FIRST MEASUREMENT OF $b \to s\gamma$

JESSE ERNST *

Dept. of Physics and Astronomy, University of Rochester
Rochester, NY 14627, USA

ABSTRACT

The CLEO Collaboration has measured the inclusive branching ratio for $b \to s\gamma$ and finds $\mathcal{B}\ (b \to s\gamma) = (2.32 \pm 0.51 \pm 0.29 \pm 0.32) \times 10^{-4}$.
This article summarizes the techniques and results of the measurement.

I also report CLEO's upper limits on two other radiative penguin decays; $\mathrm{B} \to \rho(\omega)\gamma$ and $\mathrm{B} \to \mathrm{K}(\mathrm{K}^*)\mathrm{l}^+\mathrm{l}^-$.

1. Introduction

The April 1993, discovery of $B \to K^*(890)\gamma$ by the CLEO collaboration[1] verified the existence of One-Loop Flavor Changing Neutral Current decays (penguins). While measurement of the branching ratio for this exclusive mode demonstrates penguins' existence, the sensitivity of $b \to s\gamma$ to physics beyond the Standard Model[2,3] is realized only with a determination of the inclusive rate. Unfortunately, theoretical predictions for the fraction of $b \to s\gamma$ that hadronizes into $\mathrm{K}^*(890)$ range from 5 to 40%; this uncertainty prevents us from converting the branching ratio for $B \to K^*(890)\gamma$ into an inclusive branching ratio. Here, I discuss the techniques and results of the fully inclusive measurement. $b \to s\gamma$ is an ideal decay channel in which to measure penguin contribution because 1) the GIM mechanism prevents Flavor Changing Neutral Current (FCNC) through tree diagrams, 2) QCD corrections significantly enhance the one-loop rate, and 3) contributions from non-penguin diagrams are expected to be at least an order of magnitude smaller than that from the penguin.

2. Detector and Dataset

CLEO II is a general purpose solenoidal detector. It measures charged particles over 90% of 4π steradians with three nested cylindrical drift chambers. Its barrel and endcap CsI electromagnetic calorimeters cover 98% of 4π. The region used for detecting the high energy photons for this analysis ($\cos(\theta) < .7$) has an energy resolution σ_E/E of 2% near 2.5 GeV. A more complete description of CLEO II is available.[4]

CLEO data are collected at two center of mass energies. The "ON" data are taken at $\mathrm{E}_{cm} = 10.58$ GeV, which is on the $\Upsilon(4S)$ resonance. At this energy, roughly 25% of hadronic events are B meson pairs; the rest are continuum ($\mathrm{u\overline{u}}$, $\mathrm{d\overline{d}}$, $\mathrm{c\overline{c}}$, or $\mathrm{s\overline{s}}$). The "OFF" data are taken at $\mathrm{E}_{cm} = 10.52$ GeV. This energy is below the B meson production threshold, so the events are strictly continuum. Within small corrections, the difference between an ON and an OFF spectrum represents the spectrum from B decays. The dataset used here is 2.01 fb^{-1} ON (2.15×10^6 $\mathrm{B\overline{B}}$ events and 6.6×10^6 continuum events) and .96 fb^{-1} OFF.

---

*Representing the CLEO Collaboration

3. Inclusive $b \to s\gamma$

Our general strategy is simply to count the number of photons from B decays that have energy between 2.2 and 2.7 GeV. $b \to s\gamma$ will contribute photons between 2 and 2.8 GeV, while contributions from most other B decay channels lie below 2.4 GeV. The background from B decays between 2.2 and 2.7 GeV is small and can be estimated. The background from continuum events is large and must be suppressed. We chose 2.2–2.7 GeV rather than 2.0–3.0 GeV as the signal range because 1) B backgrounds become significant below 2.2 GeV and thus including this range would have required an unacceptable reliance on Monte Carlo, and 2) there is little signal above 2.7 Gev, yet the continuum background remains large. Model calculations indicate that 75–90% of $b \to s\gamma$ photons lie between 2.2 and 2.7 GeV. In selecting photon candidates, we require that the shower be well isolated (not too close to a charged track), that it be in the highest resolution portion of the calorimeter ($\cos(\theta) < .7$), and that the energy distribution among the crystals in the shower be consistent with that from a photon. If the photon can be paired with another photon to form a π^0 or an η then the candidate photon is discarded.

Two distinct event types contribute to the continuum background; generic $q\overline{q}$ events $(q\overline{q})$, and $q\overline{q}$ events that undergo Initial State Radiation (ISR). In $q\overline{q}$ events, 83% of the photon background is from π^0 or η decay. In ISR events, the background is the radiated photon. The size of these continuum backgrounds is such that a signal for $b \to s\gamma$ would be lost in statistical noise if we measured the difference between the ON and OFF photon spectra without continuum suppression. We have measured $b \to s\gamma$ with two independent analyses. Both analyses follow the same general strategy, but use different techniques to suppress the continuum background.

3.1. Neural Net Analysis

This analysis uses eight event shape variables to separate signal from the two types of background.

Since the ISR events are $q\overline{q}$ events recoiling against the radiated photon, we can restore the $q\overline{q}$ shape of an ISR event by transforming it to this recoiling frame *. We call this the primed frame.

The B mesons at CLEO are produced nearly at rest and decay isotropically; continuum events are more jet-like. We use R_2, R_2', and $S_\perp$ to measure isotropy. R_2 is the normalized second Fox-Wolfram moment. In the lab frame, R_2 separates signal from $q\overline{q}$; in the primed frame it separates signal from ISR. $S_\perp$ is the fraction of the momentum carried by particles more than 45° from the photon; it separates signal from $q\overline{q}$.

In the primed frame, an ISR photon's direction is uncorrelated with the direction of the continuum jet. We define $Cos(\theta')$ as the cosine of the angle between the candidate photon and the thrust axis of the rest of the event. For ISR events, $|Cos(\theta')|$ will be flat; for signal events, it will peak near +1.

To make use of differences in the energy/angular distribution of particles in signal

*Having lost the energy of the radiated photon, an ISR event in the primed frame has the shape of a reduced energy continuum event.

and in background events, we define four additional variables. We compute the energy in four cones (20° and 30° along and opposed to the photon direction).

All eight variables have some power to discriminate between signal and one or both of the continuum backgrounds. However, in order to gain enough continuum suppression, we must allow for correlations between the variables. To do this, we create a single variable (r) which is a function of all eight shape variables and allows for linear and non-linear correlations between them. The functional form of r is such that its free parameters can be optimized using a neural network algorithm.[5] The network [†] was trained (i.e., the parameters of r were optimized) on a Monte Carlo sample. Once optimized, r tends toward +1 for signal events and −1 for either type of background. We do not make a cut on r. Instead, we gain additional statistical power by weighting events according to their r value. The improvement in continuum suppression over using uncorrelated cuts on the eight variables is equivalent to having 77% more data.

3.2. B-Reconstruction Analysis

In this analysis, we suppress the continuum backgrounds by reconstructing some subset of the particles in the event into a $b \to s\gamma$ candidate. For each event, we consider only one combination of particles; the combination that best satisfies a B$\to X_s\gamma$ decay hypothesis. A fully inclusive analysis is achieved by using a large number of exclusive decay modes. We allow γ with $K^{\pm}$ or K_s^0 (reconstructed in $\pi^+\pi^-$ decays), and up to four pions (which may include zero or one π^0).

For each set of candidate particles, we form the total energy E and momentum P. The energy difference ΔE, beam-constrained mass M, and mass difference ΔM are then calculated as $\Delta E = E - E_{beam}$, $M = \sqrt{E_{beam}^2 - P^2}$, and $\Delta M = M -$ 5.279 GeV. For the correct set of candidate particles, ΔE and ΔM will be zero, within measurement error. Consequently we construct a variable $\chi_B^2 = (\Delta M/\sigma_M)^2 + (\Delta E/\sigma_E)^2$, where σ_E and σ_M are the expected measurement errors. χ_B^2 reflects how well the kinematic properties of the set of candidate particles represent a B meson.

Once we have the best combination of particles, we make cuts on two variables; χ_B^2 and $|\cos(\theta_{tt})|$. χ_B^2 is defined above, and $|\cos(\theta_{tt})|$ is the cosine of the angle between the thrust axis of the particles used to form the candidate B and the thrust axis of the rest of the particles in the event. For a B$\overline{\text{B}}$ event, there is no angular correlation between the decay products of the two B mesons, so $|\cos(\theta_{tt})|$ should be flat. For continuum events, the two thrust angles will tend to align and hence $|\cos(\theta_{tt})|$ will peak near 1. Based on Monte Carlo studies, we require $\chi_B^2 < 6.0$ and $|\cos(\theta_{tt})| < 0.6$.

3.3. Yield and Background From the Υ(4S)

The Υ(4S) background is dominated by γ's from π^0 and η decays. To determine this background, we treat high momentum π^0's and η's as the candidate photon and compare the data with the Monte Carlo predictions. We then apply a correction to the Monte Carlo. With this procedure, we rely on the Monte Carlo predictions only for the very small backgrounds other than π^0 and η decays. Table 1 lists the yields and the Υ(4S) backgrounds for the two analyses. Figures 1a and 2a show the yield from

[†] We use a two layer, fully-connected neural network with 10 nodes in the first layer, and 7 in the second.

ON, OFF (scaled to correct for the smaller OFF luminosity), and OFF+B backgrounds. Figures 1b and 2b show the final yields (the curve gives the shape expected for $b \to s\gamma$).

3.4. Combining the Two Analyses

Both analyses use a spectator model from Ali-Greub[6] to determine efficiency. The spectator model gives the X_s mass distribution. We use conventional models of quark hadronization to describe the decay of the X_s.

The results (with statistical errors only) from the neural-net and the B-reconstruction analyses are $\mathcal{B}(b \to s\gamma) = (1.88 \pm 0.67) \times 10^{-4}$, and $\mathcal{B}(b \to s\gamma) = (2.76 \pm 0.63) \times 10^{-4}$ respectively. Monte Carlo indicates that the two methods have approximately equal statistical power, so we weight the results equally. To compute the combined error, we use Monte Carlo samples to determine the degree of correlation between the analyses. From this, we determine that the results agree within 1.1σ and that the average has a statistical error of $.51 \times 10^{-4}$.

3.5. Systematic Errors

The largest sources of additive systematic errors are uncertainties in subtracting the OFF data from the ON data and uncertainty in the π^0 and η corrections to the Monte Carlo when estimating the $\Upsilon(4S)$ background. The ON–OFF subtraction error is from uncertainty in a correction factor which accounts for the lower E_{cm} of the OFF data. The error in the π^0 and η corrections is statistical.

The largest multiplicative systematic error is from model dependence. We use the Ali-Greub spectator model with $< m_b >= 4.87 \pm 0.10$ GeV and $P_f = 270 \pm 40$ MeV. The uncertainties on P_f and m_b correspond to uncertainties in the X_s mass distribution. Because the efficiency is not completely flat across the various X_s masses and decay modes for either analysis, the variations in X_s mass distribution cause a variation in the overall efficiency.

3.6. Final Results and Implications

We find

$$\mathcal{B}(b \to s\gamma) = (2.32 \pm 0.51 \pm 0.29 \pm 0.32) \times 10^{-4} \quad ,$$

where the first error is statistical, the second is the additive systematic error from uncertainty in yield, and the third is the multiplicative systematic error from uncertainty in efficiency, which includes model dependence. This result is in good agreement with the Standard Model expectations of 2.0×10^{-4} to 3.8×10^{-4}.[7] At 95% CL, we find $1 \times 10^{-4} < \mathcal{B}(b \to s\gamma) < 4 \times 10^{-4}$

One example of the sensitivity of $b \to s\gamma$ to new physics is the limit it places on the charged Higgs mass in a Two-Higgs-Doublet-Model II. In this model, the down and up type quarks acquire mass from separate vacuum expectation values ($\tan(\beta) \equiv v_2/v_1$). All points below the curve in figure 3 are excluded by our upper limit (where we have allowed for theoretical uncertainty in renormalization scale, and an additional theoretical uncertainty of 10%). J. Hewett has recently published a review which covers a wide variety of $b \to s\gamma$ implications.[8]

4. Upper Limits on B $\rightarrow \rho(\omega)\gamma$ and B $\rightarrow$ K(K*)l$^+$l$^-$

CLEO has recently obtained new upper limits from two other radiative penguin decays.

For B $\rightarrow \rho(\omega)\gamma$ we find (at 90% CL) $\mathcal{B}(B^- \rightarrow \rho^-\gamma) < 2.0 \times 10^{-5}$, $\mathcal{B}(\bar{B}^0 \rightarrow \rho^0\gamma) < 2.4 \times 10^{-5}$, and $\mathcal{B}(\bar{B}^0 \rightarrow \omega\gamma) < 1.1 \times 10^{-5}$. If long-distance effects are negligible, we can combine these results with the measurement of $\mathcal{B}(B \rightarrow K^*\gamma)$ to set upper limits on $|V_{td}/V_{ts}|$. Depending on the choice of model, $|V_{td}/V_{ts}| < 0.64 - 0.76$.

For B $\rightarrow$ K(K*)l$^+$l$^-$ modes, we find the following upper limits at 90% CL:
$\mathcal{B}(B^- \rightarrow K^- e^+ e^-) < 1.2 \times 10^{-5}$, $\mathcal{B}(\bar{B}^0 \rightarrow \bar{K}^{*0} e^+ e^-) < 1.6 \times 10^{-5}$,
$\mathcal{B}(B^- \rightarrow K^- \mu^+ \mu^-) < 0.9 \times 10^{-5}$, $\mathcal{B}(\bar{B}^0 \rightarrow \bar{K}^{*0} \mu^+ \mu^-) < 3.1 \times 10^{-5}$,
$\mathcal{B}(B^- \rightarrow K^- e^\pm \mu^\mp) < 1.2 \times 10^{-5}$, and $\mathcal{B}(\bar{B}^0 \rightarrow \bar{K}^{*0} e^\pm \mu^\mp) < 2.7 \times 10^{-5}$. The upper limit for $\mathcal{B}(\bar{B}^0 \rightarrow \bar{K}^{*0} e^+ e^-)$ is within a factor of 3 of the Standard Model prediction.

5. Acknowledgements

I thank Ed Thorndike and Peter C. Kim for assistance in preparing this article.
I also thank W.I. for agreeing to delay a clear definition of non-nutritive variables.

References

1. R. Ammar *et al.* (CLEO Collaboration), Phys. Rev. Lett. **71**, 674 (1993).
2. T. G. Rizzo, Phys Rev. **D38**, 820 (1988); W.-S. Hou and R.S. Wiley, Phys Lett. **B202**, 591 (1988); C.Q.Geng and J.N. Ng, Phys Rev. **D38**, 2858 (1988); V. Barger, J.L. Hewett, and R.J.N. Phillips, Phys. Rev. **D41**, 3421 (1990); J.L. Hewett, Phys. Rev. Lett. **70**, 1045 (1993); V. Barger, M. Berger, and R.J.N. Phillips, Phys. Rev. Lett. **70**, 1368 (1993).
3. S.-P. Chia, Phys. Lett. **B240**, 465 (1990); K.A. Peterson, Phys. Lett. **B282**, 207 (1992); T.G. Rizzo, Phys. Lett. **B315**, 471 (1993); U. Baur, in Proceedings of the *Summer Workshop on B Physics*, Snowmass, CO, 1993; X.-G. He and B. McKellar, Phys. Lett. **B320**, 165 (1994).
4. Y. Kubota *et al.* (CLEO Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. **A320**, 66 (1992).
5. Rumelhart D.E. and McClelland J.L., *Parallel Distributed Processing*, Vol. **1–2** (Bradford Books, MIT Press, Cambridge, MA) 1986.
6. A. Ali and C. Greub, Phys. Lett. **B259**, 182 (1991).
7. S. Bertolini *et al.*, Phys. Rev. Lett. **59**, 180 (1987); N. Deshpande *et al.*, Phys. Rev. Lett. **59**, 183 (1987); B. Grinstein *et al.*, Phys. Lett. **B202**,138 (1988); R. Grigjanis *et al.*, Phys. Lett. **B224**, 209 (1989); G. Cella *et al.*, Phys. Lett. **B248**, 181 (1990); A.J. Buras *et al.*, Max Planck Institute preprint MPI-Ph/93-77.
8. J.L. Hewett, SLAC-PUB-6521 (May 1994).

Table 1: Yields for the two analysis procedures.

	Neural Net	B Reconstruction
On	3013 $\pm$ 59	281 $\pm$ 17
Off(scaled)	2618 $\pm$ 73	155 $\pm$ 18
$\Upsilon(4S)$ background		
$b \to c$	50.7 $\pm$ 5.1	12 $\pm$ 2
$b \to u$	11.9 $\pm$ 4.0	2 $\pm$ 1
π^0 correction	50.2 $\pm$ 27.7	-0.7 ± 2.3
η correction	16.5 $\pm$ 33.7	2.0 $\pm$ 8.5
Non-$B\bar{B}$	2.3	
$\Upsilon(4S)$ total	132 $\pm$ 44	15 $\pm$ 9
On$-$Off$-\Upsilon(4S)$	263 $\pm$ 94	110 $\pm$ 25
	$\pm$ 49	$\pm$ 10

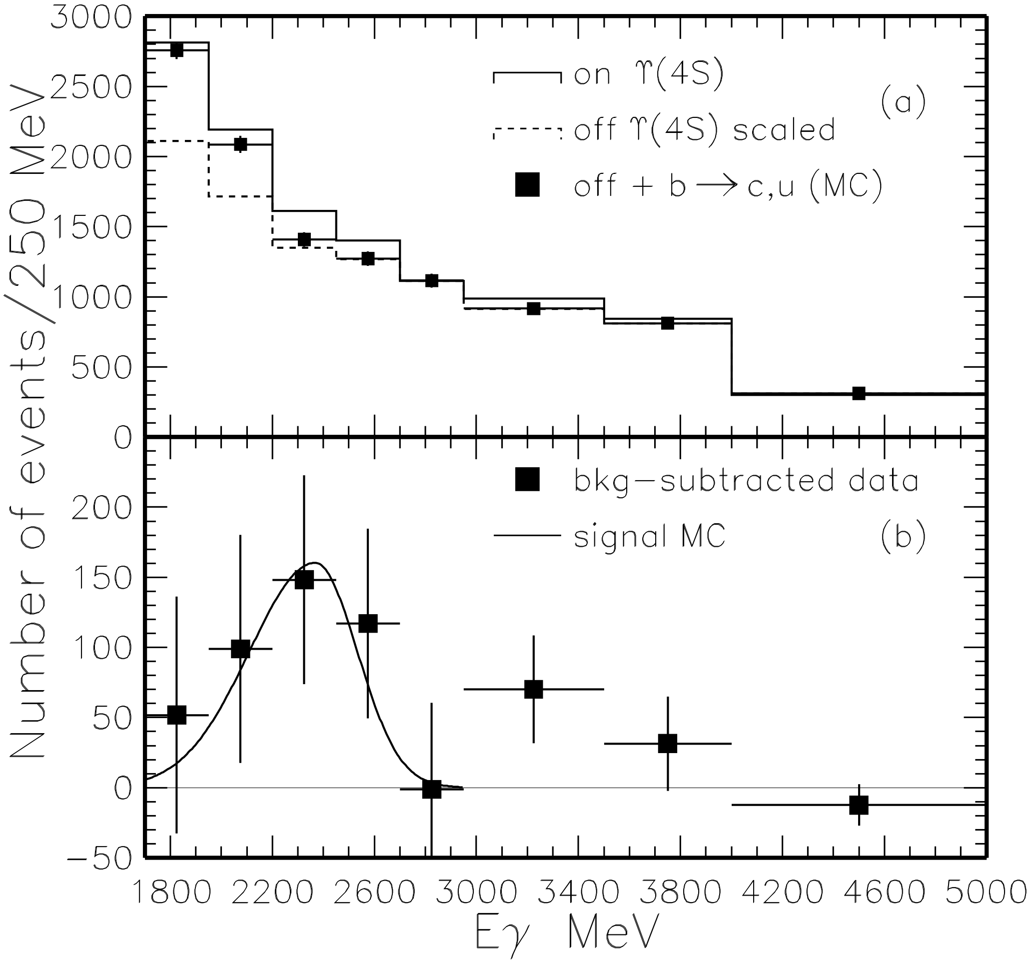

Fig. 1. Photon energy spectra from the neural net analysis.

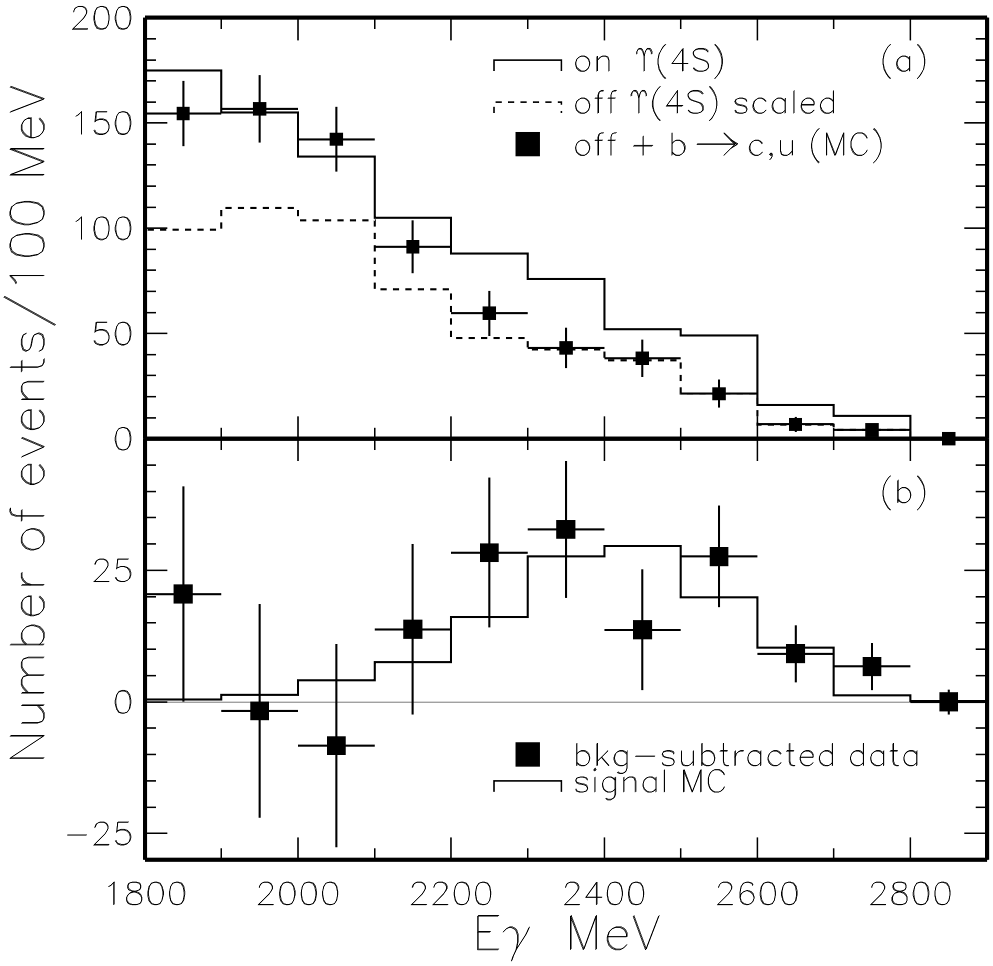

Fig. 2. Photon energy spectra from the B reconstruction analysis.

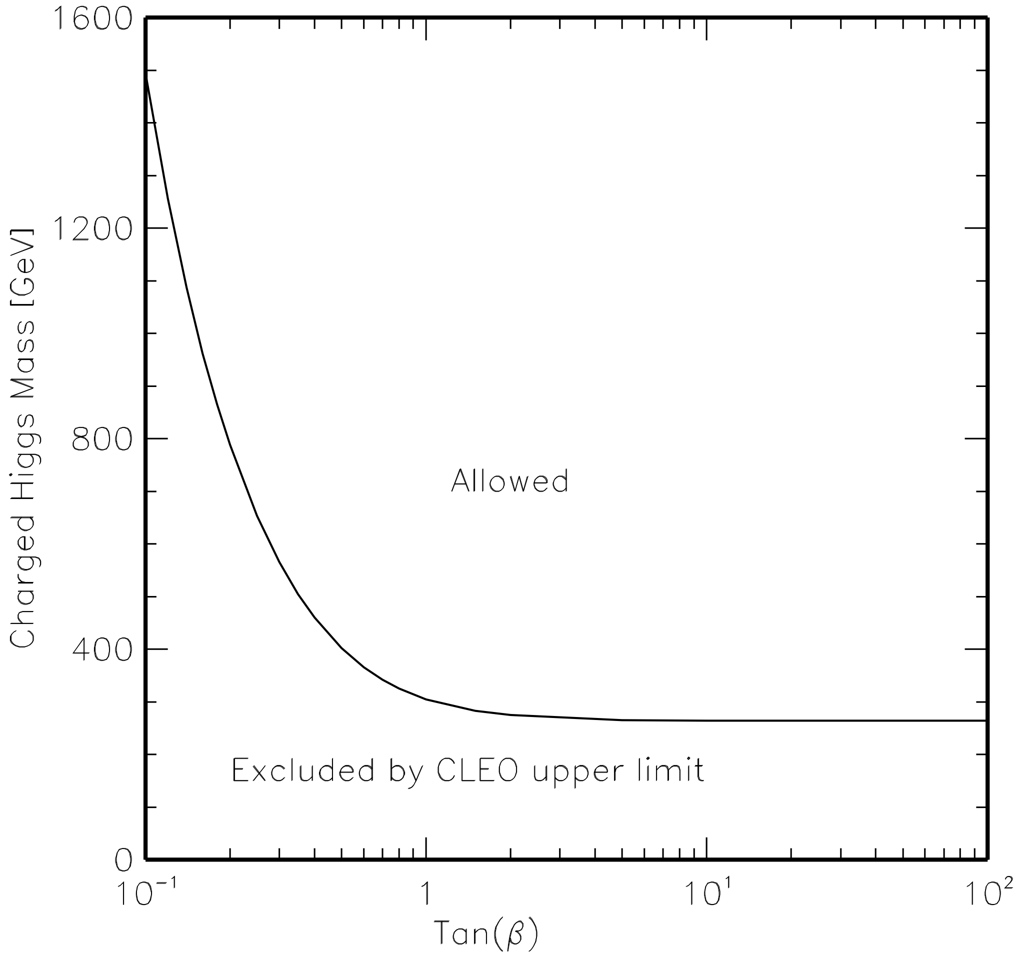

Fig. 3. Lower limit on Two-Higgs-Doublet-Model-II Higgs mass as a function of $\tan(\beta)$. The region below the curve is ruled out by the upper limit on the $b \to s\gamma$ branching ratio reported here.