First Measurement of the Rate for the Inclusive Radiative Penguin Decay $b \to s\gamma$

M. S. Alam,[1] I. J. Kim,[1] Z. Ling,[1] A. H. Mahmood,[1] J. J. O'Neill,[1] H. Severini,[1] C. R. Sun,[1] F. Wappler,[1] G. Crawford,[2] C. M. Daubenmier,[2] R. Fulton,[2] D. Fujino,[2] K. K. Gan,[2] K. Honscheid,[2] H. Kagan,[2] R. Kass,[2] J. Lee,[2] M. Sung,[2] C. White,[2] A. Wolf,[2] M. M. Zoeller,[2] F. Butler,[3] X. Fu,[3] B. Nemati,[3] W. R. Ross,[3] P. Skubic,[3] M. Wood,[3] M. Bishai,[4] J. Fast,[4] E. Gerndt,[4] J. W. Hinson,[4] R. L. McIlwain,[4] T. Miao,[4] D. H. Miller,[4] M. Modesitt,[4] D. Payne,[4] E. I. Shibata,[4] I. P. J. Shipsey,[4] P. N. Wang,[4] M. Battle,[5] J. Ernst,[5] L. Gibbons,[5] Y. Kwon,[5] S. Roberts,[5] E. H. Thorndike,[5] C. H. Wang,[5] T. Coan,[6] J. Dominick,[6] V. Fadeyev,[6] I. Korolkov,[6] M. Lambrecht,[6] S. Sanghera,[6] V. Shelkov,[6] T. Skwarnicki,[6] R. Stroynowski,[6] I. Volobouev,[6] G. Wei,[6] M. Artuso,[7] M. Gao,[7] M. Goldberg,[7] D. He,[7] N. Horwitz,[7] G. C. Moneti,[7] R. Mountain,[7] F. Muheim,[7] Y. Mukhin,[7] S. Playfer,[7] Y. Rozen,[7] S. Stone,[7] X. Xing,[7] G. Zhu,[7] J. Bartelt,[8] S. E. Csorna,[8] Z. Egyed,[8] V. Jain,[8] D. Gibaut,[9] K. Kinoshita,[9] P. Pomianowski,[9] B. Barish,[10] M. Chadha,[10] S. Chan,[10] D. F. Cowen,[10] G. Eigen,[10] J. S. Miller,[10] C. O'Grady,[10] J. Urheim,[10] A. J. Weinstein,[10] M. Athanas,[11] W. Brower,[11] G. Masek,[11] H. P. Paar,[11] J. Gronberg,[12] C. M. Korte,[12] R. Kutschke,[12] S. Menary,[12] R. J. Morrison,[12] S. Nakanishi,[12] H. N. Nelson,[12] T. K. Nelson,[12] C. Qiao,[12] J. D. Richman,[12] A. Ryd,[12] D. Sperka,[12] H. Tajima,[12] M. S. Witherell,[12] R. Balest,[13] K. Cho,[13] W. T. Ford,[13] D. R. Johnson,[13] K. Lingel,[13] M. Lohner,[13] P. Rankin,[13] J. G. Smith,[13] J. P. Alexander,[14] C. Bebek,[14] K. Berkelman,[14] K. Bloom,[14] T. E. Browder,[14,*] D. G. Cassel,[14] H. A. Cho,[14] D. M. Coffman,[14] D. S. Crowcroft,[14] P. S. Drell,[14] D. J. Dumas,[14] R. Ehrlich,[14] P. Gaidarev,[14] M. Garcia-Sciveres,[14] B. Geiser,[14] B. Gittelman,[14] S. W. Gray,[14] D. L. Hartill,[14] B. K. Heltsley,[14] S. Henderson,[14] C. D. Jones,[14] S. L. Jones,[14] J. Kandaswamy,[14] N. Katayama,[14] P. C. Kim,[14] D. L. Kreinick,[14] G. S. Ludwig,[14] J. Masui,[14] J. Mevissen,[14] N. B. Mistry,[14] C. R. Ng,[14] E. Nordberg,[14] J. R. Patterson,[14] D. Peterson,[14] D. Riley,[14] S. Salman,[14] M. Sapper,[14] F. Würthwein,[14] P. Avery,[15] A. Freyberger,[15] J. Rodriguez,[15] S. Yang,[15] J. Yelton,[15] D. Cinabro,[16] T. Liu,[16] M. Saulnier,[16] R. Wilson,[16] H. Yamamoto,[16] T. Bergfeld,[17] B. I. Eisenstein,[17] G. Gollin,[17] B. Ong,[17] M. Palmer,[17] M. Selen,[17] J. J. Thaler,[17] K. W. Edwards,[18] M. Ogg,[18] A. Bellerive,[19] D. I. Britton,[18] E. R. F. Hyatt,[19] D. B. MacFarlane,[19] P. M. Patel,[19] B. Spaan,[19] A. J. Sadoff,[20] R. Ammar,[21] P. Baringer,[21] A. Bean,[21] D. Besson,[21] D. Coppage,[21] N. Copty,[21] R. Davis,[21] N. Hancock,[21] M. Kelly,[21] S. Kotov,[21] I. Kravchenko,[21] N. Kwak,[21] H. Lam,[21] Y. Kubota,[22] M. Lattery,[22] M. Momayezi,[22] J. K. Nelson,[22] S. Patton,[22] R. Poling,[22] V. Savinov,[22] S. Schrenk,[22] and R. Wang[22]

(CLEO Collaboration)

[1]State University of New York at Albany, Albany, New York 12222
[2]Ohio State University, Columbus, Ohio, 43210
[3]University of Oklahoma, Norman, Oklahoma 73019
[4]Purdue University, West Lafayette, Indiana 47907
[5]University of Rochester, Rochester, New York 14627
[6]Southern Methodist University, Dallas, Texas 75275
[7]Syracuse University, Syracuse, New York 13244
[8]Vanderbilt University, Nashville, Tennessee 37235
[9]Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[10]California Institute of Technology, Pasadena, California 91125
[11]University of California, San Diego, La Jolla, California 92093
[12]University of California, Santa Barbara, California 93106
[13]University of Colorado, Boulder, Colorado 80309-0390
[14]Cornell University, Ithaca, New York 14853
[15]University of Florida, Gainesville, Florida 32611
[16]Harvard University, Cambridge, Massachusetts 02138
[17]University of Illinois, Champaign-Urbana, Illinois 61801
[18]Carleton University, Ottawa, Ontario, Canada K1S 5B6
and the Institute of Particle Physics, Montreal, Canada H3A 2T8
[19]McGill University, Montréal, Québec, Canada H3A 2T8
and the Institute of Particle Physics, Montreal, Canada H3A 2T8
[20]Ithaca College, Ithaca, New York 14850
[21]University of Kansas, Lawrence, Kansas 66045
[22]University of Minnesota, Minneapolis, Minnesota 55455
(Received 13 December 1994)

We have measured the inclusive $b \to s\gamma$ branching ratio to be $(2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$, where the first error is statistical and the second is systematic. Upper and lower limits on the branching ratio,

each at 95% C.L., are $\mathcal{B}(b \rightarrow s\gamma) < 4.2 \times 10^{-4}$ and $\mathcal{B}(b \rightarrow s\gamma) > 1.0 \times 10^{-4}$. These limits restrict the parameters of extensions of the standard model.

PACS numbers: 13.25.Hw, 12.60.–i, 13.40.Hq

The transition $b \rightarrow s\gamma$ is a flavor-changing neutral current process. It is described by a penguin diagram in which a virtual W is exchanged in a loop with a top quark, with a photon emitted from any of the lines [1] (see Fig. 1 of Ref. [2]). There are large QCD corrections to the penguin diagram. Progressively more complete calculations, based on the method of renormalization-group-improved perturbation theory, have been applied to the problem [3]. There is now general agreement on the full leading-log calculation [4]. Some of the next-to-leading-logarithmic QCD corrections have been calculated [5]. The QCD corrections increase the rate by a factor of 2–3. The standard model branching ratio for a leading-log calculation is $(2.8 \pm 0.8) \times 10^{-4}$ [6], where the error is dominated by the uncertainty in the renormalization scale $m_b/2 < \mu < 2m_b$. If those next-to-leading-log terms that have been calculated are included, the branching ratio falls to 1.9×10^{-4} [7]. The branching ratio is sensitive to the existence of a charged Higgs [8], anomalous $WW\gamma$ coupling [9], and other non-standard-model phenomena [10].

CLEO's observation [2] of the decay $B \rightarrow K^*(892)\gamma$, the first conclusive evidence for a penguin decay, established the existence of penguin diagrams generally and of $b \rightarrow s\gamma$ in particular. Since there are large theoretical uncertainties in the hadronization process, $\Gamma(B \rightarrow K^*\gamma)$ gives only a rough measure of $\Gamma(b \rightarrow s\gamma)$, the quantity of theoretical interest. Here we present a measurement of the branching ratio for the inclusive process $b \rightarrow s\gamma$.

The data were taken with the CLEO detector at the Cornell Electron Storage Ring (CESR), and consist of 2.01 fb^{-1} on the $\Upsilon(4S)$ resonance and 0.96 fb^{-1} at a center-of-mass energy 60 MeV below the resonance. The on-resonance sample contains 2.15×10^6 $B\bar{B}$ events and 6.6×10^6 continuum events. The CLEO detector [11] measures charged particles over 95% of 4π sr with a system of cylindrical drift chambers. Its barrel and end-cap CsI electromagnetic calorimeters cover 98% of 4π. The energy resolution for photons near 2.5 GeV in the central angular region ($|\cos\theta_\gamma| < 0.7$) is 2%.

Our signature for $b \rightarrow s\gamma$ is a photon from B-meson decay with energy between 2.2 and 2.7 GeV. The Fermi motion of the b quark in a B meson and the momentum of the B meson in the laboratory result in this Doppler-broadened photon line. Spectator model calculations [12] indicate that (75–90)% of the signal lies in this range. Backgrounds from other B-decay processes are small and calculable. There are very large backgrounds from the continuum, both from the initial-state-radiation (ISR) process $e^+e^- \rightarrow q\bar{q}\gamma$, and from the continuum reaction $e^+e^- \rightarrow q\bar{q}$, with the high energy photon arising from the hadronic debris (π^0, η, $\omega \rightarrow \pi^0\gamma$, etc.). We suppress the

continuum with two methods and subtract what remains using off-resonance data.

We select events that pass general hadronic event selection criteria [13]. We further require that the event contain a high energy calorimeter cluster with $|\cos\theta_\gamma| < 0.7$ and unmatched to a charged particle track. The vast majority of photons, from both continuum and $B\bar{B}$ events, are π^0 and η decay products. We discard those high energy clusters which, when paired with another γ in the event, have a $\gamma\gamma$ mass consistent with a π^0 or η. Finally, we require that the lateral energy distribution of the cluster be consistent with that of a single isolated photon, thus suppressing random overlaps, single-cluster π^0's, and nonphoton clusters.

The first method [13] for suppressing continuum background uses eight carefully chosen event-shape variables: R_2, $S_\perp$, R_2', $\cos\theta'$ as defined in Ref. [2], and the energies in 20° and 30° cones, parallel and antiparallel to the high energy photon direction. While no individual variable has strong discriminating power, each possesses some. Consequently, we combine the eight variables into a single variable r which tends towards $+1$ for $b \rightarrow s\gamma$ and tends towards -1 for ISR and $q\bar{q}$. A neural network is used for this task. Distributions in r are shown in Fig. 1. There is substantial discrimination between signal and background, and good agreement between Monte Carlo (MC) background and off-resonance data. We perform a weighted sum of candidate high energy photon events, with an r-dependent weighting, optimized with MC samples of sig-
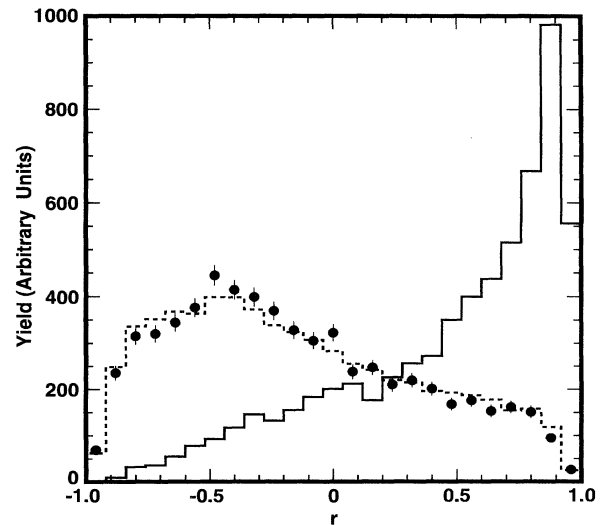

FIG. 1. Distributions in the neural net variable r, for Monte Carlo samples of $b \rightarrow s\gamma$ signal (solid histogram) and continuum background (dashed histogram), and for the off-resonance data sample (points).

nal and background. This weighting procedure is equivalent to performing a one-parameter fit to the r distribution.

In the second method for suppressing the continuum, we search each event for combinations of particles that reconstruct to a $B \rightarrow X_s \gamma$ decay. For X_s we use $K_s^0 \rightarrow \pi^+ \pi^-$ or a charged track with dE/dx consistent with a kaon, and $1-4$ π's, of which one may be a π^0. There are reconstruction ambiguities and cross feed between decay modes, but these are not important because this method is used only to suppress background and *not* for a mode-by-mode B-reconstruction analysis. In each event, we pick the combination that minimizes an overall χ^2, which includes χ_B^2 (see below), together with contributions from dE/dx and K_s^0 and π^0 mass deviations, where relevant. We calculate the momentum P_B, the energy E_B, and the beam-constrained mass $M_B = \sqrt{E_{\rm beam}^2 - P_B^2}$ of the combination, and also $\cos\theta_{tt}$, where θ_{tt} is the angle between the thrust axis of the candidate B and the thrust axis of the rest of the event. We discriminate between signal and background by requiring $|\cos\theta_{tt}| < 0.6$ and $\chi_B^2 < 6.0$, where

$$\chi_B^2 = \left(\frac{M_B - 5.279}{\sigma_M}\right)^2 + \left(\frac{E_B - E_{\rm beam}}{\sigma_E}\right)^2. \quad (1)$$

The two methods for suppressing the continuum are complementary. The B-reconstruction method has lower efficiency for $b \rightarrow s\gamma$ (9% vs 32%) but a factor of 4 better signal-to-noise ratio, so the two methods have nearly equal sensitivity, and are only slightly correlated. To correctly represent on-resonance continuum, a small correction to the off-resonance data is required to account for the change in event features due to the 60 MeV in center-of-mass energy difference. This is obtained from MC, separately for each of the two methods.

There are backgrounds from B decay processes other than $b \rightarrow s\gamma$, in particular, from $b \rightarrow cW^-$ and $b \rightarrow uW^-$. As a first approximation, we take these from MC. We then correct for any difference between the π^0-momentum spectra from data and MC, and similarly for the η-momentum spectra. This procedure corrects for any flaws in the event generators for $b \rightarrow cW^-$ and $b \rightarrow uW^-$, and for any omissions (e.g., $b \rightarrow sg$). Thus we rely on MC only for the π^0 and η veto efficiencies, and for those small B-decay backgrounds *not* from π^0 or η. Finally, there are very small backgrounds from the non-$B\bar{B}$ decays of $\Upsilon(4S)$, which we obtain from $\Upsilon(3S)$ and $\Upsilon(1S)$ data.

The photon energy spectra from the event-shape and B-reconstruction analyses are shown in Figs. 2 and 3, respectively. In both cases, the on-resonance yield exceeds the background in the energy interval $2.2-2.7$ GeV, demonstrating the presence of $b \rightarrow s\gamma$. A signal of the expected shape is evident in the subtracted spectra [14]. Yields between 2.2 and 2.7 GeV are given in Table I.

The B-reconstruction technique selects a "best" X_s candidate for $B \rightarrow X_s \gamma$, thus providing an "apparent
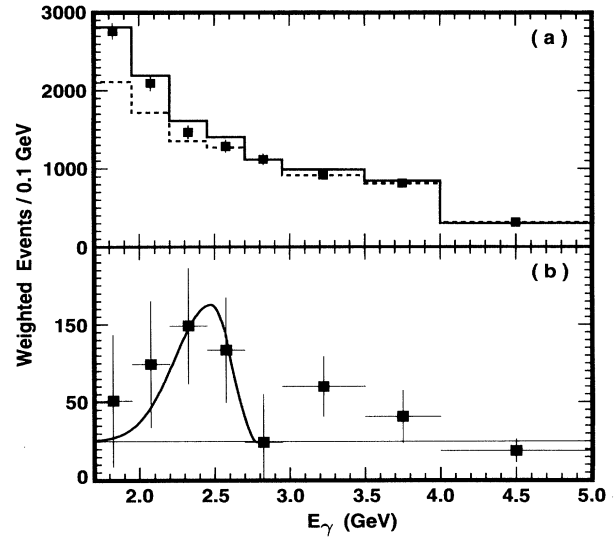
FIG. 2. Photon energy spectra from the event-shape analysis. (a) On-resonance (solid histogram), scaled off-resonance (dashed histogram), and sum of off-resonance and background from $\Upsilon(4S)$ (squares). (b) Background-subtracted data (points) and Monte Carlo prediction for the shape of the $b \rightarrow s\gamma$ signal (solid curve) [14].

X_s mass spectrum" (Fig. 4). Although reconstruction ambiguities and cross feed must be taken into account before quantitative use can be made of it, there is clear evidence both for $B \rightarrow K^*(892)\gamma$ and for X_s systems in the $1-2$ GeV range. The rate for $K^*(892)\gamma$ extracted by a fit to the distribution in Fig. 4 is consistent with our previous measurement [10].

To calculate detection efficiencies, we model the particle content of X_s in $b \rightarrow s\gamma$ with conventional models of quark hadronization [13]. We model the X_s mass dis-
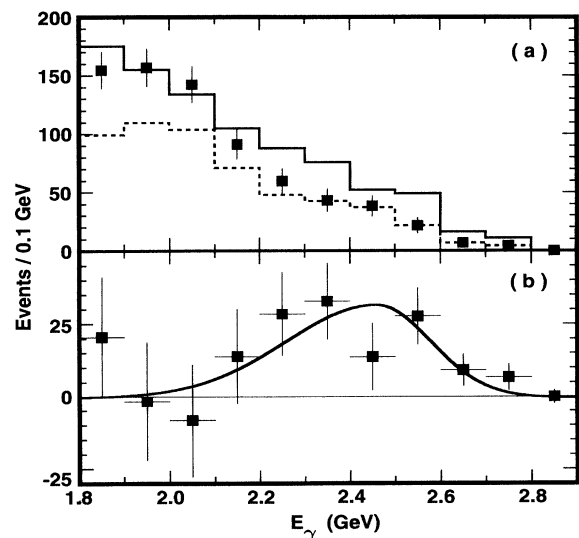
FIG. 3. Photon energy spectra from the B-reconstruction analysis. Symbols are defined as in Fig. 2.

TABLE I.   Yields of events with 2.2–2.7 GeV photons, for the two $b \to s\gamma$ analysis procedures.

	Event shape	B reconstruction
On	3013 ± 59	281 ± 17
Off (scaled)	2618 ± 73	155 ± 18
$\Upsilon(4S)$ background		
$b \to c$ MC	50.7 ± 5.1	12 ± 2
$b \to u$ MC	11.9 ± 4.0	2 ± 1
π^0 correction	50.2 ± 27.7	−0.7 ± 2.3
η correction	16.5 ±33.7	2.0 ± 8.5
Non-$B\bar{B}$	2.3	
$\Upsilon(4S)$ total	132 ± 44	15 ± 9
On-off-4S background	263 ± 104	110 ± 26

tribution with the spectator model of Ali and Greub [12], which includes gluon bremsstrahlung and higher-order radiative effects. In the Ali-Greub model, we vary the Fermi-momentum parameter P_F and spectator-quark mass simultaneously so that the b-quark average mass $\langle m_b \rangle$ is constant at 4.87 ± 0.10 GeV, a value suggested by recent theoretical work [15]. We take $P_F = 270 \pm 40$ MeV/c, based on fits to CLEO $B \to X\ell\nu$ data with the same $\langle m_b \rangle$.

We find $\mathcal{B}(b \to s\gamma) = (1.88 \pm 0.74) \times 10^{-4}$ with the event-shape analysis, and $(2.75 \pm 0.67) \times 10^{-4}$ with the B-reconstruction analysis (statistical errors only). Allowing for correlations, the difference is 1.1 standard deviation. We combine the two results, allowing for correlations, obtaining $\mathcal{B}(b \to s\gamma) = (2.32 \pm 0.57 \pm 0.35) \times 10^{-4}$, where the first error is statistical and the second is systematic (including model dependence). Conservatively allowing for the
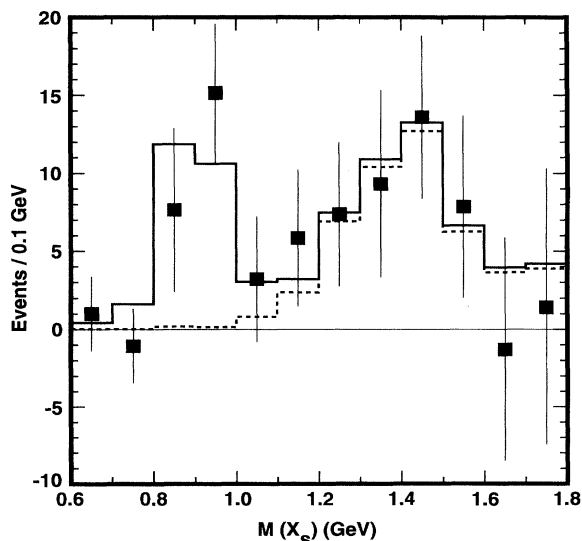

FIG. 4.   Apparent X_s mass distribution from B-reconstruction analysis. Background-subtracted data, not corrected for efficiency or cross feed (points); Monte Carlo fit, using several kaon resonances (solid histogram); component of fit from resonances other than $K^*(892)$ (dotted histogram).

systematic error, we find $\mathcal{B}(b \to s\gamma) < 4.2 \times 10^{-4}$, $\mathcal{B}(b \to s\gamma) > 1.0 \times 10^{-4}$, each limit at 95% C.L.

Uncertainties in yield and efficiency comprise the systematic error in $\mathcal{B}(b \to s\gamma)$. In units of 10^{-4}, these contributions (labeled by source of uncertainty) are ±0.03 (on-off luminosity ratio [16]), ±0.08 (π^0 veto efficiency), ±0.10 (on-off energy difference correction), ±0.08 ($B^+B^-/B^0\bar{B}^0$ production ratio), ±0.23 ($\langle m_b \rangle$), ±0.002 (P_F), ±0.12 (particle content of X_s), and ±0.16 (MC modeling).

Our measurement is in good agreement with the standard model expectation. We illustrate the implications for nonstandard models with two examples [17]. A charged Higgs boson with model II coupling would increase the $b \to s\gamma$ branching ratio [8], so our upper limit on $b \to s\gamma$ provides a lower limit on charged Higgs boson mass: $M_{H^\pm} > [244 + 63/(\tan\beta)^{1.3}]$ GeV, where $\tan\beta = v_2/v_1$, the ratio of vacuum expectation values for the two doublets. With additional non-standard-model effects this limit can be circumvented. For example, in super symmetry, if chargino and scalar top are light, their contribution can cancel the charged Higgs contribution [18].

Anomalous $WW\gamma$ couplings could either increase or decrease the $b \to s\gamma$ branching ratio [9], so both our upper and lower limits rule out portions of the $\Delta\kappa$-λ space that describes these anomalous couplings (Fig. 5). Also shown in Fig. 5 are the regions allowed and excluded by $\bar{p}p \to W\gamma X$ measurements [19]. The two types of measurements are complementary.

In summary, the inclusive branching ratio for $b \to s\gamma$ has been determined using two analysis methods yielding consistent results. Our measurement is in good agreement with standard model predictions, and places constraints on other models. These constraints will improve with completion of a next-to-leading-log calculation.
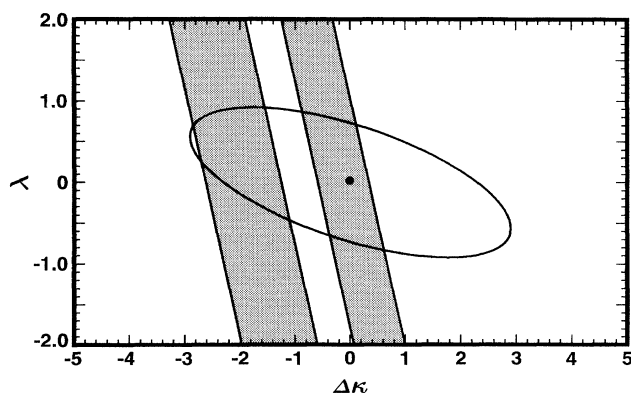

FIG. 5.   Limits on anomalous $WW\gamma$ coupling parameters λ and $\Delta\kappa$. The shaded regions are consistent with the $b \to s\gamma$ branching ratio reported here. The region between the two shaded strips is excluded by the $b \to s\gamma$ lower limit, the outer unshaded regions by the upper limit. D0's yield of $p\bar{p} \to W\gamma X$ limits the allowed range to the interior of the ellipse (CDF obtains a similar ellipse) [19]. The standard model value is shown as the dot at $\Delta\kappa = \lambda = 0$.

*Permanent address: University of Hawaii at Manoa, Honolulu, HI 96822.

[1] B. Campbell and P. O'Donnell, Phys. Rev. D **25**, 1989 (1982).

[2] CLEO Collaboration, R. Ammar *et al.*, Phys. Rev. Lett. **71**, 674 (1993).

[3] S. Bertolini *et al.*, Phys. Rev. Lett. **59**, 180 (1987); N. Deshpande *et al.*, *ibid.* **59**, 183 (1987); B. Grinstein *et al.*, Phys. Lett. B **202**, 138 (1988); R. Grigjanis *et al.*, *ibid.* **213**, 335 (1988); G. Cella *et al.*, *ibid.* **248**, 181 (1990); M. Misiak, *ibid.* **269**, 161 (1991).

[4] M. Ciuchini *et al.*, Phys. Lett. B **316**, 127 (1993); M. Misiak, *ibid.*, **321**, 113 (1994); G. Cella *et al.*, *ibid.* **325**, 227 (1994).

[5] A. Ali and C. Greub, Phys. Lett. B **259**, 182 (1991); Z. Phys. C **60**, 433 (1993); A. Buras *et al.*, Nucl. Phys. **B370**, 69 (1992); **B375**, 501 (1992); **B400**, 37 (1993); **B400** 75 (1993); M. Ciuchini *et al.*, Phys. Lett. B **301**, 263 (1993); Nucl. Phys. **B415**, 403 (1994).

[6] A. Buras *et al.*, Nucl. Phys. **B424**, 374 (1994).

[7] M. Ciuchini *et al.*, Phys. Lett. B **334**, 137 (1994).

[8] R. Ellis *et al.*, Phys. Lett. B **179**, 119 (1986); T. Rizzo, Phys. Rev. D **38**, 820 (1988); B. Grinstein and M. Wise, Phys. Lett. B **201**, 274 (1988); W. Hou and R. Wiley, *ibid.* **202**, 591 (1988); C. Geng and J. Ng, Phys. Rev. D **38**, 2858 (1988); V. Barger, J. Hewett, and R. Phillips, *ibid.* **41**, 3421 (1990).

[9] S. Chia, Phys. Lett. B **240**, 465 (1990); K. Peterson, *ibid.* **282**, 207 (1992); T. Rizzo, *ibid.* **315**, 471 (1993); U. Baur, in Proceedings of the Summer Workshop on B Physics, Snowmass, CO, 1993 (to be published); X. He and B. McKellar, Phys. Lett. B **320**, 165 (1994).

[10] For a review of implications of $b \rightarrow s\gamma$ for nonstandard models, see J. Hewett, Report No. SLAC-PUB-6521 (1994).

[11] CLEO Collaboration, Y. Kubota *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **320**, 66 (1992).

[12] A. Ali and C. Greub, Phys. Lett. B **259**, 182 (1991).

[13] J. A. Ernst, Ph.D. thesis, University of Rochester, 1995.

[14] The fit of the MC prediction to the data in Fig. 2(b) has a χ^2 of 5.8 for 7 degrees of freedom. We have searched for explanations of the high point at 3.0–3.5 GeV, and found none (other than statistical fluctuation). The B-reconstruction method has no efficiency above 2.9 GeV.

[15] M. Voloshin and Y. Zaitsev, Sov. Phys. Usp. **30**, 553 (1987); C. Davies *et al.*, Phys. Rev. Lett. **73**, 2654 (1994).

[16] CLEO Collaboration, G. Crawford *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **345**, 429 (1994).

[17] We use $m_t = 175$ GeV, leading-log terms only, and $m_b/2 < \mu < 2m_b$. We assume a 10% theoretical uncertainty in addition to the renormalization scale uncertainty.

[18] S. Bertolini *et al.*, Nucl. Phys. **B294**, 321 (1987); **B353**, 591 (1991); Y. Okada, Phys. Lett. B **315**, 119 (1993); R. Garisto and J. N. Ng, *ibid.* **315**, 372 (1993).

[19] CDF Collaboration F. Abe *et al.*, Fermilab Report No. Fermilab-Pub-94/236-E; D0 Collaboration, J. Ellison *et al.*, in Proceedings of 1994 Meeting of Division of Particles and Fields, Albuquerque, NM, 1994 (to be published).