained for correlation functions [20] and the $b\rightarrow ul^-\overline{\nu_l}$ rate [13] by moments of R_3, as estimated in Eq. (27), over the entire ultraviolet region [e.g., $\mu^2/(m_b m_c)\geqslant 1$ for the case at hand]. These moments are then equated to corresponding moments of $R_3=c_0+c_1L+c_2L^2+c_3L^3$ in order to obtain values for $\{c_0,c_1,c_2,c_3\}$. If we define $w=m_b m_c/\mu^2$ ($L=-\log w$), the moments

$$N_j=(j+2)\int_0^1 w^{j+1}R_3(w)\,\mathrm{d}w \tag{32}$$

can be obtained using Eq. (27) for the integrand R_3, which becomes a function of w for our pole mass case by virtue of the w-dependence of R_2: $R_1=a_0, R_2=b_0-b_1\log w$. After (numerical) computation of the values of N_j, the coefficients c_i are obtained by equating such values to the corresponding integrals

$$N_j=(j+2)\int_0^1 w^{j+1}(c_0-c_1\log w+c_2\log^2 w-c_3\log^3 w)\,\mathrm{d}w. \tag{33}$$

In particular, we see that

$$N_{-1}=c_0+c_1+2c_2+6c_3 \tag{34}$$

$$N_0=c_0+\frac{1}{2}c_1+\frac{1}{2}c_2+\frac{3}{4}c_3 \tag{35}$$

$$N_1=c_0+\frac{1}{3}c_1+\frac{2}{9}c_2+\frac{2}{9}c_3 \tag{36}$$

$$N_2=c_0+\frac{1}{4}c_1+\frac{1}{8}c_2+\frac{3}{32}c_3. \tag{37}$$

If $k\neq-1$, such a procedure is seen to lead to a non-zero value of c_3, in contradiction to the result $c_3=0$ (16) necessarily following from application of the RG equation (7) within the pole mass scheme. For the case $k=-1$, however, the result (27) collapses to the naive estimate (23), which in the pole mass scheme is necessarily a degree-2 polynomial in L:

$$R_3^{(k=-1)}=\frac{R_2^2}{R_1}=\frac{1}{a_0}(b_0+b_1L)^2=\frac{b_1^2}{a_0}L^2+2\frac{b_1b_0}{a_0}L+\frac{b_0^2}{a_0}. \tag{38}$$

The moment procedure described above then reduces to fitting the degree-3 polynomial $R_3=c_3L^3+c_2L^2+c_1L+c_0$ to the degree-2 polynomial in Eq. (38). Such a fit necessarily reduces to equating the powers of L in these two expressions. We thus find that c_3 must be zero, and that

$$c_2=\frac{b_1^2}{a_0}=a_0\beta_0^2 \tag{39}$$

$$c_1=2\frac{b_1b_0}{a_0}=2\beta_0 b_0 \tag{40}$$

$$c_0=\frac{b_0^2}{a_0}. \tag{41}$$

Equation (39) is in exact agreement with the result (16) obtained via RG-invariance from Eq. (7). Equation (40) represents the first (and dominant) contribution to the RG determination of c_1 in Eq. (16). For the central value $b_0=-8.9$ (and $n_f=4$) the (40) prediction $c_1=-37.1$ is not far from the true value $c_1=-42.4$. The accuracy of these results provide some support to the naive estimate $c_0=-47.4$ obtained via Eq. (41) for the RG-inaccessible coefficient c_0.

We can further improve our estimate for c_0 by finding the value of k within Eq. (27) which, when used within the integrand of Eq. (32) to match the moments N_j to Eqs. (34)–(37), most closely reproduces the true values of c_2 and c_1, as determined by RG methods in Eqs. (16) and (17). In such a procedure we obtain estimated values for c_3, c_2, c_1, and c_0 that depend explicitly on k:

$$c_0(k) = -\frac{1}{6}N_{-1}(k) + 4N_0(k) - \frac{27}{2}N_1(k) + \frac{32}{3}N_2(k) \tag{42}$$

$$c_1(k) = \frac{3}{2}N_{-1}(k) - 32N_0(k) + \frac{189}{2}N_1(k) - 64N_2(k) \tag{43}$$

$$c_2(k) = -\frac{13}{6}N_{-1}(k) + 38N_0(k) - \frac{189}{2}N_1(k) + \frac{176}{3}N_2(k) \tag{44}$$

$$c_3(k) = \frac{2}{3}N_{-1}(k) - 8N_0(k) + 18N_1(k) - \frac{32}{3}N_2(k) \tag{45}$$

where

$$N_j(k) = (j+2)\int_0^1 w^{j+1}\left[\frac{(2+k)(b_0 - b_1\log w)^3}{(1+k)a_0^3 + a_0(b_0 - b_1\log w)}\right] dw. \tag{46}$$

We then use the explicit values for c_1 and c_2 obtained from Eq. (17) by RG methods to optimize the sum of the squares of the relative error of estimated values of c_1 and c_2,

$$\Delta(k) = \left(\frac{c_2(k) - 2a_0\beta_0^2}{2a_0\beta_0^2}\right)^2 + \left(\frac{c_1(k) - (2\beta_0 b_0 - a_0\beta_1)}{2\beta_0 b_0 - a_0\beta_1}\right)^2 \tag{47}$$

with respect to k. For the set of values $a_0 = -1.67, b_0 = -8.9, \beta_0 = 25/12, \beta_1 = 77/24$, we find a clear minimum of $\Delta(k)$ at $k = -0.94$, as evident from Fig. 3, consistent with the case made in the preceding paragraph for an optimal k value close to $k = -1$. Corresponding $k = -0.94$ values for the moments N_j are obtained numerically via Eq. (46)

$$N_{-1} = -106.3, \quad N_0 = -74.92, \quad N_1 = -66.17, \quad N_2 = -62.12 \tag{48}$$

and these values lead via Eqs. (42)–(45) to the following estimates for the third subleading order coefficients:

$$c_3 = 2.0\times 10^{-4}, \quad c_2 = -7.68, \quad c_1 = -39.7, \quad c_0 = -51.2. \tag{49}$$

These values reflect excellent agreement with the RG values (17). The above estimate for the RG-inaccessible coefficient c_0 is only 20% larger in magnitude than the naive prediction (41), indicative of the internal consistency of the methodology.

We conclude by noting that four separate estimates of c_0 can be obtained from Eqs. (34)–(37) by substituting into these equations the optimal N_j values (48) as well as the true values $c_3 = 0$, $c_1 = -42.4$, and $c_2 = -7.25$ as determined by RG invariance (17) in the previous section. Solving each equation separately for c_0, we find that

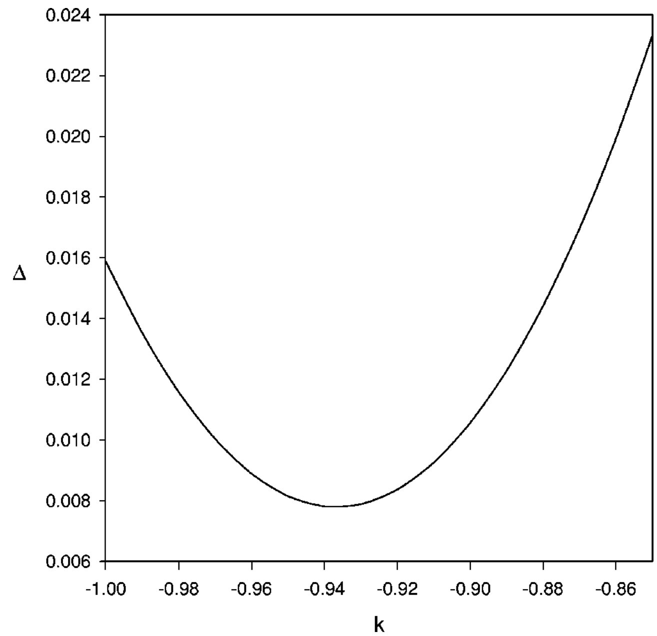

FIG. 3. The quantity Δ measuring the sum of the squares of the relative errors in the Padé estimate (43),(44) of the RG-accessible three-loop coefficients $\{c_1, c_2\}$ plotted as a function of the error-formula parameter k (27). The location of the curve's minimum represents the value of k leading to an optimized Padé estimate.

$$c_0 = N_{-1} - c_1 - 2c_2 - 6c_3 = -49.4, \tag{50}$$

$$c_0 = N_0 - \frac{1}{2}c_1 - \frac{1}{2}c_2 - \frac{3}{4}c_3 = -50.1, \tag{51}$$

$$c_0 = N_1 - \frac{1}{3}c_1 - \frac{2}{9}c_2 - \frac{2}{9}c_3 = -50.4, \tag{52}$$

$$c_0 = N_2 - \frac{1}{4}c_1 - \frac{1}{8}c_2 - \frac{3}{32}c_3 = -50.6. \tag{53}$$

The estimates (50)–(53) are remarkably consistent with each other and only slightly smaller in magnitude than the estimate in Eq. (49) obtained without RG inputs for $\{c_1, c_2, c_3\}$.

In estimating the third subleading order of Eq. (14), the series within the rate (1), we choose Eq. (51) as the most central of our c_0 estimates, in conjunction with the explicit RG determinations of c_1 and c_2 $(c_3 = 0)$. This set of values is obtained, as noted earlier, for the central value estimate [2] of b_0:

$$b_0 = -8.9: \ \{c_0, c_1, c_2\} \cong \{-50.1, -42.4, -7.25\}. \tag{54}$$

Precisely the same procedure can be used to obtain corresponding estimates of $\{c_0, c_1, c_2\}$ for the extremes of the $b_0 = -8.9 \pm 0.3$ range obtained in [2] (see footnote 1):

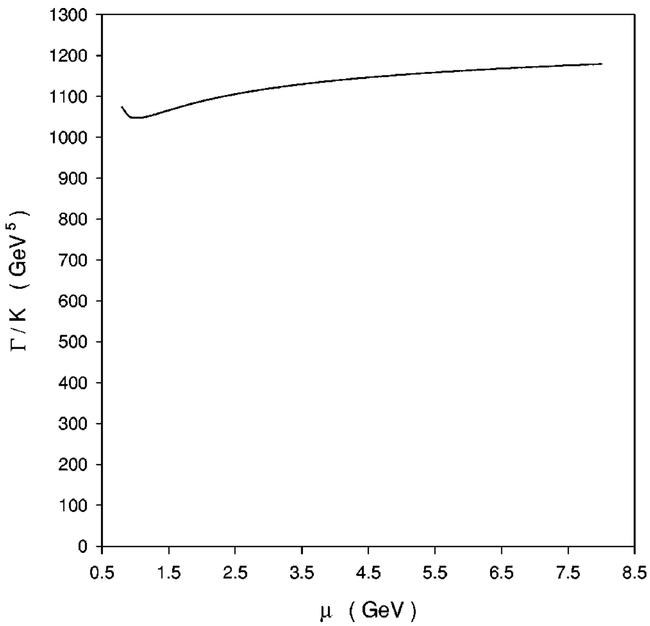

FIG. 4. Renormalization scale (μ) dependence of the three-loop reduced rate Γ/K in the pole mass scheme. The PMS point is represented by the local minimum of the curve.

$$b_0=-9.2: \ \{c_0,c_1,c_2\}\cong\{-53.6,-43.7,-7.25\} \tag{55}$$

$$b_0=-8.6: \ \{c_0,c_1,c_2\}\cong\{-47.5,-41.2,-7.25\}. \tag{56}$$

These estimates are obtained in precisely the same way as those of Eq. (54): c_1 and c_2 are identified with their RG values via Eq. (17), and c_0 is determined via Eq. (51) with $N_0(k)$ [see Eq. (46)] evaluated at a value of k which minimizes $\Delta(k)$ [see Eq. (47)]. The coefficient $c_3=0$ for all cases, as evident in the previous section from RG invariance. For the $b_0=-9.2$ case, the minimizing value $k=-0.94$ is the same as for $b_0=-8.9$. For $b_0=-8.6$, the minimum value of $\Delta(k)$ is found to occur when $k=-0.93$.

IV. SCALE DEPENDENCE OF THE THREE-LOOP RATE

If $b_0=-8.9$, the three-loop $b\to cl^-\overline{\nu_l}$ inclusive estimated rate is given by Eq. (1) with

$$S[x(\mu),L(\mu)]=1-1.67x+(-8.9-3.479L)x^2 + (-50.1-42.4L-7.25L^2)x^3. \tag{57}$$

The coefficients of x^3 are those of Eq. (54), as estimated in the previous section. Figure 4 displays a plot of the μ (scale) dependence of the "reduced" three-loop rate (10) with $S[x,L]$ given by Eq. (57). The pole masses m_b and m_c are assumed to be $m_b=4.9$ GeV and $m_c=0.3m_b=1.47$ GeV consistent with values used in [2]. Figure 4 clearly displays a much flatter μ-dependence than the two-loop rate plotted in Fig. 1. In addition to this diminished dependence on the renormalization scale μ, the rate plotted in Fig. 4 also exhibits a distinct minimum at $\mu=1.0$ GeV. At this principal of minimal sensitivity (PMS) value of μ, the successive terms of the series $S[x,L]$ exhibit reasonable convergence:

$$S[x(1.0\ \text{GeV}),L(1.0\ \text{GeV})]=1-0.258-0.049+0.020 \tag{58}$$

$$\frac{\Gamma_{PMS}}{K}=\frac{\Gamma(1.0\ \text{GeV})}{K}=1047\ \text{GeV}^5. \tag{59}$$

Equation (59) corresponds to the minimal-sensitivity value for the rate, as discussed in Sec. II.

Note that the small three-loop contribution to Eq. (58) can be tuned to zero by making only a small change in the choice of the renormalization-scale parameter μ. The value of μ at which the three-loop term vanishes (i.e., the value of μ at which the Fig. 1 two-loop and Fig. 4 three-loop curves intersect) corresponds to the renormalization scale associated with the "fastest apparent convergence" (FAC) of the series $S[x,L]$. This occurs at $\mu=1.18$ GeV:

$$S[x(1.18\ \text{GeV}),L(1.18\ \text{GeV})]=1-0.226-0.058+0 \tag{60}$$

$$\frac{\Gamma_{FAC}}{K}=\frac{\Gamma(1.18\ \text{GeV})}{K}=1051\ \text{GeV}^5. \tag{61}$$

It is striking that the FAC [5] and PMS [4] criteria predict virtually identical rates; moreover, a similar equivalence of rates obtained via these same two criteria is found for the estimated three-loop contribution to the $b\to ul^-\overline{\nu_l}$ rate [13]. In both semileptonic processes, the PMS and FAC momentum scales are comparably small. The PMS and FAC scales (1.0 GeV and 1.18 GeV) for $b\to cl^-\overline{\nu_l}$ are respectively 37% and 44% of the logarithm reference scale $\sqrt{m_b m_c}\cong 2.7$ GeV. For $b\to ul^-\overline{\nu_l}$, $\mu_{PMS}=1.78$ GeV and $\mu_{FAC}=1.84$ GeV, numbers which are 42% and 44% respectively of the logarithm reference scale $m_b(m_b)=4.2$ GeV [3,13].

Although the estimated three-loop rate plotted in Fig. 4 exhibits much less dependence on the renormalization scale μ than the two-loop rate of Fig. 1, we anticipate the existence of residual μ-dependence as a consequence of the truncation of the series (57) after three-loop order. A way to eliminate much of this residual scale dependence is to "undo the truncation" by choosing an appropriate Padé approximant to the series (57). For example, a $[2|1]$ approximant to the series (57) that reproduces its power series to $\mathcal{O}(x^3)$ is

$$S^{[2|1]}(x,L)=\frac{1+A_1(L)x+A_2(L)x^2}{1+B_1(L)x} \tag{62}$$

where

$$A_1(L)=-1.67-\frac{50.1+42.4L+7.25L^2}{8.9+3.48L} \tag{63}$$

$$A_2(L)=-(8.9+3.48L)+1.67\frac{50.1+42.4L+7.25L^2}{8.9+3.48L} \tag{64}$$

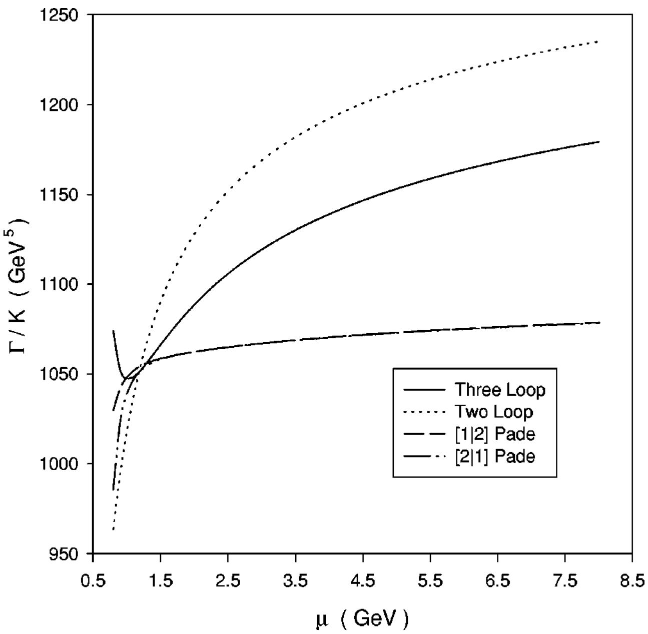

FIG. 5. Scale dependence of different estimates of the reduced rate in the pole scheme. The solid curve represents the three-loop estimate presented in Fig. 4, and the dotted curve represents the two-loop estimate also presented in Fig. 1. The $[1|2]$ and $[2|1]$ Padé approximants obtained from the three-loop estimated rate are represented by the dashed curves which overlap almost completely above $\mu=1.5$ GeV. The PMS point is represented by the local minimum of the three-loop curve, and the FAC point occurs at the intersection of the two- and three-loop curves. Note the convergence of *all* the estimates near the FAC-PMS points.

$$B_1(L)=-\frac{50.1+42.4L+7.25L^2}{8.9+3.48L}. \tag{65}$$

Similarly, the known series terms (57) are also reproduced in the power series of the $[1|2]$ approximant

$$S^{[1|2]}(x,L)=\frac{1+D_1(L)x}{1+E_1(L)x+E_2(L)x^2} \tag{66}$$

$$E_1(L)=-\frac{50.1+42.4L+7.25L^2+1.67(8.9+3.48L)}{1.67^2+8.9+3.48L} \tag{67}$$

$$E_2(L)=\frac{(8.9+3.48L)^2-1.67(50.1+42.4L+7.25L^2)}{1.67^2+8.9+3.48L} \tag{68}$$

$$D_1(L)=E_1(L)-1.67. \tag{69}$$

In Fig. 5 we have superimposed plots of the reduced rate (10) using

(1) Eqs. (4) and (9), the two-loop version of $S[x,L]$ leading to the reduced rate also plotted in Fig. 1,

(2) Eq. (57), the three-loop version of $S[x,L]$ leading to the reduced rate plotted in Fig. 4,

(3) $S^{[2|1]}$, as determined via Eqs. (62)–(65), and

(4) $S^{[1|2]}$, as determined via Eqs. (66)–(69).

The vertical scale of Fig. 5 is magnified compared to that of Figs. 1 and 4 in order to accentuate the differences between reduced rates obtained for each of the above scenarios. We observe from Fig. 5 that both Padé-approximant versions of the rate coincide after $\mu=1.5$ GeV and are considerably flatter than the rate devolving from the three-loop version of $S[x,L]$. Indeed, the three-loop reduced rate is itself quite stable, increasing slowly from 1066 GeV^5 to 1180 GeV^5 as μ increases from 1.5 GeV to 8 GeV. Nevertheless, the two Padé approximant versions of $S[x,L]$ vary only minimally over the same range of μ, increasing from 1058 GeV^5 at $\mu=1.5$ GeV to 1079 GeV^5 at $\mu=8$ GeV. Thus Padé-improvement of the three-loop rate virtually eliminates the residual scale dependence of the naively truncated expression. Such use of Padé approximants to eliminate residual scale dependence is also evident in Fig. 3 of Ref. [13] for the $b\rightarrow ul^-\overline{\nu_l}$ rate, and has been previously discussed in the context of the Bjorken sum-rule [21] as well as in more general terms [22]. Of particular interest, however, is the convergence of all four curves in Fig. 5 to virtually the same PMS or FAC point. This convergence lends further support to the PMS or FAC estimates (59),(61) for the reduced rate.

V. DISCUSSION

The entire analysis presented in Sec. IV can be repeated using the extreme values $b_0=-8.6$ and $b_0=-9.2$, as estimated in [2] (again, see footnote 1), utilizing Eqs. (55) and (56) for the appropriate determinations of three-loop coefficients in conjunction with the known values of a_0 and b_1 (9). One finds the uncertainty in b_0 is reflected in a ± 14 GeV^5 spread in the PMS or FAC value 1050 GeV^5 for the reduced rate.

Other sources of theoretical uncertainty arise from $\alpha_s(m_\tau)=0.33\pm 0.02$ [6,7], $m_b^{pole}=(4.9\pm 0.1)$ GeV [8], and the error that may occur in estimating c_0. We estimate that Padé determinations of c_0 are subject to errors comparable to those of Padé determinations of the RG-accessible coefficients c_1 and c_2; e.g., $|(c_1^{Pade}-c_1^{RG})/c_1^{RG}|\cong 7\%$. If we are conservative and estimate the uncertainty of c_0 to be double that of c_1,

$$\left|\frac{\delta c_0}{c_0}\right|\cong 14\%, \tag{70}$$

the corresponding uncertainty in the reduced rate is ± 38 GeV^5. Consequently, our estimate of the purely perturbative three-loop order $b\rightarrow cl^-\overline{\nu_l}$ rate, as defined by Eq. (10), is

$$\frac{\Gamma^{pert}}{K}=(1050\pm 14\pm 44\pm 115\pm 38)\ \mathrm{GeV}^5, \tag{71}$$

where the listed theoretical uncertainties respectively devolve from the uncertainty in b_0, $\alpha_s(m_\tau)$, m_b^{pole}, and c_0.

Nonperturbative (NP) contributions to the rate may be extracted from Eq. (5.8) of [23], and correspond to the following additional contributions to the series $S[x,L]$:

$$\Delta S^{NP} = \frac{\lambda_1 + 3\lambda_2}{2m_b^2} - \frac{6\left(1 - \frac{m_c^2}{m_b^2}\right)^4 \lambda_2}{m_b^2 F\left(\frac{m_c^2}{m_b^2}\right)}, \qquad (72)$$

where the form factor F is given by Eq. (2) and where

$$-0.5\ \ \mathrm{GeV}^2 \leqslant \lambda_1 \leqslant 0,\ \ \lambda_2 = 0.12\ \ \mathrm{GeV}^2. \qquad (73)$$

This additional NP contribution entails a $\sim 6\%$ reduction in the reduced rate (71):

$$\frac{\Gamma}{K} = \frac{\Gamma^{pert}}{K} - (58.5 \pm 5.5)\ \ \mathrm{GeV}^5 = (992 \pm 217)\ \ \mathrm{GeV}^5, \qquad (74)$$

where the independent sources of uncertainty in Eq. (71) and in the intermediate step of Eq. (74) have been combined additively.

If we identify the predicted $b \to c l^- \overline{\nu_l}$ decay rate with the inclusive semileptonic process $B \to X_c l^- \overline{\nu_l}$ [$l = e$ or μ, but not their sum], we can then relate the aggregate theoretical uncertainty in Eq. (74) to the concomitant theoretical uncertainty in the determination of $|V_{cb}|$:

$$|V_{cb}| = \left[\frac{192\pi^3 \hbar\ \mathrm{BR}(B \to X_c l^- \overline{\nu_l})}{G_F^2 \tau_B [(992 \pm 217)\ \ \mathrm{GeV}^5]} \right]^{1/2}. \qquad (75)$$

To factorize experimental and theoretical uncertainties, we employ recent central values for the average B lifetime τ_B [25] and the $B \to X_c e^- \overline{\nu_e}$ branching ratio [26] to rewrite Eq. (75) in the following form:

$$|V_{cb}| = (0.0453^{+0.0060}_{-0.0043}) \times \left(\frac{1.564 \times 10^{-12}\ \mathrm{s}}{\tau_B} \right)^{1/2} \left(\frac{\mathrm{BR}(B \to X_c l^- \overline{\nu_l})}{0.1105} \right)^{1/2}. \qquad (76)$$

The first factor in Eq. (76) reflects the summed theoretical uncertainties in Eq. (74), which are separately broken down in Eq. (71). We have been conservative in identifying and assessing the magnitude of each such independent source of error—the theoretical uncertainty estimated from a (truncated) two-loop calculation of the $b \to c l^- \overline{\nu_l}$ rate should be *larger* than that obtained by us in Eq. (76).

Note also that we can compare our $\Gamma^{pert}/K = 1050\ \ \mathrm{GeV}^5$ central value estimate (71) for the reduced rate (exclusive of NP effects) with the corresponding estimate one would obtain using low-scale masses [2]:[3]

$$\frac{\Gamma^{pert}}{K} = \tilde{m}_b^5 F\left(\frac{\tilde{m}_c^2}{\tilde{m}_b^2}\right) \left[1 - 1.14 \frac{\alpha_s(\sqrt{\tilde{m}_b \tilde{m}_c})}{\pi} - (3.5 \pm 0.3) \times \left(\frac{\alpha_s(\sqrt{\tilde{m}_b \tilde{m}_c})}{\pi} \right)^2 \right]. \qquad (77)$$

If we utilize the low-scale mass values $\tilde{m}_b = 4.64\ \ \mathrm{GeV}, \tilde{m}_c = 1.25\ \ \mathrm{GeV}$, as quoted in [2] from Ref. [24] and find via devolution from $\alpha_s(m_\tau) = 0.33$ that $\alpha_s(\sqrt{\tilde{m}_b \tilde{m}_c})/\pi = 0.091$, we observe that the rate predicted via Eq. (77) is $1097\ \ \mathrm{GeV}^5$, an answer in surprisingly good agreement with our $1050\ \ \mathrm{GeV}^5$ three-loop estimate in the pole-mass renormalization scheme.

We conclude by noting that the RG equations (6),(7) may be used to determine additional higher-order corrections to the decay rate (1). The leading-log corrections to all orders in x are determined by the one-loop β-function; next-to-leading-log corrections to all orders in x are determined by the two-loop β-function etc. Although we have made use of RG invariance to $\mathcal{O}(x^3)$ in present work, it is in fact possible to incorporate these logarithmic corrections to all subsequent orders. The full exploitation of RG invariance within perturbative-QCD expressions for physical processes is presently under study.

ACKNOWLEDGMENTS

We are grateful for the hospitality of the KEK Theory Group, where this research was conducted with financial support from the International Opportunity Fund of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the International Collaboration Program 2001 of Enterprise Ireland.

[3]The $\mathcal{O}(\alpha_s^2)$ coefficient in Eq. (77) [A. Czarnecki (personal communication)] differs slightly from the value -2.65 ± 0.4 appearing in Ref. [2].

[1] A. Czarnecki, Phys. Rev. Lett. **76**, 4124 (1996); A. Czarnecki and K. Melnikov, *ibid.* **78**, 3630 (1997); A. Czarnecki, K. Melnikov, and N. Uraltsev, Phys. Rev. D **57**, 1769 (1998).
[2] A. Czarnecki and K. Melnikov, Phys. Rev. D **59**, 014036 (1999).
[3] T. van Ritbergen, Phys. Lett. B **454**, 353 (1999).
[4] P.M. Stevenson, Phys. Rev. D **23**, 2916 (1981).
[5] G. Grunberg, Phys. Lett. **95B**, 70 (1980); Phys. Rev. D **29**, 2315 (1984).
[6] ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. C **4**, 409 (1998); G. Cvetic and T. Lee, Phys. Rev. D **64**, 014030 (2001); C.J. Maxwell and A. Mirjalili, Nucl. Phys. **B611**, 423 (2001).
[7] T.G. Steele and V. Elias, Mod. Phys. Lett. A **13**, 3151 (1998).
[8] A.H. Hoang, Phys. Rev. D **61**, 034005 (2000); **59**, 014039

(1999).
[9] T. van Ritbergen, J.A.M. Vermaseren, and S.A. Larin, Phys. Lett. B **400**, 379 (1997).
[10] N. Gray, D.J. Broadhurst, W. Grafe, and K. Schilcher, Z. Phys. C **48**, 673 (1990); D.J. Broadhurst, N. Gray, and K. Schilcher, *ibid.* **52**, 111 (1991); J. Fleischer, F. Jegerlehner, O.V. Tarosov, and O.L. Veretin, Nucl. Phys. **B539**, 671 (1999); K.G. Chetyrkin and M. Steinhauser, *ibid.* **B573**, 617 (2000); Phys. Rev. Lett. **83**, 4001 (1999).
[11] A. Hoang, Z. Ligeti, and A.V. Manohar, Phys. Rev. D **59**, 074017 (1999).
[12] K.G. Chetyrkin, Phys. Lett. B **404**, 161 (1997).
[13] M.R. Ahmady, F.A. Chishtie, V. Elias, and T.G. Steele, Phys. Lett. B **479**, 201 (2000).
[14] J. Ellis, M. Karliner, and M.A. Samuel, Phys. Lett. B **400**, 176 (1997).
[15] I. Jack, D.R.T. Jones, and M.A. Samuel, Phys. Lett. B **407**, 143 (1997); I. Jack, D.R.T. Jones, and A. Pickering, *ibid.* **435**, 61 (1998).
[16] F.A. Chishtie and V. Elias, Phys. Lett. B **499**, 270 (2001).
[17] V. Elias, T.G. Steele, F. Chishtie, R. Migneron, and K. Sprague, Phys. Rev. D **58**, 116007 (1998); V. Elias, F.A. Chishtie, and T.G. Steele, J. Phys. G **26**, 1239 (2000).
[18] F.A. Chishtie, V. Elias, and T.G. Steele, J. Phys. G **26**, 93 (2000).
[19] F.A. Chishtie and V. Elias, Phys. Lett. B **521**, 434 (2001).
[20] F. Chishtie, V. Elias, and T.G. Steele, Phys. Rev. D **59**, 105013 (1999).
[21] J. Ellis, E. Gardi, M. Karliner, and M.A. Samuel, Phys. Lett. B **366**, 4380 (1996); Phys. Rev. D **54**, 6986 (1996).
[22] E. Gardi, Phys. Rev. D **56**, 68 (1997).
[23] E. Bagan, P. Ball, V.M. Braun, and P. Godzinsky, Nucl. Phys. **B432**, 3 (1994).
[24] I.I.Y. Bigi, M.A. Shifman, and N. Uraltsev, Annu. Rev. Nucl. Part. Sci. **47**, 591 (1997).
[25] M. Battaglia, Nucl. Phys. B (Proc. Suppl.) **96**, 443 (2001).
[26] BELLE double-lepton tag data reported by H. Ishino at YITP (Kyoto) Conference, 2001.