PHYSICAL REVIEW D VOLUME 51, NUMBER 3 1 FEBRUARY 1995

Inconclusive inclusive nonleptonic B decays

Adam F. Falk*
Department of Physics, University of California, San Diego, La Jolla, California 92093

Mark B. Wise
California Institute of Technology, Pasadena, California 91125

Isard Dunietz
Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510

(Received 26 May 1994)

We reconsider the conflict between recent calculations of the semileptonic branching ratio of the B meson and the experimentally measured rate. Such calculations depend crucially on the application of "local duality" in nonleptonic decays, and we discuss the relation of this assumption to the weaker assumptions required to compute the semileptonic decay rate. We suggest that the discrepancy between theory and experiment might be due to the channel with two charm quarks in the final state, either because of a small value for m_c or because of a failure of local duality. We examine the experimental consequences of such solutions for the charm multiplicity in B decays.

PACS number(s): 13.25.Hw, 12.38.Bx, 12.39.Hg, 13.20.He

I. INTRODUCTION

Because of the large energy which is released, the decay of a heavy quark is essentially a short distance process. This simple observation has led to much recent progress in the calculation of the inclusive decays of hadrons containing a heavy quark [1–7]. The method relies on the construction of a systematic expansion in the inverse of the energy release, given approximately by the heavy quark mass, and hence works most reliably in the bottom system. In fact, it is expected that certain features of inclusive bottom hadron decays may be reliably predicted with the accuracy of a few percent.

Considerable attention has been paid to inclusive semileptonic [2–5] and rare [6,7] B decays, both to total rates and to lepton and photon energy spectra. There is little controversy that these calculations rest on a firm theoretical foundation. However, it has been suggested to extend these methods to include nonleptonic decays as well [6,8]. This proposal has led to an intriguing conflict with experiment, as the predicted nonleptonic widths differ significantly from those which may be extracted from the measured semileptonic branching ratio of the B meson [9]. In this calculation, the short-distance expansion has been carried out to third order in the inverse mass $1/m_b$, and a reasonable analysis leads the authors of Ref. [9] to the conclusion that it would be unnatural to find the source of the discrepancy in uncalculated terms of higher dimension or higher order in α_s.

It is the purpose of this article to reconsider this problem, in particular the assumptions on which the computation is based. In Sec. II, we review the techniques used to treat inclusive decay rates, with an eye to emphasizing the differences between the theoretical foundations underlying the calculations of semileptonic and nonleptonic decays. In Sec. III, we discuss the possible discrepancy between theory and experiment in the B semileptonic branching ratio. This might be resolved by an unusually small value for m_c, or might involve the failure of the key assumption, "local duality," underlying the calculation of the nonleptonic rate. In either case the enhancement of decays into final states with two charm quarks is a likely consequence.

In Sec. IV we examine the implications of this for the charm multiplicity in B decays, for which present data do not seem to support an enhancement resulting from the $b \to c\bar{c}s$ process. The unusual feature of the data on inclusive B decays is neither the semileptonic branching ratio alone, nor the charm multiplicity alone, but rather the combination of the two. Brief concluding remarks are given in Sec. V.

II. THEORETICAL TECHNIQUES

The weak decay of b quarks is mediated by operators of the form

$$\mathcal{O} = J_h^\mu J_{\ell\mu}\,, \tag{2.1}$$

where

$$\begin{aligned} J_h^\mu &= \bar{q}\gamma^\mu(1-\gamma^5)b\,, \\ J_\ell^\mu &= \bar{q}_1\gamma^\mu(1-\gamma^5)q_2, \qquad \text{or} \qquad \bar{\ell}\gamma^\mu(1-\gamma^5)\nu_\ell \end{aligned} \tag{2.2}$$

*Current address: The Johns Hopkins University, Baltimore, MD 21218.

0556-2821/95/51(3)/1183(9)/$06.00

are fermion bilinears. The inclusive decay rate is given by a sum over all possible final states X with the correct quantum numbers:

$$\Gamma \sim \sum_X \langle B|\, \mathcal{O}^\dagger \,|X\rangle\langle X|\, \mathcal{O}\,|B\rangle\,. \qquad (2.3)$$

In this article we adopt the notation that a generic B meson contains a b quark, rather than a $\bar{b}$ quark. The optical theorem may be used to rewrite Eq. (2.3) as the imaginary part of a forward scattering amplitude:

$$\Gamma \sim \mathrm{Im}\,\langle B|\, T\left\{\mathcal{O}^\dagger, \mathcal{O}\right\}\,|B\rangle\,. \qquad (2.4)$$

One then would like to use perturbative QCD to extract information about the time-ordered product appearing in Eq. (2.4). The extent to which this is possible is precisely the extent to which inclusive decay rates may be calculated reliably.

In the case of semileptonic decays, one may follow a systematic procedure to justify the application of perturbative QCD [1]. Up to negligible corrections of order $\alpha_{\rm EM}$ and G_F, one may factorize the matrix element of the four-fermion operator,

$$\langle X\,\ell(p_\ell)\bar{\nu}(p_{\bar\nu})|\, J_h^\mu J_{\ell\mu}\,|B\rangle = \langle X|\, J_h^\mu\,|B\rangle\langle \ell(p_\ell)\bar{\nu}(p_{\bar\nu})|\, J_{\ell\mu}|0\rangle\,, \qquad (2.5)$$

and consider only the time-ordered product of the quark currents. One then finds an expression in which the integral over the momenta of the leptons is explict:

$$\Gamma \sim \int dy\, dv\cdot\hat{q}\, d\hat{q}^2\, L_{\mu\nu}(v\cdot\hat{q},\hat{q}^2,y)\, W^{\mu\nu}(v\cdot\hat{q},\hat{q}^2)\,, \qquad (2.6)$$

where $L_{\mu\nu}$ is the lepton tensor and $W^{\mu\nu}$ the hadron tensor. Here the momentum of the external b quark is written as $p_b^\mu = m_b v^\mu$. The other independent kinematic variables are $q^\mu = p_\ell^\mu + p_{\bar\nu}^\mu$ and $y = 2E_\ell/m_b$. It is convenient to scale all momenta by m_b, so $\hat{q} = q/m_b$. The hadronic tensor is given by

$$\begin{aligned} W^{\mu\nu} &= \sum_X \langle B|\, J_h^{\mu\dagger}\,|X\rangle\langle X|\, J_h^\nu\,|B\rangle \\ &= -2\mathrm{Im}\,\langle B|\, i\int dx\, e^{iq\cdot x}\, T\left\{J_h^{\mu\dagger}(x), J_h^\nu(0)\right\}\,|B\rangle \\ &\equiv -2\mathrm{Im}\, T^{\mu\nu}\,. \end{aligned} \qquad (2.7)$$

One may perform the integrations in y, $v\cdot\hat{q}$, and $\hat{q}^2$ in Eq. (2.6) to compute the total semileptonic decay rate, or leave some of them unintegrated to obtain various differential distributions.

The doubly differential distribution $d\Gamma/dy\,d\hat{q}^2$ is a useful case to consider. Here we must perform the integration over $v\cdot\hat{q}$, for y and $\hat{q}^2$ fixed. The range of integration for $v\cdot\hat{q}$ is given by $(y+\hat{q}^2/y)/2 \le v\cdot\hat{q} \le (1+\hat{q}^2-\hat{m}_q^2)/2$, where m_q is the mass of the quark to which the b decays, and $\hat{m}_q = m_q/m_b$. This integration is pictured in Fig. 1(a), along with the analytic structure of $T^{\mu\nu}$ in the $v\cdot\hat{q}$ plane[1] [1,10]. The absence of a cut along the real axis in the region $(1+\hat{q}^2-\hat{m}_q^2)/2 < v\cdot\hat{q} < [(2+\hat{m}_q)^2-\hat{q}^2-1]/2$ is simple to understand in terms of the invariant mass p_H of the intermediate hadronic state. Such a state may contain no b quarks, in which case it is subject to the restriction $p_H^2 = (m_b v - q)^2 \ge m_q^2$ (the left-hand cut), or it may contain $bb\bar{q}$, in which case $p_H^2 = (m_b v + q)^2 \ge (2m_b + m_q)^2$ (the right-hand cut). Except in the limit $\hat{q}^2 = \hat{q}^2_{\rm max} = (1-\hat{m}_q)^2$ and $m_q \to 0$, the two cuts do not pinch.

In Fig. 1(a), we have already included only the imaginary part of $T^{\mu\nu}$ by integrating over the top of the cut and then back underneath it. In general, $T^{\mu\nu}$ along the physical cut will depend on $v\cdot\hat{q}$ in a complicated nonperturbative way. We do not necessarily know how to compute in QCD in the physical region where there are threshold effects. However, we may use Cauchy's theorem to deform the contour of integration until it lies away from the cut everywhere except at its end points, as illustrated in Fig. 1(b). Along the new contour, we are far from the physical region, and we may perform an operator product expansion for $T^{\mu\nu}$ in perturbative QCD. Only far from any physical intermediate states is such a calculation necessarily valid. However, this is enough to allow us to compute reliably certain smooth integrals of $T^{\mu\nu}$ by deforming the contour of integration into the un-

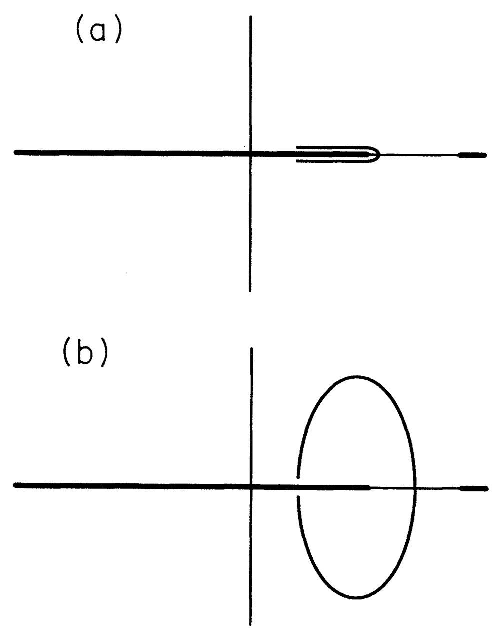

FIG. 1. Contours in the complex $v\cdot\hat{q}$ plane, for fixed $\hat{q}^2$ and y. The gap between the cuts extends for $(1+\hat{q}^2-\hat{m}_q^2)/2 < v\cdot\hat{q} < [(2+\hat{m}_q)^2-\hat{q}^2-1]/2$. The end points of the contour integral are at $v\cdot\hat{q} = (y+\hat{q}^2/y)/2 \pm i\epsilon$.

[1]The discussion of the analytic structure of $T^{\mu\nu}$ given in Ref. [3] is erroneous. We thank B. Grinstein and A.I. Vainshtein for discussions of this point.

physical region. That we can compute integrals of $T^{\mu\nu}$ in perturbation theory in this way is the property of "global duality."

Unfortunately, the contour in Fig. 1 must still approach the physical cut near the end points of the integration. This introduces an uncertainty into the calculation which cannot be avoided. Still, one has two arguments that this uncertainty is likely to be small. First, for large m_b, the portion of the contour which is within $\Lambda_{\rm QCD}$ of the physical cut scales as $\Lambda_{\rm QCD}/m_b$ and thus makes a small contribution to the total integral. Second, if the energy release into the intermediate hadronic system is large compared to $\Lambda_{\rm QCD}$, it is reasonable to expect that $T^{\mu\nu}$ will be well approximated by perturbative QCD even in the physical region. This is because in this region the cut is dominated by multiparticle states, and hence the strength of the imaginary part of $T^{\mu\nu}$ is a relatively smooth function of the energy. While new thresholds associated with the production of additional pions are found along the cut even in this region, their effect is small compared to the smooth background of states to which they are being added.

This intuition, that for large enough energies one may perform the operator product expansion directly in the physical region, is "local duality." While it is a reasonable property for QCD to have, it is obviously a stronger assumption than that of global duality. In particular, it cannot be justified by analytic continuation into the complex plane. Rather, it rests on one's sense of how QCD ought to behave at high energies. It is clear, as well, that the energy at which local duality takes effect will depend on the operators which appear in the time-ordered product. Hence the fact that local duality appears to work at a given energy in one process, such as in electron-positron annihilation into hadrons, may be suggestive but does not prove that it should hold at the same energy in another process.

To compute the inclusive semileptonic decay rate, then, one may use global duality except in a region along the contour of order $\Lambda_{\rm QCD}/m_b$, where one must approach the physical cut. In this small region one must resort to local duality to justify the operator product expansion.

Let us now turn to inclusive nonleptonic decays. Here there is no analogue of the factorization (2.5) which we had in the semileptonic case. Hence there is no "external" momentum q in which one may deform the contour away from the physical region, leaving one unable to use global duality in the transition to perturbative QCD. In this case, one is forced to invoke local duality from the outset if one is to argue that the the time-ordered product $T^{\mu\nu}$ is computable. This clearly puts the calculation of inclusive nonleptonic B decays on a less secure theoretical foundation than that of inclusive semileptonic B decays.

Nonetheless, we do not mean to assert that the assumption of local duality in nonleptonic decays is inherently unreasonable, merely that it is the least reliable aspect of the computation. In fact, it is not entirely clear what it is reasonable to expect in this case. On the one hand, the energy released when a b quark decays is certainly large compared to $\Lambda_{\rm QCD}$. On the other, the decay is initially into three strongly interacting particles (rather than into only one for semileptonic decays), and the energy per strongly interacting particle is not really so large. (Note that in the semileptonic case, the point at which the contour approaches the cut and local duality must be invoked is conveniently the point of *maximum* recoil of the final state quark, where local duality is expected to work best.) What we propose is that the comparison of the nonleptonic decay rate, as computed via the operator product expansion, with experiment be taken as a direct test of local duality in this process. As such, it is a probe of a property of QCD in an interesting kinematic region, and nonleptonic B decay well deserves the intense scrutiny which it has recently been accorded.

III. THE SEMILEPTONIC BRANCHING FRACTION OF B MESONS

The experimental implications of inclusive nonleptonic decays of B mesons have recently been discussed in great detail by Bigi, Blok, Shifman, and Vainshtein [9]. Since the semileptonic branching ratio of the B is relatively well measured, they use their calculation of the nonleptonic decay rate to predict this quantity. Their conclusion is that the semileptonic branching ratio which comes out of their computation is unacceptably high, corresponding to a nonleptonic width which is too low by at least 15–20 %. In this section we will reconsider their analysis.

The inclusive decay rate of the B meson may be divided into parts based on the flavor quantum numbers of the final state,

$$\Gamma_{\rm tot} = \Gamma(b\to c\,\ell\bar\nu) + \Gamma(b\to c\bar u d') + \Gamma(b\to c\bar c s')\,. \quad (3.1)$$

Here we neglect rare processes, such as those mediated by an underlying $b \to u$ transition or penguin-induced decays. By d' and s' we mean the approximate flavor eigenstates ($d' = d\cos\theta_1 - s\sin\theta_1$, $s' = d\sin\theta_1 + s\cos\theta_1$) which couple to u and c, respectively, and we ignore the effect of the strange quark mass. It is convenient to normalize the inclusive partial rates to the semielectronic rate, defining

$$R_{ud} = \frac{\Gamma(b\to c\bar u d')}{3\Gamma(b\to c\,e\bar\nu)}\,, \qquad R_{cs} = \frac{\Gamma(b\to c\bar c s')}{3\Gamma(b\to c\,e\bar\nu)}\,. \quad (3.2)$$

The full semileptonic width may be written in terms of the semielectronic width as

$$\Gamma(b\to c\,\ell\bar\nu) = 3f(\hat m_\tau)\Gamma(b\to c\,e\bar\nu)\,, \quad (3.3)$$

where the factor $3f(\hat m_\tau)$ accounts for the three flavors of lepton, with a phase space suppression which takes into account the τ mass. Then, since the semileptonic branching ratio is given by $B(b\to c\,\ell\bar\nu) = \Gamma(b\to c\,\ell\bar\nu)/\Gamma_{\rm tot}$, we may rewrite Eq. (3.1) in the form

$$R_{ud}+R_{cs}=f(\hat m_\tau)\,\frac{1-B(b\to c\ell\bar\nu)}{B(b\to c\ell\bar\nu)}\,. \tag{3.4}$$

The measured partial semileptonic branching fractions are [11,12]

$$\begin{aligned} B(B\to Xe\bar\nu)&=10.7\pm0.5\%\,,\\ B(B\to X\mu\bar\nu)&=10.3\pm0.5\%\,,\\ B(B\to X\tau\bar\nu)&=2.8\pm0.6\%\,, \end{aligned} \tag{3.5}$$

leading to a total semileptonic branching fraction $B(b\to c\ell\bar\nu)$ of $23.8\pm0.9\%$, with the experimental errors added in quadrature. Of the semileptonic rate, 11% comes from decays to τ, corresponding to a phase space suppression factor $f(\hat m_\tau)=0.74$, consistent with what one would expect in free quark decay [5,13]. If we substitute the measured branching fractions into the right-hand side of Eq. (3.4), we find

$$R_{ud}+R_{cs}=2.37\pm0.12\,. \tag{3.6}$$

We now compare this constraint with the theoretical calculations of R_{ud} and R_{cs}.

The ratios R_{ud} and R_{cs} depend on the total rates $\Gamma(b\to ce\bar\nu)$, $\Gamma(b\to c\bar ud')$, and $\Gamma(b\to c\bar cs')$. Each of these has a theoretical expansion in terms of $\alpha_s(\mu)$ and $1/m_b$. Since corrections of order $1/m_b$ vanish and those of order $1/m_b^2$ are numerically expected to be at the few percent level [1–4,6,7,9], we include here only the radiative corrections. Neglecting terms of order $\alpha_s^2(\mu)$, the expansions take the form

$$\begin{aligned} \Gamma(b\to ce\bar\nu)&=\Gamma_0 I(\hat m_c,0)\left\{1-\frac{2\alpha_s(\mu)}{3\pi}\left(\pi^2-\frac{25}{4}+\delta_{\rm s.l.}(\hat m_c)\right)\right\},\\ \Gamma(b\to c\bar ud')&=\Gamma_0 I(\hat m_c,0)3\eta(\mu)\left\{1-\frac{2\alpha_s(\mu)}{3\pi}\left(\pi^2-\frac{31}{4}+\delta_{ud}(\hat m_c)\right)+J_2(\mu)\right\},\\ \Gamma(b\to c\bar cs')&=\Gamma_0 I(\hat m_c,0)3\eta(\mu)G(\hat m_c)\left\{1-\frac{2\alpha_s(\mu)}{3\pi}\left(\pi^2-\frac{31}{4}+\delta_{cs}(\hat m_c)\right)+J_2(\mu)\right\}. \end{aligned} \tag{3.7}$$

The prefactor $\Gamma_0=G_F^2m_b^5|V_{cb}|^2/192\pi^3$ will cancel in the ratios R_{ud} and R_{cs}, as will the charm quark phase space suppression $I(\hat m_c,0)$ [13], to be discussed below.

The radiative corrections have been computed analytically to order α_s in the limit $m_c=0$, and for semileptonic decays up to one numerical integration for general m_c [14]. For semileptonic decays we absorb the correction due to $m_c\neq0$ into $\delta_{\rm sl}(\hat m_c)$ and present the numerical value of $\delta_{\rm sl}(\hat m_c)$ below. Finite charm mass effects for nonleptonic decays are absorbed into $\delta_{ud}(\hat m_c)$ and $\delta_{cs}(\hat m_c)$. Because these quantities have not been computed, we present numerical results in the case of nonleptonic decays only for $m_c=0$. The expressions for $\Gamma(b\to c\bar ud')$ and $\Gamma(b\to c\bar cs')$ in Eq. (3.7) are somewhat more complicated than that for $\Gamma(b\to ce\bar\nu)$, due to renormalization group running between $\mu=M_W$ and $\mu=m_b$. The leading logarithms are resummed into $\eta=(2L_+^2+L_-^2)/3$, where [15]

$$L_+=\left[\frac{\alpha_s(\mu)}{\alpha_s(M_W)}\right]^{-6/23},\quad L_-=\left[\frac{\alpha_s(\mu)}{\alpha_s(M_W)}\right]^{12/23}. \tag{3.8}$$

The subleading logarithms, which must be included if terms of order $\alpha_s(\mu)$ are also to be kept, are assembled into J_2:

$$\begin{aligned} J_2=&\frac{2\alpha_s(\mu)}{3\pi}\left(\frac{19}{4}+6\ln\frac{\mu}{m_b}\right)\frac{L_-^2-L_+^2}{2L_+^2+L_-^2}\\ &+2\left(\frac{\alpha_s(\mu)-\alpha_s(M_W)}{\pi}\right)\frac{2L_+^2\rho_++L_-^2\rho_-}{2L_+^2+L_-^2}, \end{aligned} \tag{3.9}$$

where $\rho_+=-\frac{6473}{12696}=-0.51$ and $\rho_-=\frac{9371}{6348}=1.47$ arise from two-loop anomalous dimensions [16]. The factor 3 in Eq. (3.7) is for the sum over colors in the final state. Finally, there is an additional phase space suppression $G(\hat m_c)$ in $\Gamma(b\to c\bar cs')$ because of the masses of the two charm quarks. This factor is given by [13]

$$G(\hat m_c)=\frac{I(\hat m_c,\hat m_c)}{I(\hat m_c,0)}\,, \tag{3.10}$$

where

$$\begin{aligned} I(x,0)&=(1-x^4)(1-8x^2+x^4)-24x^4\ln x\,,\\ I(x,x)&=\sqrt{1-4x^2}(1-14x^2-2x^4-12x^6)\\ &\quad+24x^4(1-x^4)\ln\left(\frac{1+\sqrt{1-4x^2}}{1-\sqrt{1-4x^2}}\right). \end{aligned} \tag{3.11}$$

In terms of the theoretical expressions (3.7) for the partial widths, the ratios take the form

$$\begin{aligned} R_{ud}&=P(\mu)+\delta P_{ud}(\mu,\hat m_c)\,,\\ R_{cs}&=G(\hat m_c)\Big[P(\mu)+\delta P_{cs}(\mu,\hat m_c)\Big]\,, \end{aligned} \tag{3.12}$$

where

$$P(\mu) = \eta(\mu)\left[1 + \frac{\alpha_s(\mu)}{\pi} + J_2(\mu)\right] ,$$
$$\delta P_{ud}(\mu, \hat{m}_c) = \eta(\mu)\frac{2\alpha_s(\mu)}{3\pi}\left[\delta_{\rm sl}(\hat{m}_c) - \delta_{ud}(\hat{m}_c)\right] , \quad (3.13)$$
$$\delta P_{cs}(\mu, \hat{m}_c) = \eta(\mu)\frac{2\alpha_s(\mu)}{3\pi}\left[\delta_{\rm sl}(\hat{m}_c) - \delta_{cs}(\hat{m}_c)\right]$$

parametrize the radiative corrections. As emphasized in Ref. [9], if the theoretical expressions (3.12) are inserted, then Eq. (3.6) is not well satisfied. For example, if one simply takes the reasonable values $\mu = m_b = 4.8$ GeV, $m_c = 1.5$ GeV, $\Lambda_{\overline{\rm MS}}^{(5)} = 180$ MeV, and $\delta P_{ud} = \delta P_{cs} = 0$, then $P(\mu) = 1.27$, $G(\hat{m}_c) = 0.36$ and the left-hand side of Eq. (3.6) is only 1.73 (where $\overline{\rm MS}$ denotes the modified minimal subtraction scheme). We are thus tempted to push the uncertainties in the calculation as far as is reasonable, in order to see how much of the discrepancy can be resolved within the context of the operator product expansion.

The largest uncertainty in the theoretical expression for $R_{ud} + R_{cs}$ comes from the choice of the charm and bottom masses. Up to certain ambiguities which have recently been discussed [17], within perturbation theory these masses should be taken to be the pole masses [3,18]. These masses have not been determined with much precision. However, within the heavy quark expansion, the difference between m_c and m_b is much more precisely known, in terms of the spin-averaged D meson and B meson masses:

$$m_b - m_c = \langle M_B\rangle_{\rm ave} - \langle M_D\rangle_{\rm ave} = 3.34 \text{ GeV} . \quad (3.14)$$

In what follows, we will hold $m_b - m_c$ fixed, and consider variations of m_b only. A reasonably conservative range for m_b might be 4.4 GeV $\leq m_b \leq$ 5.0 GeV, which corresponds to $0.24 \leq \hat{m}_c \leq 0.33$. In Fig. 2, we plot $G(\hat{m}_c)$ as a function of m_b, using the constraint (3.14). In Fig. 3, we plot $P(\mu)$ for a variety of values of the QCD scale $\Lambda_{\overline{\rm MS}}^{(5)}$ [11].

We start by considering R_{ud}, for which the calculation is likely to be more reliable, since it is less sensitive to $\hat{m}_c$. There is uncertainty in the radiative correction $P(\mu)$ from the choice of the renormalization scale μ. The usual choice $\mu = m_b$ is motivated by the fact that the total energy released in the decay is m_b. However, this energy has to be divided between three particles, so perhaps the appropriate scale is lower. For $\mu = 1.6$ GeV $\approx m_b/3$, a reasonable lower limit, and $\Lambda_{\overline{\rm MS}}^{(5)} = 180$ MeV, we find $P(\mu) = 1.45$, a modest enhancement over $\mu = 4.8$ GeV. If $\Lambda_{\overline{\rm MS}}^{(5)}$ is taken as high as 220 MeV, we have $P(\mu) = 1.52$, which makes a small additional difference. The uncertainty in δP_{ud} is harder to estimate, since $\delta_{ud}(\hat{m}_c)$ has not been calculated. However, one may extract $\delta_{\rm sl}(\hat{m}_c)$ by doing a numerical integration of the formulas in Ref. [14]. For $\hat{m}_c = 0.30$, we find $\delta_{\rm sl} = -1.11$, corresponding to $[2\alpha_s(m_b)/3\pi]\delta_{\rm sl}(\hat{m}_c) = -0.050$. The magnitude of this correction grows approximately linearly with $\hat{m}_c$, and for $\hat{m}_c = 0.33$, we have $\delta_{\rm sl} = -1.20$. Hence the term is small and actually reduces R_{ud}, although one might expect it to cancel in whole or in part against the term proportional to $\delta_{ud}(\hat{m}_c)$. What we can conclude at this point is that the error associated with ignoring the charm quark mass in the radiative corrections is likely to be no larger than ± 0.05, and henceforth we will neglect this effect.

The leading nonperturbative strong interaction corrections to R_{ud} and R_{cs} are characterized by the two dimensionless quantities $K_b = -\langle B(v)|\,\bar{b}_v(iD)^2 b_v\,|B(v)\rangle/2m_b^2$ and $G_b = \langle B(v)|\,\bar{b}_v g_s G_{\mu\nu}\sigma^{\mu\nu} b_v\,|B(v)\rangle/4m_b^2$. Because it breaks the heavy quark spin symmetry, the parameter G_b may be determined from the measured B^*-B mass splitting, but the value of K_b is not known. Fortunately, K_b does not occur in the nonperturbative correction to R_{ud}. (Using the "smearing" technique of Ref. [3], this cancellation arises because $\Gamma(b \to c\bar{u}d)$ and $\Gamma(b \to ce\bar{\nu})$ have the same dependence on m_b.) For R_{ud}, then, we are more confident than for R_{cs} that the nonperturbative QCD corrections are small. Note, however, that there is a con-

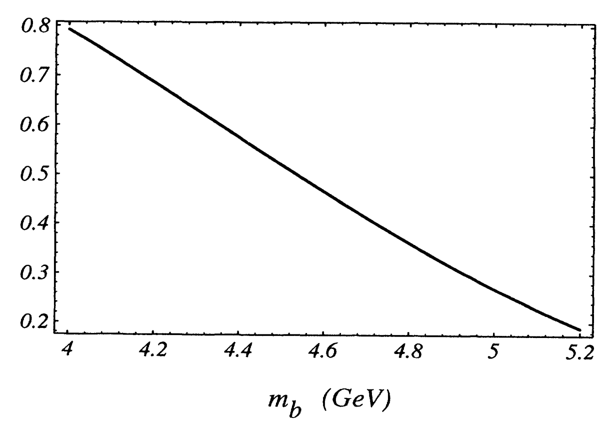

FIG. 2. The phase space suppression factor $G(\hat{m}_c)$, as an implicit function of m_b with $m_b - m_c = 3.34$ GeV held fixed.

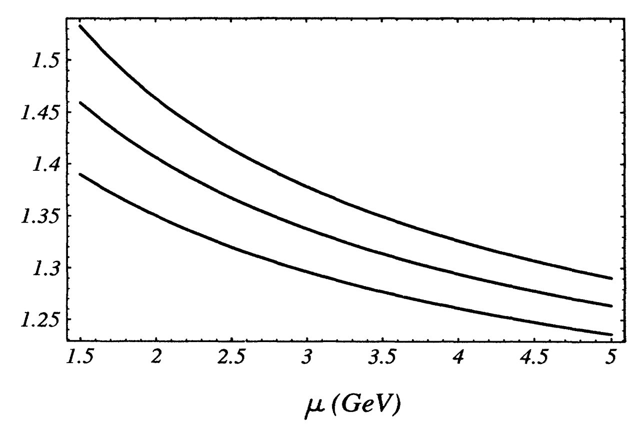

FIG. 3. The radiative correction $P(\mu)$. The upper curve corresponds to $\Lambda_{\overline{\rm MS}}^{(5)} = 220$ MeV, the middle curve to $\Lambda_{\overline{\rm MS}}^{(5)} = 180$ MeV, and the lower curve to $\Lambda_{\overline{\rm MS}}^{(5)} = 140$ MeV. We take $m_b = 4.8$ GeV.

tribution to the mass difference in Eq. (3.14) involving K_b and K_c, which we have neglected.

The above estimates lead us to the conclusion that with the effects we have included in the operator product expansion, it is difficult to avoid the upper bound $R_{ud} \leq 1.52$. If this is true, then Eq. (3.6) would imply $R_{cs} \geq 0.85$. This can barely be achieved in the theoretical expressions we have given. If we vary $4.4\ \text{GeV} \leq m_b \leq 5.0\ \text{GeV}$, as suggested above, then $0.27 \leq G(\hat{m}_c) \leq 0.58$. Estimating the radiative corrections as before, with $\Lambda^{(5)}_{\overline{\text{MS}}} = 220$ MeV, this suggests the upper limit $R_{cs} \leq 0.89$, or $R_{ud} + R_{cs} \leq 2.43$. This is in agreement with experiment, but on the other hand, it requires us to push all the freedom in the calculation in the same direction, perhaps further than is reasonable. If one were to take the point of view that $\mu = 2.4\ \text{GeV} \approx m_b/2$ were the lowest reasonable value for μ, then one would have the constraints $R_{ud} \leq 1.44$, $R_{cs} \leq 0.83$, and $R_{ud} + R_{cs} \leq 2.27$. If one were further to require $m_b \geq 4.6$ GeV, one would have $R_{cs} \leq 0.67$ and $R_{ud} + R_{cs} \leq 2.10$. In this case, one might consider the discrepancy with experiment to be a more serious issue.

Another possibility is that the relevant scale for the radiative corrections in the decay to two charm quarks is considerably lower than that for the final state with a single charm. Since the rest masses of the two charm quarks absorb approximately 60% of the energy available in the decay, the strongly interacting particles are not emitted with very large momenta. For example, the average energy of the strange quark in the decay $b \to c\bar{c}s$, computed at the tree level, is only about 1 GeV. With such a low energy the procedure of estimating the value of higher order QCD corrections by varying the subtraction point μ is of dubious value. In fact one might question whether any finite order of perturbation theory is adequate and whether threshold effects that cause a violation of local duality are important.

It is evident from this discussion that nothing is particularly clear. Although the data on inclusive nonleptonic decays can almost be accounted for by squeezing the input parameters, one might feel a little nervous about the necessity of such a conspiracy. After all, as mentioned earlier the "reasonable" values $\mu = m_b = 4.8$ GeV, $m_c = 1.5$ GeV, and $\Lambda^{(5)}_{\overline{\text{MS}}} = 180$ MeV lead to $R_{ud} = 1.27$ and $R_{cs} = 0.46$, far short of the mark. An enhancement of approximately 40% in the nonleptonic rate is called for. If one were to require this effect to be found entirely in R_{cs}, it would amount to more than a factor of two. While we are less inclined than the authors of Ref. [9] to insist that something is amiss, it is nonetheless intriguing to consider the possibility that the data indicate an enhancement of the nonleptonic rate over and above what we have included in the operator product expansion. Where might such an enhancement come from?

The simplest explanation would be that due to a failure of local duality, the inclusive nonleptonic decay rate is simply not calculable to better than 40% or so. This is certainly a discouraging explanation, in that if it were true then there would be very little one could say in detail about why local duality, and hence the calculation, had failed. One was simply unlucky. On the other hand, this explanation may well be correct. While we expect local duality to hold in the asymptotic limit of infinite b quark mass, we have little to guide us in estimating how heavy the b quark actually needs to be in practical terms. In particular, it is not relevant to consider, at low orders in QCD perturbation theory, the size of a few subleading terms which appear in the operator product expansion itself. The matrix elements which appear in this expansion are sensitive to details of the B meson bound state, but they are explicitly not sensitive to resonance effects in the final hadronic state.

If local duality fails, it could well fail differently in the $\Gamma(b \to c\bar{u}d')$ and $\Gamma(b \to c\bar{c}s')$ channels. In fact, we would expect it to fail worse in the channel with two charm quarks, since we expect the final states to be characterized by lower particle multiplicity and be closer to the resonance-dominated regime. Local duality, by contrast, is applicable only in the regime where the effect of individual resonance thresholds is small compared to the almost smooth "continuum" of multiparticle states. On the other hand, the phase space suppression from the two final state charm quarks means that unless m_c is unusually small, only 30% or so of the inclusive nonleptonic rate comes from the $\Gamma(b \to c\bar{c}s')$ channel. Hence, to account for an enhancement of the full nonleptonic rate by 40% purely from $b \to c\bar{c}s'$ would require a dramatic failure of local duality in this channel.

IV. EXPERIMENTAL CONSEQUENCES OF AN ENHANCEMENT OF R_{cs}

Either through a failure of local duality, or from an unusually small value for m_c, or because of a combination of these effects, the value of R_{cs} is likely to be near unity in order to account for the measured B semileptonic branching ratio. This corresponds to about one-third of B decays arising from the $b \to c\bar{c}s'$ process. One consequence of this is a large number of charmed quarks per B decay:

$$n_c = 1 + R_{cs}\,\frac{B(B \to X_c \ell\bar{\nu})}{f(\hat{m}_\tau)}\,. \tag{4.1}$$

We remind the reader that we have adopted the notation that a generic B meson contains a b quark, rather than a $\bar{b}$ quark. Using $B(B \to X_c\,\ell\bar{\nu}) = 23.8\%$ and $f(\hat{m}_\tau) = 0.74$ in Eq. (4.1) yields

$$n_c = 1.00 + 0.32\,R_{cs}\,, \tag{4.2}$$

which for the values of R_{cs} necessary to explain the semileptonic branching ratio would indicate $n_c \sim 1.3$.

There are contributions to the experimental value of n_c from charmed mesons, charmed baryons, and $c\bar{c}$ resonances. The number of charged and neutral D mesons per decay, summed over B and $\overline{B}$, has been measured to be [19]

$$n_{D^\pm} = 0.246 \pm 0.031 \pm 0.025\,,$$
$$n_{D^0,\overline{D}^0} = 0.567 \pm 0.040 \pm 0.023\,. \tag{4.3}$$

The branching ratio to $D_s^\pm$ mesons has not yet been determined, because no absolute D_s branching ratio has been measured. However, it is known that [20]

$$n_{D_s^\pm} = (0.1224 \pm 0.0051 \pm 0.0089)\left[\frac{3.7\%}{B(D_s \to \phi\pi)}\right], \tag{4.4}$$

and the branching ratio for $D_s \to \phi\pi$ is expected to be about 3.7%.

We must include in n_c twice the inclusive branching ratio to all $c\bar{c}$ resonances which are below $D\overline{D}$ threshold. The measured inclusive branching ratio to ψ is $(1.11 \pm 0.08)\%$, including feed-down from ψ' and χ_c decays [19]. It is also known that $B(B \to \psi' X) = (0.32 \pm 0.05)\%$, $B(B \to \chi_{c1} X) = (0.66 \pm 0.20)\%$, and $B(B \to \eta_c X) < 1\%$. Hence we expect that the inclusive B branching ratio to charmonium states below $D\overline{D}$ threshold is about 2%.

The inclusive B decay rate to baryons is about 6% [19]. While it is commonly believed that these baryons arise predominantly from the $b \to c\bar{u}d'$ process, giving $\Lambda_c X$ final states, we argue elsewhere [21] that a large fraction of B decays to baryons may actually arise from the $b \to c\bar{c}s'$ process, which gives final states with both a charm baryon and an anticharm baryon, such as $\Xi_c \overline{\Lambda}_c X$. While we have no firm theoretical justification for such a prediction, experimental evidence for this interpretation comes from the observed distribution of Λ_c momenta, which shows that the Λ_c's produced in B or $\overline{B}$ decay are recoiling against a state with a mass greater than or equal to the mass of the Ξ_c [22–24]. This novel interpretation of B decays to baryons can be consistent with the measured $\Lambda\ell^\pm$ correlations if $B(\Xi_c \to \Lambda X)/B(\Lambda_c \to \Lambda X)$ is large [21].

Even if B decay to baryons predominantly gives final states with both a charm and an anticharm baryon, the data summarized above do not provide supporting evidence for a value of n_c around 1.3. Given the uncertainties, however, such a large value for the number of charmed hadrons per B decay is perhaps not excluded. From our perspective the curious feature of the data on inclusive B decay is not the measured semileptonic branching ratio alone, but rather the combination of it with the data on charm multiplicity in these decays.

In this paper we have neglected B decays that do not arise from an underlying $b \to c$ transition. Other possible processes include the $b \to u$ transition and contributions from penguin-type diagrams. While it is very unlikely that such sources contribute significantly to the nonleptonic decay rate, this assumption can be tested experimentally, if enough branching ratios can be measured. The fraction of B decays arising from the $b \to c$ transition is given by the sum of the B branching ratio to charmonium states below $D\overline{D}$ threshold, the branching ratio to states containing at least one charmed baryon, and the branching ratios to the ground state charmed mesons, $B(B \to D^0 X)$, $B(B \to D^+ X)$, and $B(B \to D_s^+ X)$. Note that the inclusive charm yields reported in Eqs. (4.3) and (4.4) are actually sums of branching ratios [for example, neglecting CP violation, $n_{D^\pm} = B(B \to D^+ X) + B(B \to D^- X)$]. However, it should be possible with enough data to extract the individual branching ratios themselves. For example, one could perform a careful study of DD and $D_s D_s$ sign correlations in decays of the $\Upsilon(4S)$, properly taking into account the effects of coherence and B-$\overline{B}$ mixing. Another method for determining individual branching ratios would involve tagging the flavor of the B which produced the charmed hadron by measuring the charge of a hard primary lepton from the other B in the event.

Invoking a large violation of local duality has some implications for the pattern of B meson decays which may be different from what would be expected if local duality held and an unusually small value of m_c were used to explain the measured B semileptonic branching ratio. For example, a violation of local duality in the $b \to c\bar{c}s$ channel could lead to quite different lifetimes for the B, B_s, and Λ_b, differences which are small in the operator product expansion because they arise only from higher dimension operators. However, since the effective Hamiltonian for this process has isospin zero, the equality of the B^0 and B^- lifetimes would not be disturbed. Similarly, violations of local duality in $b \to c\bar{u}d'$ could lead to unequal B^0 and B^- lifetimes. B decay event shapes can also provide a test of the free quark decay picture for the $b \to c\bar{u}d'$ decay channel [25].

V. CONCLUDING REMARKS

We have examined whether the measured B meson semileptonic branching ratio can be explained within the conventional application of the operator product expansion, in which operators of low dimension are kept and perturbative corrections are included to a few orders in α_s.

We have found that this scenario would require an unusually small value for m_c. If instead the explanation lies outside the conventional application of the operator product expansion, then a failure of local duality in the $b \to c\bar{c}s'$ channel is the likely explanation for the discrepancy with experiment. In either case, we expect the number of charmed hadrons per B decay to be approximately 1.3. Unfortunately, the present data on charm multiplicities do not support such a large value of n_c. From our perspective, the unusual feature of inclusive B decay is not the semileptonic branching ratio alone, nor the charm multiplicity alone, but rather the combination of the two. Together, they would seem to suggest a significant violation of local duality in the $b \to c\bar{u}d'$ nonleptonic decay process. From a theoretical point of view, however, such a resolution would be somewhat unsettling, as it would indicate a breakdown in the computation of the nonleptonic decay rate in the region where it is expected to be the most reliable; we understand why such a conclusion was resisted by the authors of Ref. [9]. Still, it remains an open possibility, indicating perhaps that the invocation of local duality in quark decay requires a considerably larger energy release than has been naïvely hoped or expected. Given the apparent difficulties in performing a reliable computation of the nonleptonic decay rate, then, the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{cb} should be extracted from the B semileptonic decay width rather than from the B lifetime. The uncertainties in such an extraction arise primarily from the choice of m_b and subtraction point μ, and are discussed in detail in Refs. [26,27].

Note added

After this paper was completed, we became aware of a number of additional calculations of radiative corrections to inclusive decay widths. First, the order α_s radiative correction to the semileptonic decay width has been calculated analytically, including all dependence on m_c/m_b, by Y. Nir [28]. Second, the correction δ_{ud} to R_{ud} has recently been computed by Bagan *et al.* [29], in which they find $[2\alpha_s(m_b)/3\pi]\delta_{ud} \approx -0.05$. Hence the effect of the charm mass on the radiative correction to the nonleptonic width indeed approximately cancels the analogous effect on the semileptonic width, as we supposed in Sec. III. The authors also estimate the charm mass corrections to R_{cs}, and speculate that they may be much larger.[2] However, as we have shown, the semileptonic branching ratio cannot be reconciled with the reported charm multiplicity merely by supposing an enhancement of R_{cs}. Unfortunately, the same is true of the result of Hokim and Pham [31], who also find evidence for a large δ_{cs} in an order α_s calculation in which the large logarithms between $\mu = M_W$ and $\mu = m_b$ are not resummed.

There is also recent work [32] which suggests that the scale μ should be taken even lower than we had supposed, perhaps $\mu = m_b/10$. The one loop expression for $P(\mu)$ which we have used here is not sensible below $\mu \approx 700$ MeV, at which point it takes its maximum value of 1.7. Such a value of $P(\mu)$ is not enough of an enhancement to resolve the problem we have discussed, as it still leads to the prediction $n_c \geq 1.2$. However, this result may indicate that the calculation of R_{ud} and R_{cs} is unreliable not because of a failure of local duality *per se*, but because the appropriate renormalization scale is too small for a perturbative calculation to make sense.

Finally, in a recent measurement the CLEO Collaboration has reported that approximately 20% of B decays with baryons in the final state produce two charmed baryons [33]. In the same publication, they report an inclusive branching ratio of of B to Ξ_c of 0.04. Together, these constitute an additional contribution to n_c of about 0.05.

[2] A subsequent calculation by Voloshin [30] supports this conclusion.

ACKNOWLEDGMENTS

We are indebted to B. Blok, T. E. Browder, P. S. Cooper, B. Grinstein, J. D. Lewis, M. Luke, Y. Nir, M. Savage, M. Shifman, N. G. Uraltsev, and A. I. Vainshtein for useful conversations. A.F. and I.D. would like to thank the Institute for Theoretical Physics, where this work was initiated, for their gracious hospitality. I.D. acknowledges the support of the Texas National Research Laboratory Commission under Grant No. FCFY9303. This work was supported by the Department of Energy under Grants Nos. DOE-FG03-90ER40546, DE-AC03-81ER40050, and DE-AC02-76CHO3000, and by the National Science Foundation under Grant No. PHY89-04035.

[1] J. Chay, H. Georgi, and B. Grinstein, Phys. Lett. B **247**, 399 (1990).
[2] I.I. Bigi, M. Shifman, N.G. Uraltsev, and A.I. Vainshtein, Phys. Rev. Lett. **71**, 496 (1993); B. Blok, L. Koyrakh, M. Shifman, and A.I. Vainshtein, Phys. Rev. D **49**, 3356 (1994).
[3] A. Manohar and M.B. Wise, Phys. Rev. D **49**, 1310 (1994).
[4] T. Mannel, Nucl. Phys. **B413**, 396 (1994).
[5] A.F. Falk, Z. Ligeti, M. Neubert, and Y. Nir, Phys. Lett. B **326**, 145 (1994); L. Koyrakh, Phys. Rev. D **49**, 3379 (1994); S. Balk, J.G. Körner, D. Pirjol, and K. Schilcher, Z. Phys. C **64**, 37 (1994).
[6] I.I. Bigi, N.G. Uraltsev, and A.I. Vainshtein, Phys. Lett. B **293**, 430 (1992); **297**, 477(E) (1993); I.I. Bigi, B. Blok, M. Shifman, N.G. Uraltsev, and A.I. Vainshtein, Minnesota Report No. TPI–MINN–92/67–T, 1992 (unpublished).
[7] A.F. Falk, M. Luke, and M.J. Savage, Phys. Rev. D **49**, 3367 (1994).
[8] B. Blok and M. Shifman, Nucl. Phys. **B399**, 441 (1993); **B399**, 459 (1993); N. Bilic, B. Guberina, and J. Trampetic, *ibid.* **B248**, 261 (1984); M. Voloshin and M. Shifman, Yad. Fiz. **41**, 187 (1985) [Sov. J. Nucl. Phys. **41**, 120 (1985)]; M. Voloshin and M. Shifman, Zh. Eksp. Teor. Fiz. **91**, 1180 (1986) [Sov. Phys. JETP **64**, 698 (1986)]; V.A. Khoze, M. Shifman, N.G. Uraltsev, and M. Voloshin, Yad. Fiz. **46**, 181 (1987) [Sov. J. Nucl. Phys. **46**, 112 (1987)].
[9] I.I. Bigi, B. Blok, M.A. Shifman, and A. Vainshtein, Phys. Lett. B **323**, 408 (1994).
[10] J. Chay and S.-J. Rey, Seoul Report No. SNUTP 94–08, 1994 (unpublished).
[11] Particle Data Group, K. Hikasa *et al.*, Phys. Rev. D **45**, S1 (1992).
[12] A. Putzer, Heidelberg Report No. HD–IHEP–93–03, 1993

(unpublished).
[13] J.L. Cortes, X.Y. Pham, and A. Tounsi, Phys. Rev. D **25**, 188 (1982).
[14] M. Ježabek and J.H. Kühn, Nucl. Phys. **B314**, 1 (1989).
[15] M.K. Gaillard and B. Lee, Phys. Rev. Lett. **33**, 108 (1974); G. Altarelli and L. Maiani, Phys. Lett. **52B**, 351 (1974).
[16] G. Altarelli and S. Petrarca, Phys. Lett. B **261**, 303 (1991); G. Altarelli *et al.*, Phys. Lett. **99B**, 141 (1981); G. Altarelli *et al.*, Nucl. Phys. **B187**, 461 (1981).
[17] M. Beneke and V.M. Braun, Michigan Report No. UM-TH-94-4, 1994 (unpublished); I.I. Bigi, M.A. Shifman, N.G. Uraltsev, and A.I. Vainshtein, Phys. Rev. D **50**, 2234 (1994).
[18] B. Grinstein, Nucl. Phys. **B339**, 253 (1990); A.F. Falk, M. Luke, and M. Neubert, *ibid.* **B388**, 363 (1992).
[19] T.E. Browder, K. Honscheid, and S. Playfer, Cornell Report No. CLNS 93/1261, to appear in *B Decays*, 2nd ed., edited by S. Stone (World Scientific, Singapore, in press).
[20] CLEO Collaboration, S. Menary (private communication).
[21] I. Dunietz, P.S. Cooper, A.F. Falk, and M.B. Wise, Phys. Rev. Lett. **73**, 1075 (1994).
[22] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **210**, 263 (1988).
[23] CLEO Collaboration, G. Crawford *et al.*, Phys. Rev. D **45**, 752 (1992).
[24] CLEO Collaboration, M.M. Zoeller, Ph.D. thesis, submitted to the State University of New York, Albany, 1994.
[25] M. Luke, M. Savage, and M.B. Wise, Phys. Lett. B **322**, 154 (1994).
[26] M. Luke and M. Savage, Phys. Lett. B **321**, 88 (1994).
[27] M. Shifman, N.G. Uraltsev, and A. Vainshtein, Phys. Rev. D (to be published).
[28] Y. Nir, Phys. Lett. B **221**, 184 (1989).
[29] E. Bagan *et al.*, Barcelona Report No. UAB-FT-347, 1994 (unpublished).
[30] M.B. Voloshin, Minnesota Report No. TPI-MINN-94/35-T, 1994 (unpublished).
[31] Q. Hokim and X.Y. Pham, Ann. Phys. (N.Y.) **155**, 202 (1984).
[32] M. Luke, M. Savage, and M.B. Wise, Caltech Report No. CALT-68-1950, 1994 (unpublished).
[33] CLEO Collaboration, D. Cinabro *et al.*, CLEO Report No. CLEO-CONF-94-08, 1994 (unpublished).