CDF/ANAL/EXOTIC/CDFR/3572
Version 2.0
May 17th, 1996

Search for New Physics in Events with a Photon, b-tag, and $\not\!\!E_t$

Jeff Berryhill, Ray Culbertson, Henry Frisch, Dave Toback, and Peter Wilson
University of Chicago

Abstract

We search for anomalies in events with a photon, SECVTX b-tag and $\not\!\!E_t$. In events with two or more jets we find a peak in the dijet mass distribution. The peak consists of about 18-20 events over background around a mass of 105 GeV.

The signal is robust in that it survives higher threshold photon E_t and $\not\!\!E_t$ cuts. The signal is not robust in that the peak is prominent when we use jet cones of 1.0 but weak when we form the mass with cones of 0.7 or 0.4. Using JTC96S corrections instead of JTC90S broadens the peak significantly.

We try to determine if the second, non-btagged jet is a $\bar{b}$, and the evidence is inconclusive.

Many of the unanswered questions are listed in the conclusions section.

1 Introduction

In this note we describe a search for new physics in events which contain a photon, SECVTX b-tag and $\not\!\!E_t$. We begin with the data set and tagging results and show the dijet mass distributions for cone 0.4 and cone 1.0. There appears to be a peak when we use a cone of 1.0. In the following sections we describe our investigation into why there is a peak and why it might prefer the cone of 1.0 jets. Specifically try to answer the following questions:

- Why is the peak there when we use a cone of 1.0 and not when we use a cone of 0.4?
- Is the second jet from a b-quark?
- Is the peak 'sculpted' by the $\not\!\!E_t$ cut?

- Is the peak 'sculpted' by a trigger bias?
- Is there a bias because there is a b-jet in the event?
- Do known backgrounds give peaks in that region?
- Is the photon real and prompt?

This analysis is in a early stage and many of the answers to these questions are inconclusive. There are many more important questions to be answered. This note serves to document the work that has been completed to date.

2 Event Selection

We use the W_NOTRK trigger to select this sample. At Level 3 the trigger requires $\not\!\!E_t >$ 25 GeV and a central EM cluster with $E_t > 25$ GeV. There is no Level 2 prerequisite but it is fed mainly by photon triggers. Events that pass the trigger are written to the ECMB tapes. We have only processed the 1B data (90pb^{-1}).

We have filtered the ECMB tapes using the loose photon ID and isolation cuts described in Appendix A. The filter requirements are one central photon with $E_t > 25$ GeV, $\not\!\!E_t >$ 25 GeV, and out-of-time $E_t = 0$, to suppress the large number of cosmics. (Measured in W events, this cut is over 99% efficient, measured in photon events with no $\not\!\!E_t$, it is 95% efficient.) The $\not\!\!E_t$ is always uncorrected unless specified otherwise. The result is a sample of 18K events of $\gamma+$ $\not\!\!E_t$. Table 1 summarizes the filtering.

Cumulative cuts	Events
$E_t(\gamma) > 25$, $\not\!\!E_t > 25$	36140
$E_t(out) = 0$	18103
SECVTX tag	134
no tagged photon	133
two jets ($E_t > 10$,$\lvert\eta_{det}\rvert < 2.0$)	79
two SECVTX tags	1

Table 1: Summary of the filtering of the 1B ECMB tapes.

We recreated the SVX tracks with DO_SVX_PAD with Alignment Run 51110. The b jets are identified using the SECVTX algorithm. When a background spectrum is shown, it is the SECVTX Method 1 background developed for the top analysis. (We used C$TOP:FAKE_SUMET_QJTB_OPT3.DBT, which has been superceded recently.)

In the data sample, we find 135 btags in 134 events. The Method 1 background prediction is 112.6 tags. We find one event where the photon jet is tagged. If we remove the photon jet tag, and the corresponding background for all 18K photon jets, we find 134 tags in 133 events with a Method 1 background of 107.1 tags. We remove the photon tags and the corresponding background for the following unless otherwise noted.

We count jets with a cone radius of $R = 0.4$, uncorrected $E_t > 10$ GeV, and $\lvert\eta_{det}\rvert < 2.0$, after excluding the jet matched to the photon. If we require two or more jets, we find 80 tags in 79 events with a predicted background of 63.3 tags.

To study the effect of the $\not{E}_t$ cut we also have stripped the 1B GHIB tapes. These tapes contain isolated photon events with $E_t > 23$GeV and non-isolated photons with $E_t > 50$GeV. We took the events with a photon passing loose cuts with $E_t > 25$GeV. From this subset, we have stripped out all events with a SECVTX tag. As a control sample, we also stripped 40K events (about 5.5pb^{-1}) from the 1B GHIB tapes without the tagging requirement.

3 Results Using a Jet Cone of 0.4

Figure 1 shows the N_{jet}, photon E_t, and missing E_t distributions. The overlayed dashed histograms are the result from the Method 1 background prediction. We see most of the tagging excess in the one and three jet events, although the statistics are marginal.

Figure 2 shows six mass plots. The left three are the distributions in invariant mass of the dijet system consisting of the btagged jet and the highest E_t non-tagged jet. In this note we will always histogram dijet masses in 10GeV bins between 0 and 300GeV. The plots on the right are of the invariant mass of the three-particle system with the photon added in, *i.e.* the invariant mass of the btagged jet, the highest E_t non-tagged jet, and the photon. The jets are measured with a cone radius of 0.4. All jets are corrected with JTC90S for energy scale and out-of-cone energy ("YY,DD"). Almost all results in this note are with JTC90S, simply because that's what we started with. To see the effect of using JTC96S corrections, skip to figure 13.

The upper row is for all $\gamma \not{E}_t + \geq 2$ jet events. Since we are plotting the mass of the b and the next leading jet, the mass in events without a btag is ambiguous. In this case we choose the jet with the highest SECVTX fake-tag probability as the bjet. The "bjet" is corrected using AA corrections (including SLT corrections). When we use AA corrections this includes constraining the bjet to have a mass of 5 GeV. When we do not use AA corrections, we do not constrain the b mass. The other leading non-tagged jet is never corrected with AA corrections.

For the second row of histograms, we require a SECVTX tag and the bjet is corrected with AA corrections. For the third row of histograms, the bjet is not corrected with AA corrections and the jet mass is not constrained. The overlayed dashed histograms are the result from the Method 1 background prediction.

Figure 3 shows the dijet mass for cone 0.4 with AA corrections but with the expected tagging background subtracted. There is a possible excess near 115 GeV. A Gaussian fit over a flat background finds only the 6-event bin centered at 115 GeV to be in the peak. The double-tagged event has a dijet mass of 115 GeV. Figure 4 shows the dijet mass plotted *vs* the γ-dijet mass for cone 0.4 with AA corrections.

Appendix B shows some basic checks we performed on the tagged sample.

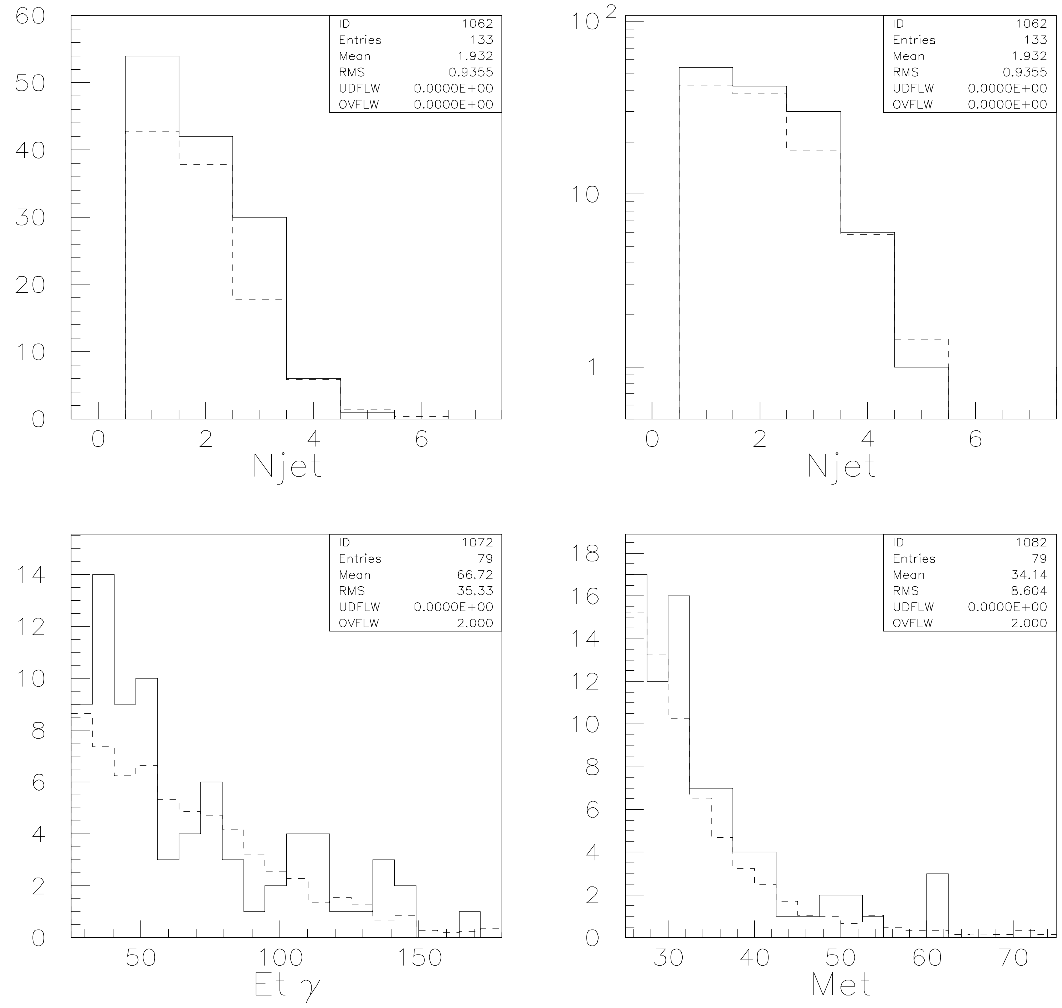


Figure 1: The N_{jet} (on linear and log scales), photon E_t, and missing E_t distributions. The overlayed dashed histograms are the result from the Method 1 background prediction.

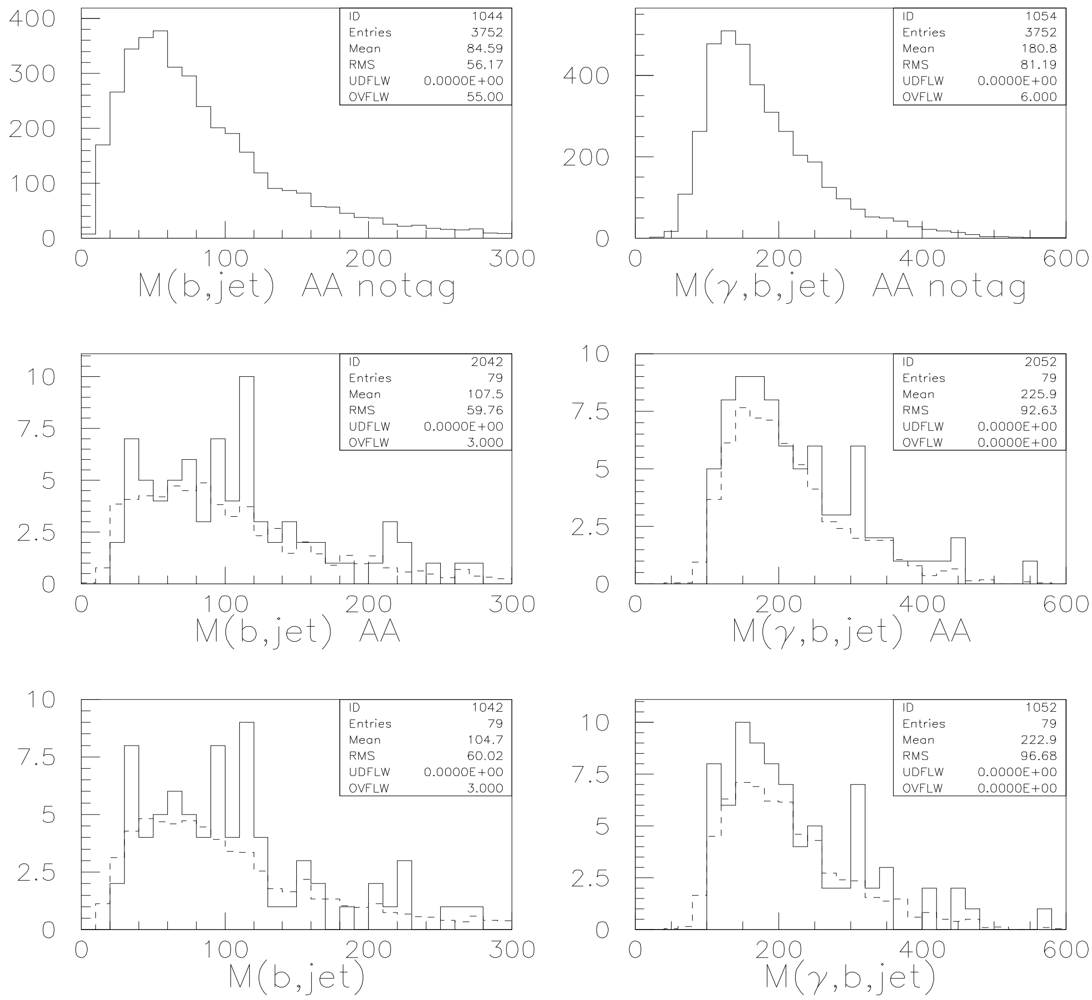


Figure 2: The mass of the b and leading jet (left) and the b, leading jet, and γ (right). The first row is for all events with γ, $\not\!\!E_t$, and two or more cone 0.4 jets. All jets are corrected with JTC90S. For the middle row we require a SECVTX tag and the b jet is forced to have a mass of 5 Gev and is corrected with AA corrections. For the bottom row the AA corrections are not used and the b jet mass is not constrained. Jet energies are measured in a cone of 0.4.

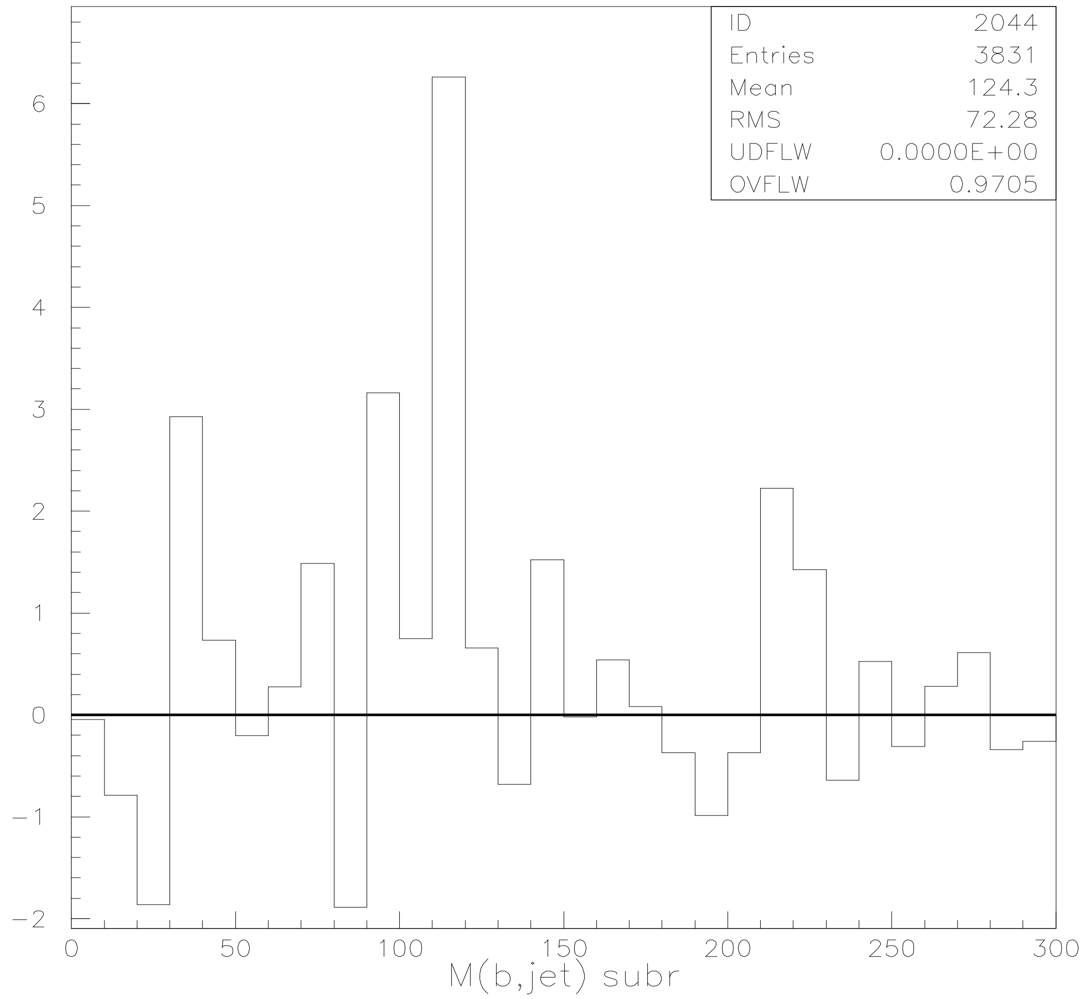


Figure 3: The mass of the b and leading jet with the expected tagging background subtracted. Jets with a cone of 0.4 with AA corrections are used.

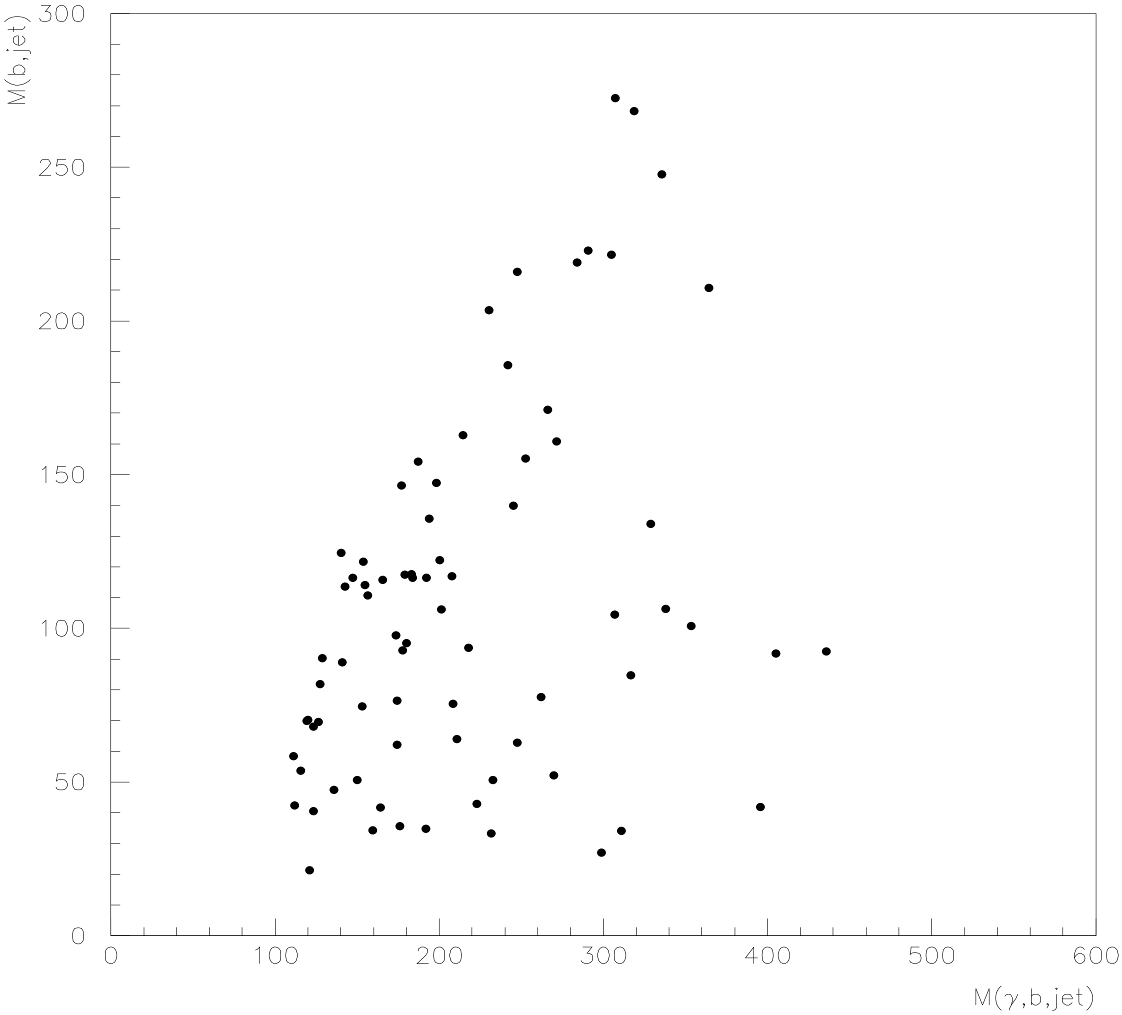


Figure 4: The mass of the b and leading jet plotted *vs* the b, leading jet, and γ mass for tagged events with two or more jets. Jet energies are measured in a cone of 0.4 and AA corrections are applied.

4 Results Using a Jet Cone of 1.0

Figure 5 shows the untagged and tagged dijet and γ-dijet masses using jets found with a cone of 1.0. The jet tagging and counting are still done with cones of radius 0.4 so these are the same events as in figure 2. The cone 1.0 jet matched to the photon is removed. The cone 1.0 jet matched to the tag is used as the bjet and the highest E_t remaining cone 1.0 jet is used for the second jet. If there is only one cone 1.0 jet left, the mass of the jet is plotted. (This occurs in 15 events.) There are no AA corrections on the cone of 1.0 masses. There is an excess near 105 GeV which does not seem to correspond to specific γ+dijet mass (See figure 7). In the three bins where the excess appears, there are 30 events. The Method 1 background for these bins is 12.5 events; fitting the sidebands by eyeball, there are about 10.5 background events.

Figure 6 shows the dijet mass with the expected tagging background subtracted. A Gaussian fit over a flat background gives a peak of 17.5 events centered at 103 ± 1 GeV with a width of 7.2 ± 0.8 GeV. This is less than the mass resolution for a dijet object which is expected to be in the range 11-15 GeV. The errors from the fit are underestimated since they are smaller than the ideal errors for a Gaussian with 17.5 events and a known width of 7.2 which would be $\sigma/\sqrt{N} = 1.7$ GeV on the mean and $\sigma/\sqrt{2N} = 1.2$ Gev on the width.

Figure 7 shows the scatter plot of dijet *vs* γ-dijet mass using cones of 1.0. The events where only one cone 1.0 jet was found are indicated with the open circles. Figure 8 shows the scatter plot and the dijet mass projection broken down into the number of cone 0.4 jets in the event. The figure indicates that the signal does not prefer a particular number of jets.

Attempting to isolate the excess around 105 GeV, we plot the cone 1.0 dijet mass *vs* several variables in figure 9. The b and second jet tend to have roughly equal energies in the signal region events. The signal events also may prefer a more energetic bjet and possibly the second jet to be not back-to-back with the bjet. However, both these variables are related to the mass.

Figure 10 shows the variables with the most correlation to the mass peak, E_t of the photon and $\not\!\!E_t$. The signal survives higher cuts on both these variables.

Figure 11 is more information about the signal and $\not\!\!E_t$. The $\not\!\!E_t$ significance is calculated as $\not\!\!E_t/\sqrt{\Sigma E_t}$. The corrected $\not\!\!E_t$ is calculated by taking each jet, subtracting out its raw energy from the $\not\!\!E_t$ and then adding back in the corrected jet energy.

Some signal events have a large $\not\!\!E_t$ significance; a cut > 3 would keep 5 signal region events and almost no others. The signal stands out in the corrected $\not\!\!E_t$ distribution about as well as the raw $\not\!\!E_t$ distribution. The third plot indicates how much the the raw $\not\!\!E_t$ is opposite the photon. The fourth shows that corrected $\not\!\!E_t$ is less directly opposite to the photon.

Figure 12 displays more information about corrected $\not\!\!E_t$. The first plot, for reference, is the GHIB sample with $\not\!\!E_t > 25$. The second plot is also the GHIB sample, but now with corrected $\not\!\!E_t > 25$. The shaded regions are events that pass both $\not\!\!E_t$ cuts. When going from cutting on $\not\!\!E_t$ to cutting on corrected $\not\!\!E_t$, the yield goes down by 25%. Despite the relatively small (20%) overlap in the event samples, both samples prefer the 100 GeV mass region. The third plot in figure 12 shows that in the sample with correct $\not\!\!E_t$, the corrected $\not\!\!E_t$ does not tend to point opposite the photon.

Figure 13 shows the effect of switching to the JTC96S correction. These were run with out-of-cone corrections and using the event-by-event class 12 vertex count to determine the underlying event correction. The cone 0.4 appears a little more peaky, but not significantly. The cone 1.0 peak has been smeared. The width with '96 corrections is much more consistent with the expected resolution for a dijet object.

Appendix B shows some basic checks we performed on the tagged sample.

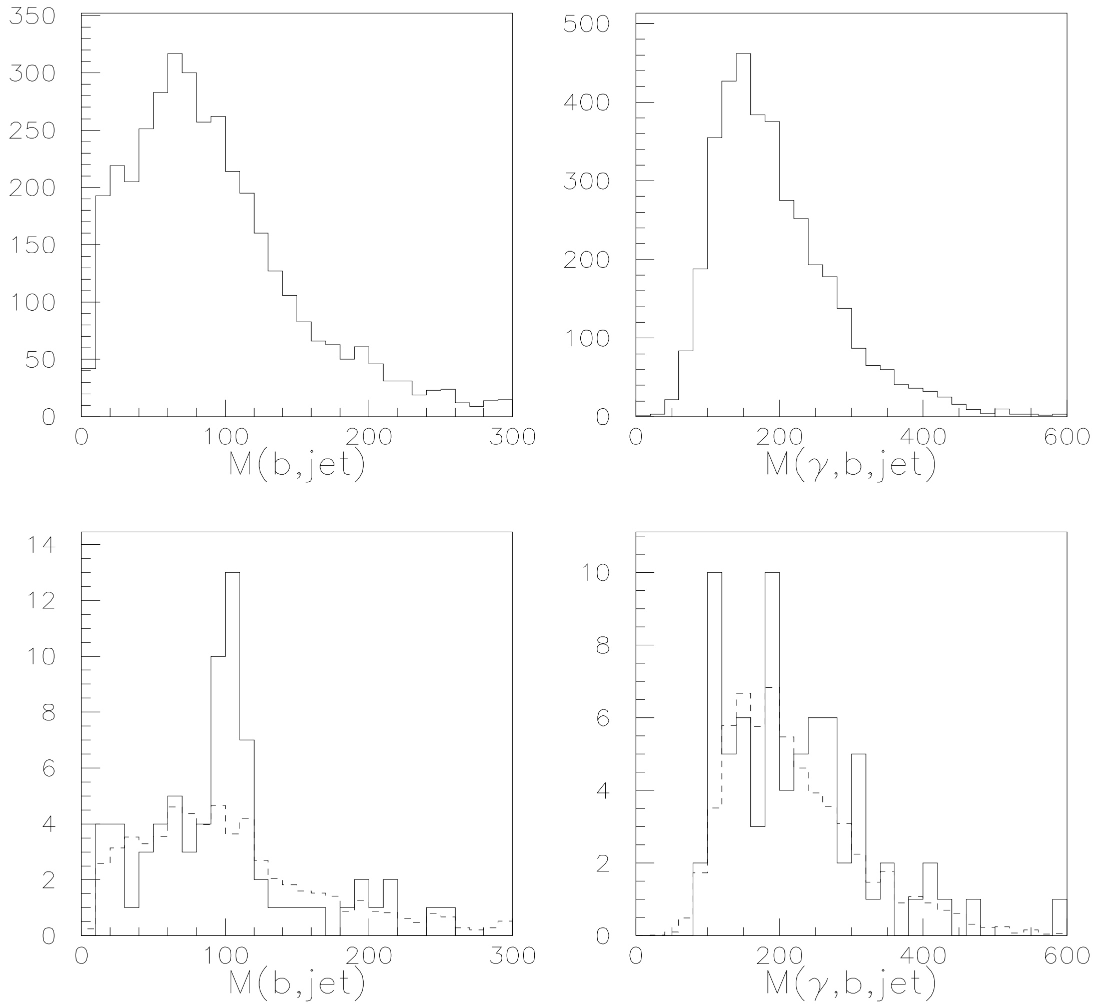


Figure 5: The mass of the b and leading jet (left) and the b, leading jet, and γ (right). The upper histograms are for all γ $\not{E}_t$ events; the lower are for tagged events with 2 or more jets found with a cone radius of $R = 0.4$. Jet energies measured with a cone radius of 1.0 are used to calculate the masses. The jets are corrected with JTC90S only (AA not available).

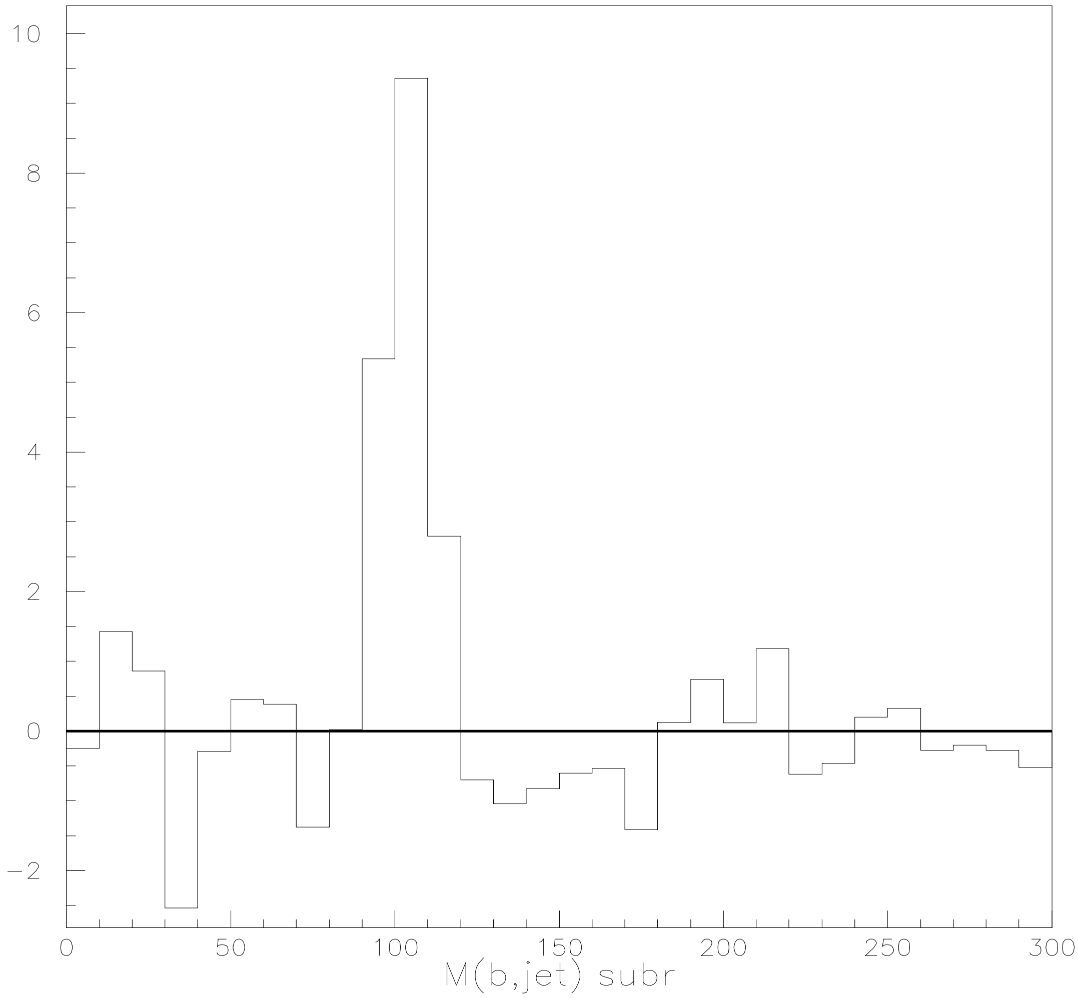


Figure 6: The mass of the b and leading jet with the expected Method 1 tagging background subtracted. Jets with a cone of 1.0 are used. The jets are corrected with JTC90S only (AA not available). A Gaussian fit over a flat background gives a peak of 17.5 events centered at 103 ± 1 GeV with a width of $7.2 \pm .8$ GeV.

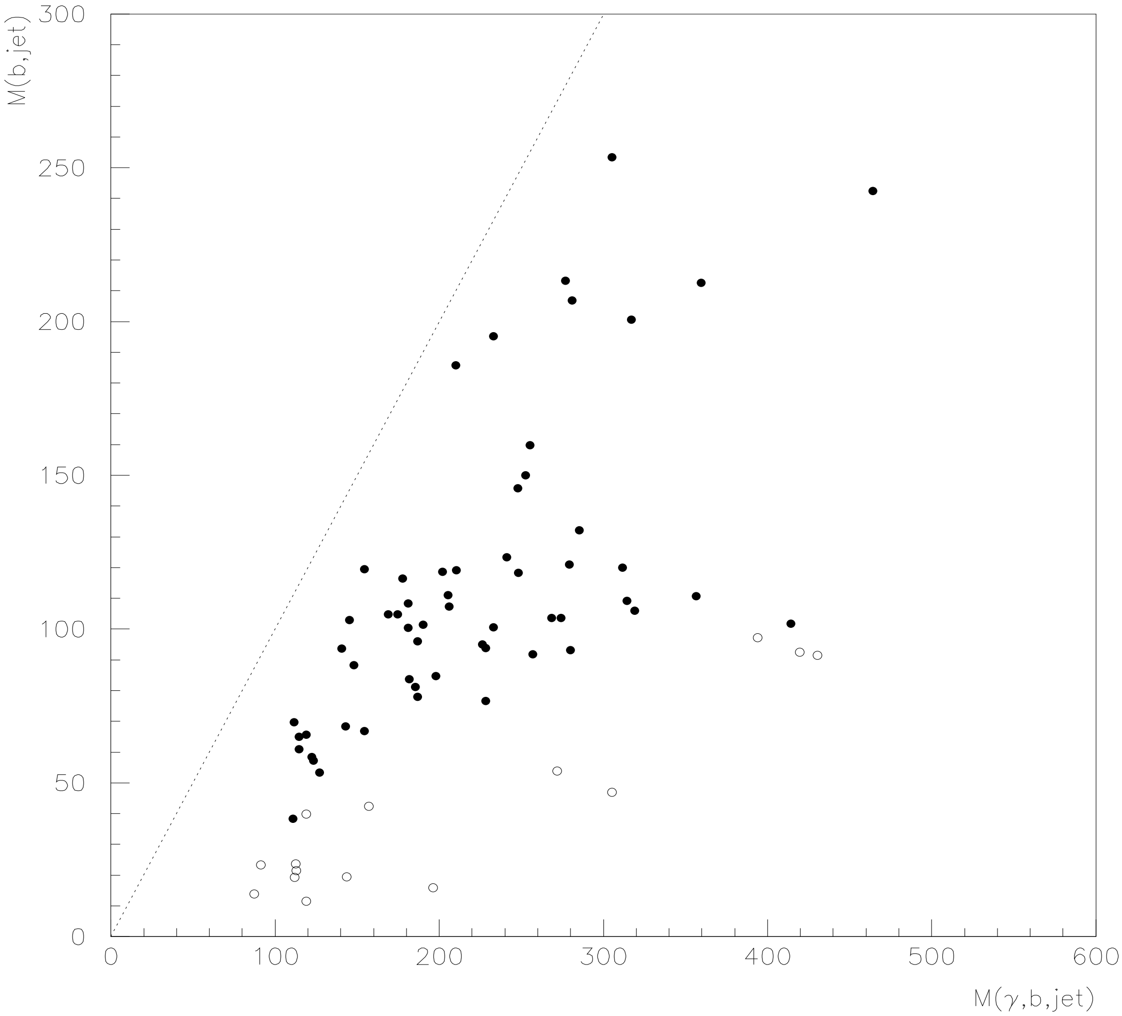


Figure 7: The mass of the b and leading jet plotted *vs* the b, leading jet, and γ mass for tagged events with two or more cone 0.4 jets. Jets measured with a cone of 1.0 are used for these masses. The events where only one cone 1.0 jet was found are indicated with the open circles. The jets are corrected with JTC90S only (AA not available). The dotted line is the diagonal.

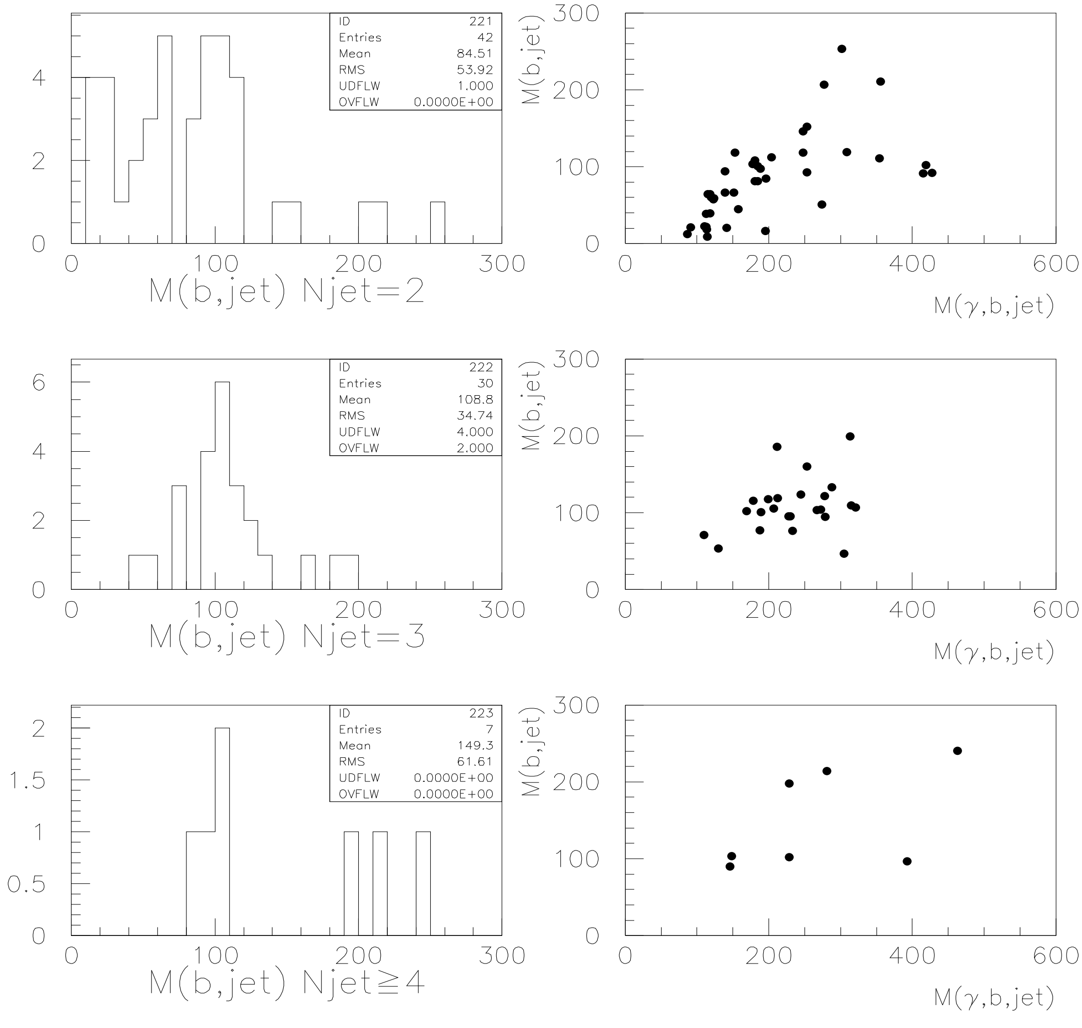


Figure 8: The scatter plots are the mass of the b and leading jet plotted *vs* the b, leading jet, and γ mass for tagged events with two or more cone 0.4 jets. Jets measured with a cone of 1.0 are used for these masses. The histograms are the projection of the mass of the b and leading jet. The first row is for events with only two cone 0.4 jets, the center for events with three cone 0.4 jets and the bottom for events with four or more.

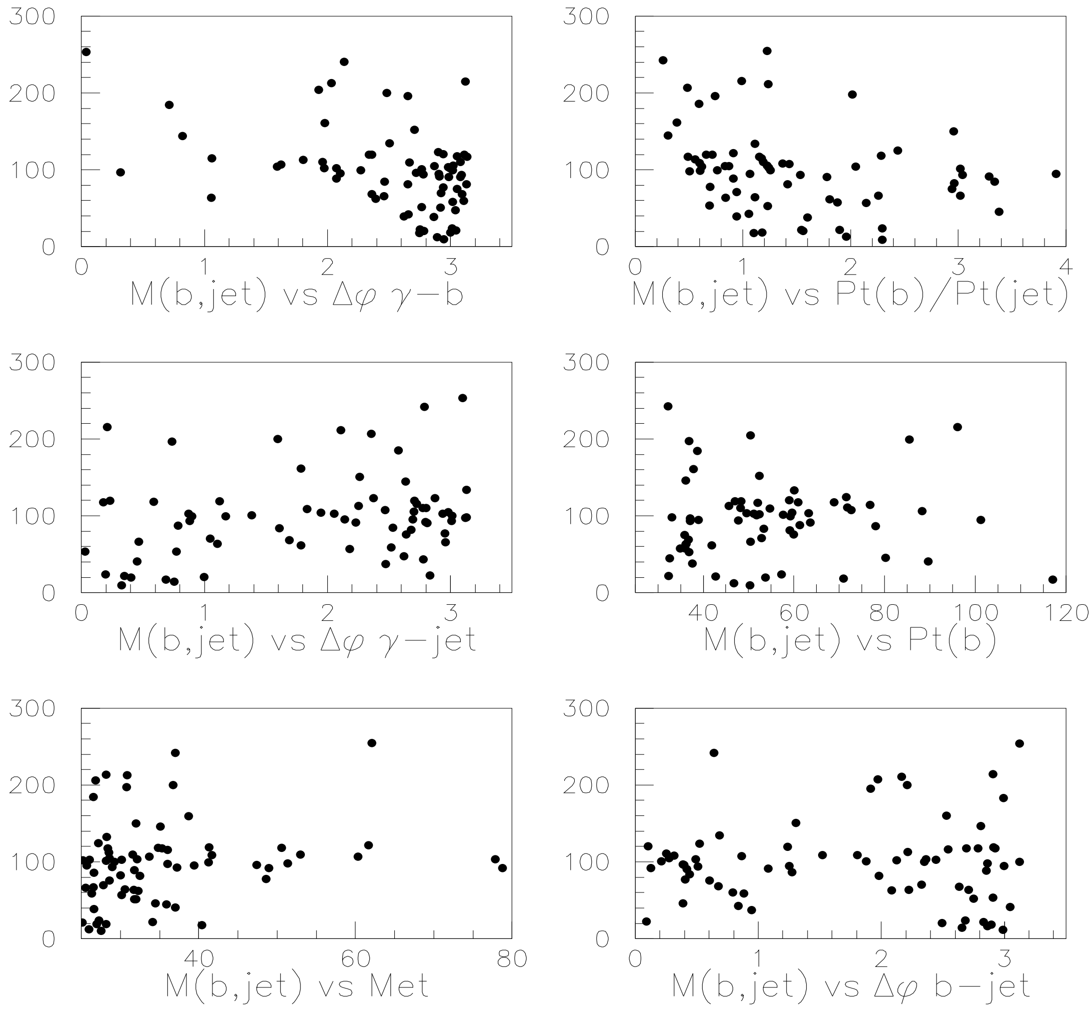


Figure 9: The $b+jet$ mass plotted *vs* several variables: $\Delta\phi(\gamma-b)$, $P_t(b)/P_t(jet)$, $\Delta\phi(\gamma-jet)$, $P_t(b)$, $\not{E}_t$, and $\Delta\phi(b-jet)$. The jet energies for the dijet mass are measured in a cone of 1.0.

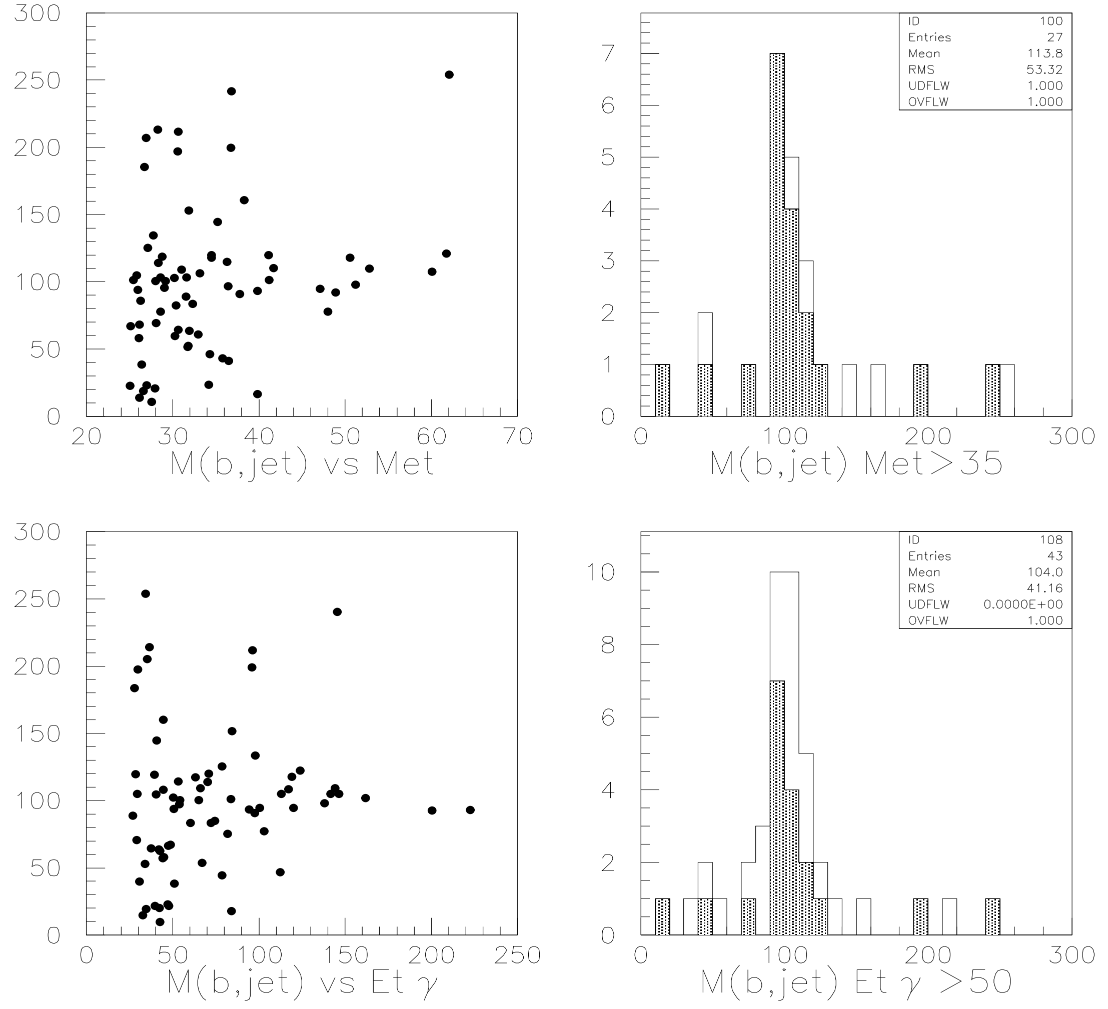


Figure 10: The cone 1.0 dijet mass plotted *vs* $\not\!\!E_t$; dijet mass for $\not\!\!E_t > 35$; the dijet mass plotted *vs* E_t of the γ; the dijet mass for photon $E_t > 50$. The shaded region is events that pass both $\not\!\!E_t > 35$ and $E_t > 50$ cuts.

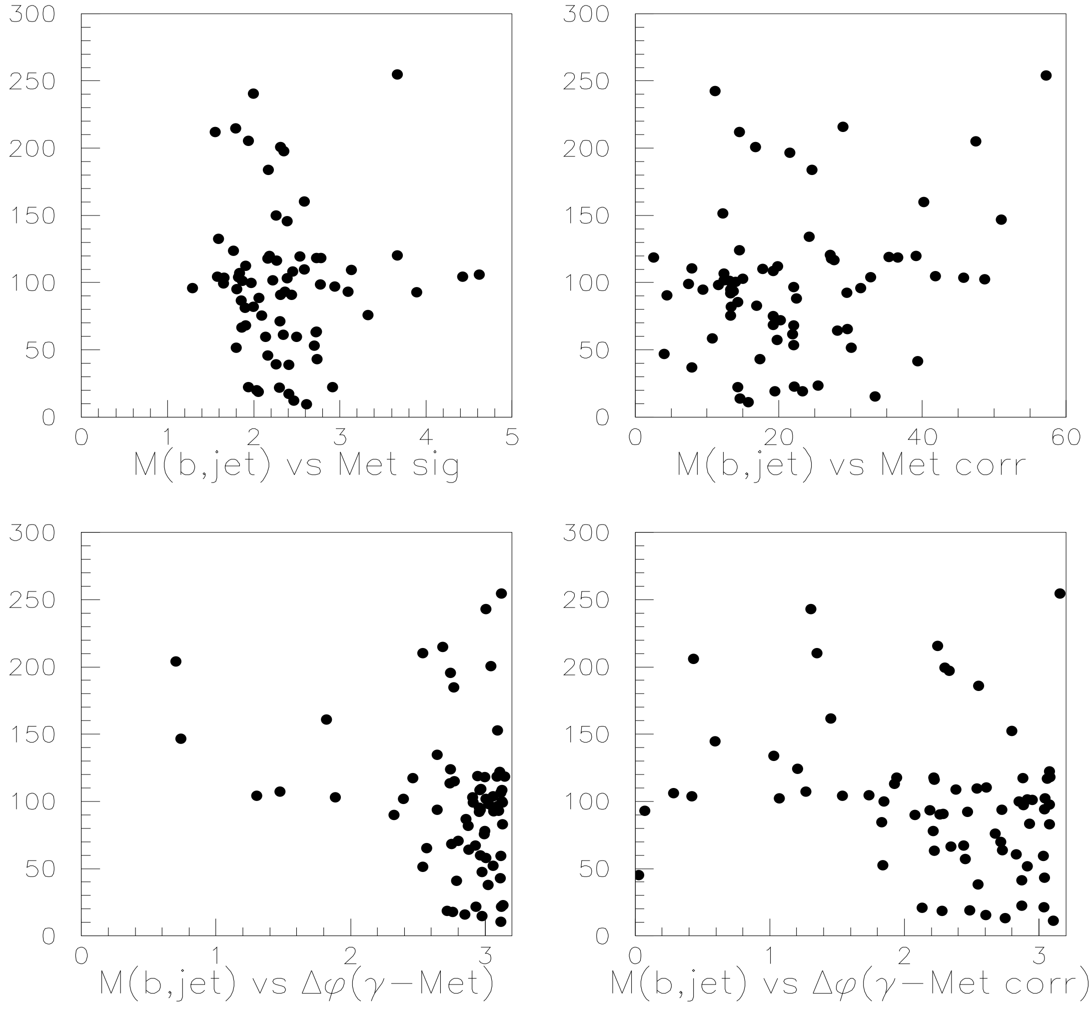


Figure 11: The upper left plot is the cone 1.0 dijet mass plotted *vs* the $\not{E}_t$ significance, upper right is the dijet mass plotted *vs* the $\not{E}_t$ corrected for jets. Lower left is the dijet mass plotted *vs* the angle between the photon and the $\not{E}_t$, lower right is the dijet mass plotted *vs* the angle between the photon and the $\not{E}_t$ corrected for jets.

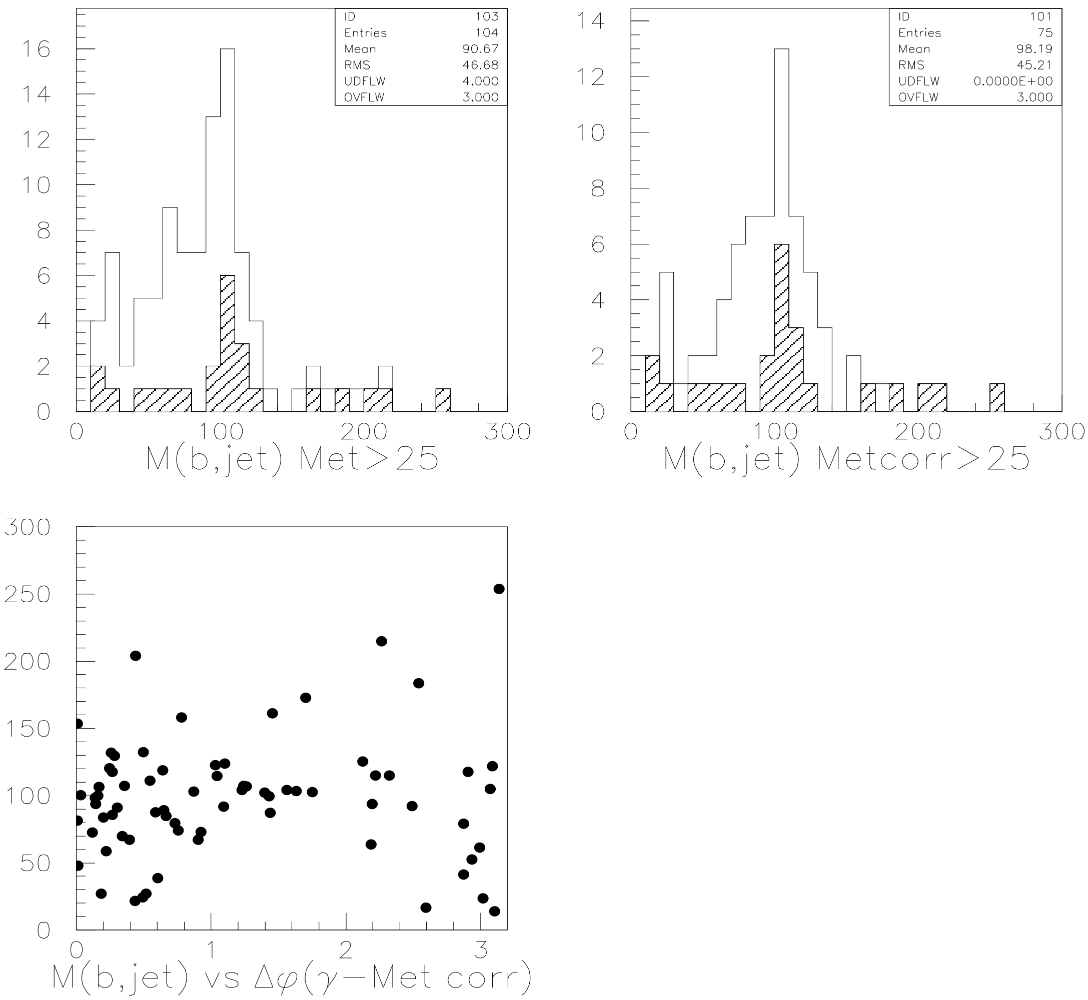


Figure 12: The upper left plot is the cone 1.0 dijet mass for events with raw $\not\!\!E_t > 25$, upper right is the dijet mass for events with corrected $\not\!\!E_t > 25$. Both these samples were taken from the inclusive photon (GHIB) stream, The shaded events in both plots are the events with both $\not\!\!E_t > 25$ and corrected $\not\!\!E_t > 25$, the overlap. Lower left is the dijet mass plotted *vs* the angle between the photon and the corrected $\not\!\!E_t$ for events with corrected $\not\!\!E_t > 25$.

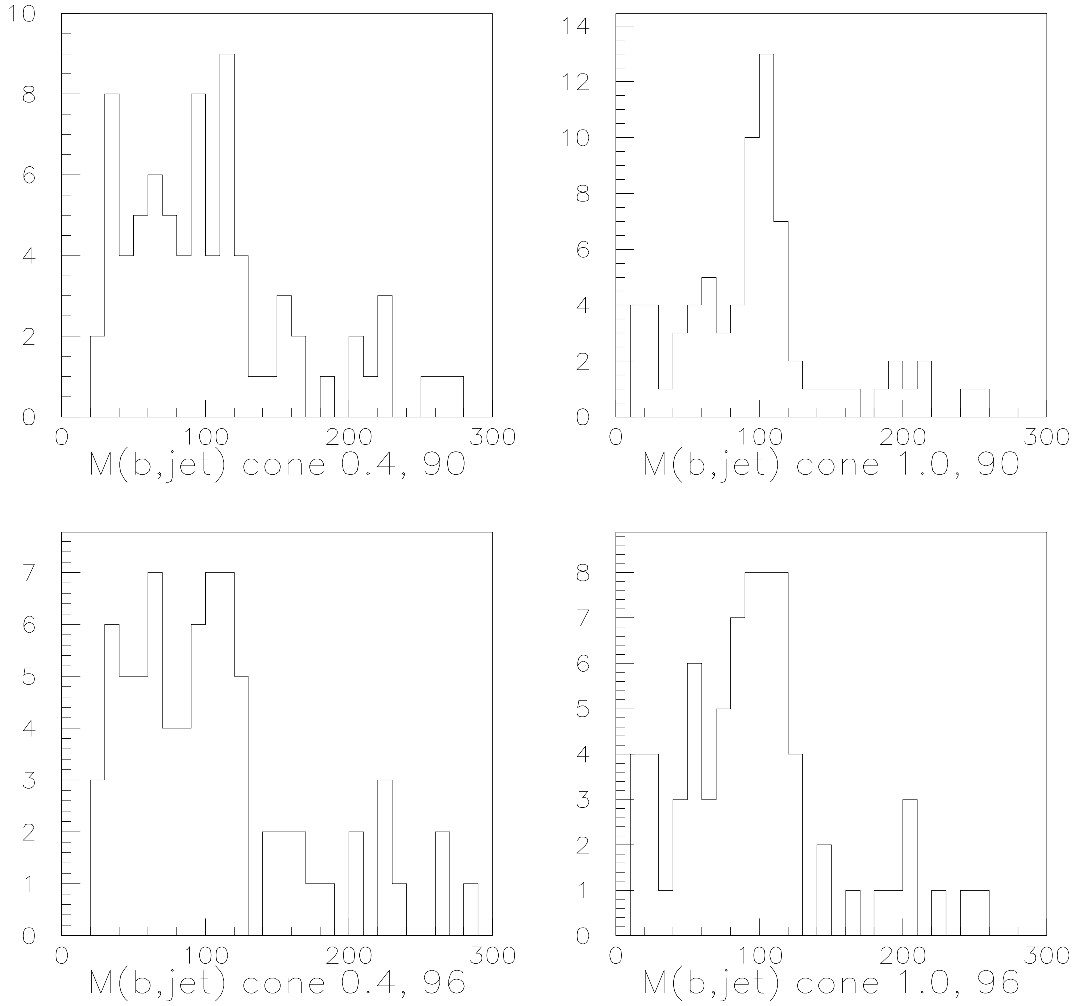


Figure 13: The dijet mass comparing JTC90S corrections and JTC96S “EY,DD” corrections. The top is the 90 corrections and bottom is 96. The left is cone 0.4 and the right is cone 1.0.

5 Why Is the Peak Only in a Cone of 1.0?

The excess is significantly more peaked when the cone 1.0 jets are used compared to the cone of 0.4 jets. This could be caused by the larger cones picking up legitimate stray energy from the jets, improving the resolution. Figure 14a shows the $Z \rightarrow jets$ Monte Carlo peak when cones of 1.0 or 0.4 are used. The FWHM of the two peaks are consistent with the same width.

Alternatively, the difference could be caused by a third jet in the system which is included in the mass when cone 1.0 is used. Figure 14b shows the cone 0.4 mass where the third leading jet with $E_t > 10$ is included, if there is one. There is a degradation in the peak excess. The last plot in the figure shows the cone 0.7 mass which does peak a little better than the cone 0.4 but not as well as the cone 1.0.

Recall that when we plot the cone 1.0 mass, in 15 events there is only one cone 1.0 to use in the "dijet" mass. In this case we plot the one jet's mass. The mass distribution of these events is shown in figure 15. This pathology does not seem to concern the mass peak. The same effect can be seen in the low-mass end of the Z Monte Carlo in figure 14.

Figure 15 shows the distribution of cone 0.4 dijet masses when the cone 1.0 dijet mass is in the peak window. In going from the cone 0.4 mass to the cone 1.0 mass, 5 events have a large mass promotion (increase in dijet mass), 5 have a moderate promotion, 5 have a moderate demotion, and 1 has a large demotion.

Table 4 lists the 16 events that have a cone 1.0 dijet mass within 15 GeV of 105 GeV but a cone 0.4 mass not in this window. Sometimes the cone 1.0 algorithm simply fails, due to the proximity of a jet to the photon. Other times the cone 1.0 collects more stray energy which it is designed to do. There is no clear pattern to what is happening in these events.

To see if there is something unusual about the events in the peak, we asked how many cone 0.4 jets are contained in the cone 1.0 jets, in and out of the signal region. Table 2 shows that the distribution of the number of number of cone 0.4 in each cone 1.0 is about the same inside and outside the peak.

N cone 0.4 in cone 1.0	Signal bjet (%)	Signal 2nd jet (%)	Non-signal bjet (%)	Non-signal 2nd jet (%)
0	0	26	0	13
1	74	67	72	78
2	19	7	25	9
3	7	0	3	0

Table 2: The percentage of events that have zero, one, two, and three cone 0.4 jets in each cone 1.0.

Another way to break down the events into categories is to look at the cone 0.4 and 1.0 matching. Figure 16 shows the cone 1.0 mass for several different criteria. In the first plot we simply remove the 15 events were there was only one cone 1.0 jet in the event. The remaining events are divided into classes. In all events the cone 0.4 btag was found matched to a cone 1.0 jet.

- Class 1. The second cone 0.4 jet does not match either of the cone 1.0 jets.

- Class 2. The second cone 0.4 jet is matched to the same cone 1.0 as the btag. The second cone 1.0 jet comes from somewhere else.

- Class 3. The second cone 0.4 is matched to the second cone 1.0 jet. (The jet selection agrees between the cone sizes.)

- Class 3.1. The same criteria as class 3 but also there are no extra cone 0.4 jets matched to the cone 1.0 jets. (The jet selection agrees and the only difference is in scattered energy.)

The mass distributions for these classes are shown in figure 16. Class 1 accounts for a few events on the high side of the signal region; class 2 accounts for some on the low side, but the center signal-region bin appears mostly in the preferred class 3. The class 3.1 requirement seems to be about 70-80% efficient over class 3, for signal and background.

The last histogram in the figure is the cone 0.4 dijet mass for the class 3 events. Even with the jet selection agreeing, the cone 0.4 refuses to peak. Maybe this isn't surprising since we are taking a one-bin peak and smearing it with some uncertainty.

Table 3 shows how the signal events, the events in the signal region, and the control samples break down into these classes. The control samples are the GHIB sample (with a tag, no $\not\!\!E_t$ cut), the ECMB sample (no tag, with a $\not\!\!E_t$ cut), and the GHIB sample (no tag, no $\not\!\!E_t$ cut). In the samples without a tagging requirement, we take the jet with the highest SECVTX Method 1 probability as the bjet. In all samples we found approximately 75% of the events had two cone 1.0 jets. For the percentages in the table, we use the sample with two cone 1.0 jets as the denominator.

Class	Signal sample	Signal Region	γ, tag	$\gamma, \not\!\!E_t$	γ
1	17	15	11	14	10
2	17	15	13	35	32
3	67	67	76	49	56
3.1	49	52	69	41	52

Table 3: The percentage of events falling in each jet-matching class. Class 3.1 is a subset of class 3, so only the first three rows should sum to 100%.

The first thing we notice is that the signal region is not distinguished from the rest of the events in the signal sample. The samples differing only by the $\not\!\!E_t$ cut seem to follow each other, indicating the $\not\!\!E_t$ is not altering the distribution. However, when a tagging requirement is added, class 3 is increased and class 2 is decreased. Basically this difference is: events without a tag tend to have the two leading jets near each other; events with a tag tend to have leading and second jets further apart. It is not clear what this means, it could just be the $E_t > 15$ and SVX fiducial cuts in the tagging or it may be a different production mechanisms for b's.

The fact that the peak is prominent with cones of 1.0 and not with cones of 0.4 is the most troubling aspect of the signal.

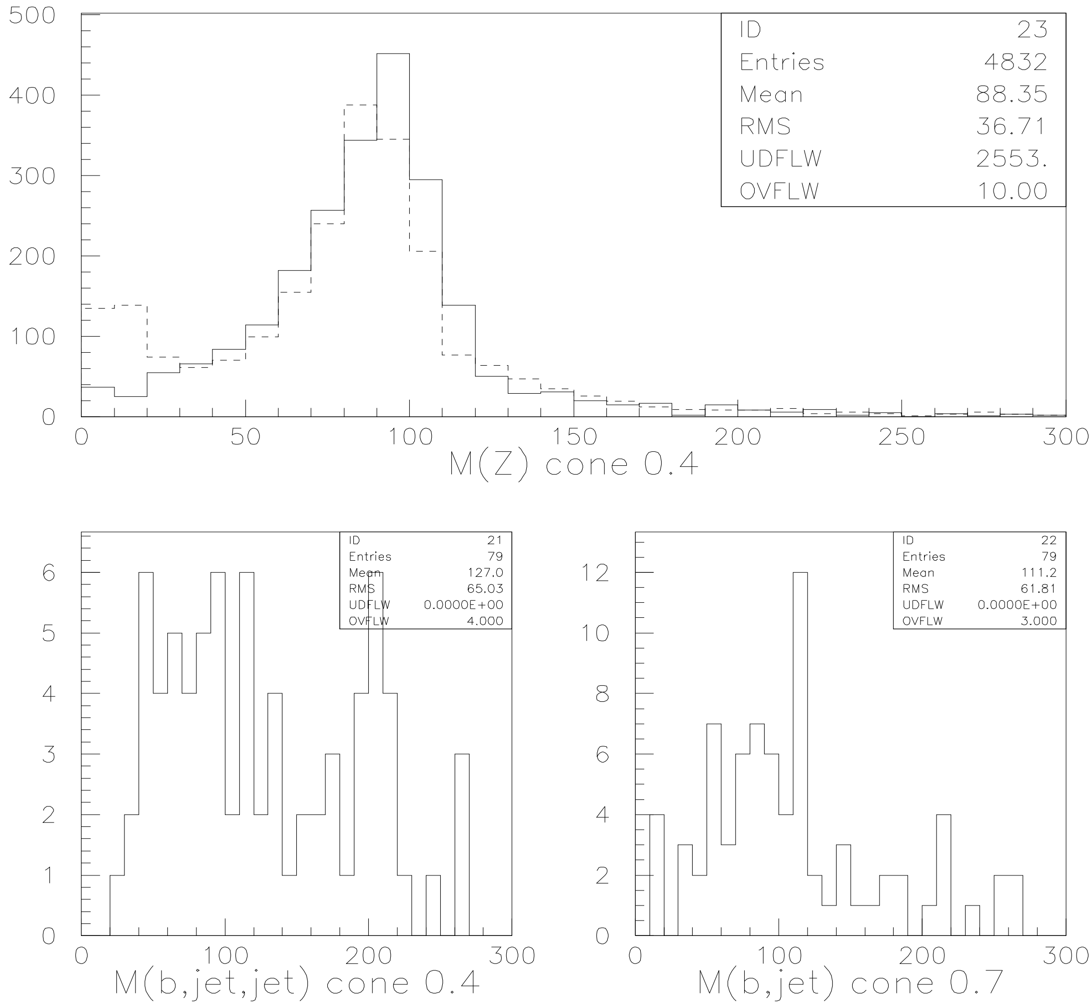


Figure 14: The top plot is a γZ with $Z \rightarrow jj$ Monte Carlo mass peak made with cone 0.4 jets (solid) and overlayed with the mass made with cone 1.0 jets (dashed). The bottom left is the cone 0.4 dijet mass with a third jet added in if one exists (AA corrections applied to the bjet). The bottom right is the cone 0.7 dijet mass distribution.

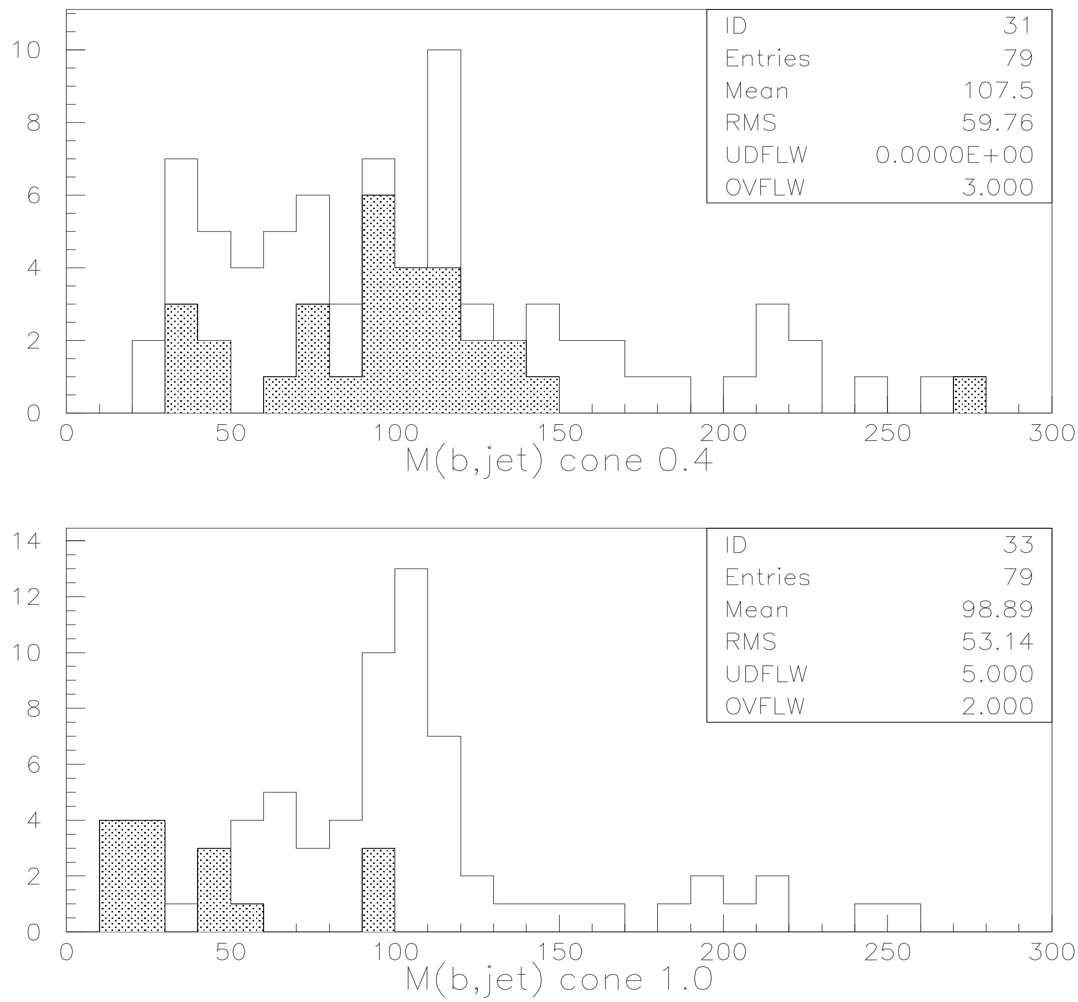


Figure 15: The upper histogram is the cone 0.4 dijet mass for events with 2 or more cone 0.4 jets. The shaded region is the events that have a cone 1.0 dijet mass in the window 105 ± 15 GeV. The lower histogram is the cone 1.0 mass for the same events. The shaded region is events where only one cone 1.0 jet contributes to the mass.

Run/Event	cone 0.4 mass	cone 1.0 mass	Comment	Winner
61442/13175	132	117	Cone 1.0 added little energy, but had significantly smaller correction	1.0
66500/493126	120	108	Cone 1.0 added little energy, but had significantly smaller correction	1.0
65004/427309	137	115	Cone 1.0 added little energy, but had significantly smaller correction	1.0
63078/45206	271	118	3 jets. The leading jet is in the same cone 1.0 as the photon, different jets selected	0.4
68592/184523	124	103	4 jets. The leading jet is in the same cone 1.0 as the photon, different jets selected	0.4
67349/33033	147	99	The leading jet is in the same cone 1.0 as the photon, different jets selected	0.4
64811/532300	87	94	Cone 1.0 picked up more scattered energy	1.0
63684/15166	78	104	Cone 1.0 picked up more scattered energy.	1.0
69998/81487	61	93	Cone 1.0 picked up more scattered energy	1.0
59857/902	42	90	Cone 1.0 clusters the two leading jets together and finds a third 13 GeV jet.	?
63262/17881	33	102	Cone 1.0 clusters the two leading jets together and finds a third 11 GeV jet.	?
69808/484976	39	93	Cone 1.0 clusters the two leading jets together and find a small third jet.	?
64997/52899	33	96	Cone 0.4 splits a fat jet	1.0
67845/12308	40	92	Cone 0.4 splits a fat jet	1.0
68393/32201	77	93	Cone 1.0 brings a small 3rd jet into the 2nd jet.	1.0
66046/566693	74	118	Cone 1.0 picked different 2nd leading jet because of adding in scattered energy.	0.4

Table 4: Events where the cone 1.0 mass is within 15 Gev of 105 GeV and the cone 0.4 mass is not. In going from cone 0.4 to cone 1.0, 10 events increased their dijet mass and 6 events decreased their mass. The “Winner” column indicates which cone size seemed to have a more appropriate interpretation for a dijet object.

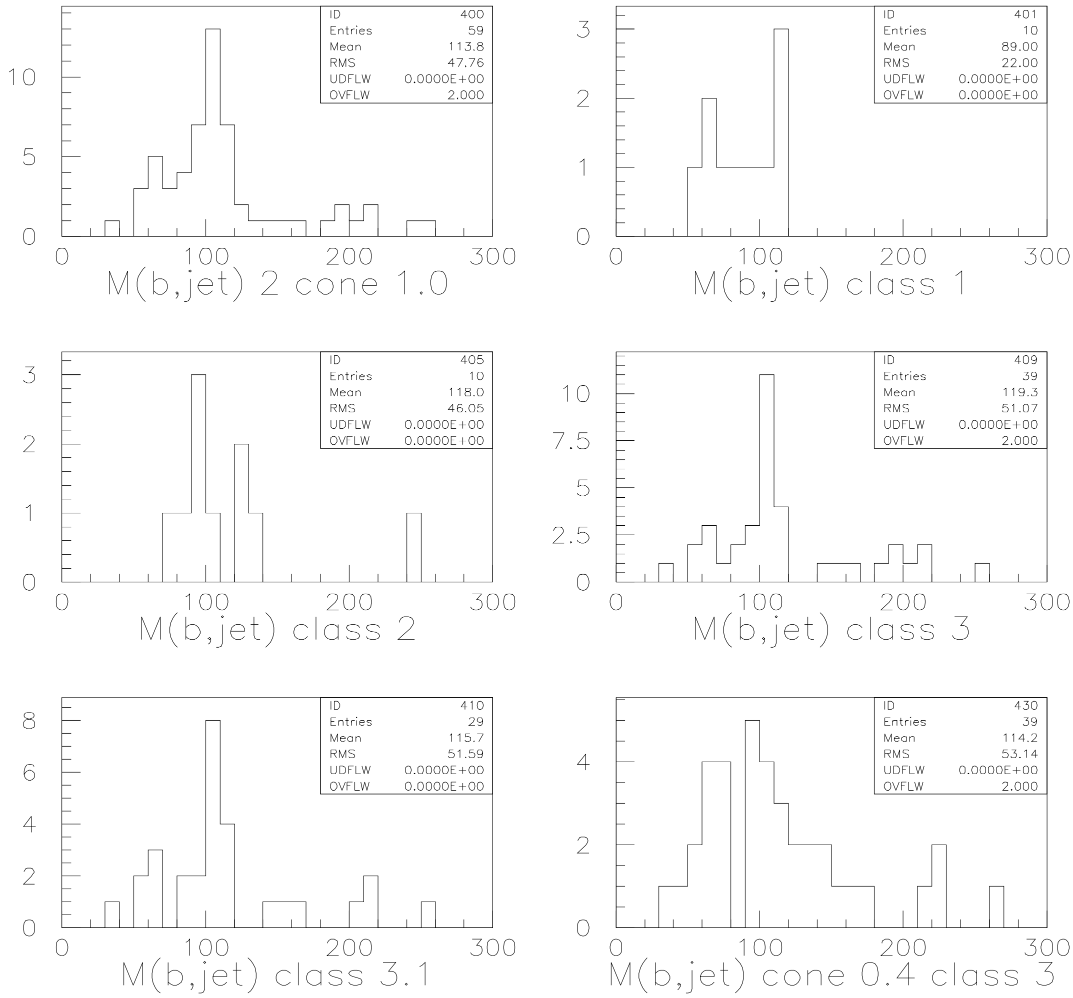


Figure 16: The first 5 plots are the cone 1.0 mass for several different requirements on the matching between cone 1.0 and cone 0.4 jets. The bottom right is a cone 0.4 dijet mass for class 3 events. The classes are described in the text.

6 Trigger Path Biases

We have investigated events which come in on separate triggers which should pass our final dataset cuts. Of the sample from the ECMB tapes (W_NOTRK), 75% pass the photon (PHOB_CEM3_23_ISOL*) trigger and therefore are also found on the GHIB inclusive isolated photon tapes. Looking at it the other way, only 57% of PHOB_CEM3_23_ISOL* events passing our loose photon cuts, with $\not{E}_t > 25$, also pass the W_NOTRK trigger.

Naively both W_NOTRK and the photon trigger with a $\not{E}_t$ cut should both accept these events. But the difference may not be unreasonable since the photon trigger requires neural net isolation while the W_NOTRK trigger doesn't have neural net isolation but does have more cluster quality cuts.

The trigger path differences are discussed more in Appendix C.

Figure 17 shows that the signal is found in both W_NOTRK and photon trigger paths.

A potential problem is that the GHIB sample is fed by a tight 23GeV trigger and a loose 50GeV trigger. This might cause a jump in the number of events in going from $E_t < 50$ to $E_t > 50$. However we find that approximately 90% of events in either trigger path pass a single low-E_t trigger. This does not rule out any effects, but it does indicate any problem would be limited. Both trigger paths have a significant decrease in efficiency at larger E_t. The trigger path differences are discussed more in Appendix C.

A small effect tending to push the E_t higher and another effect tending to push it down could contribute to some sculpting. However, if there were a serious problem, we would expect it to appear in the mass distributions of the untagged sample or the Method 1 background both of which show no sign of serious sculpting.

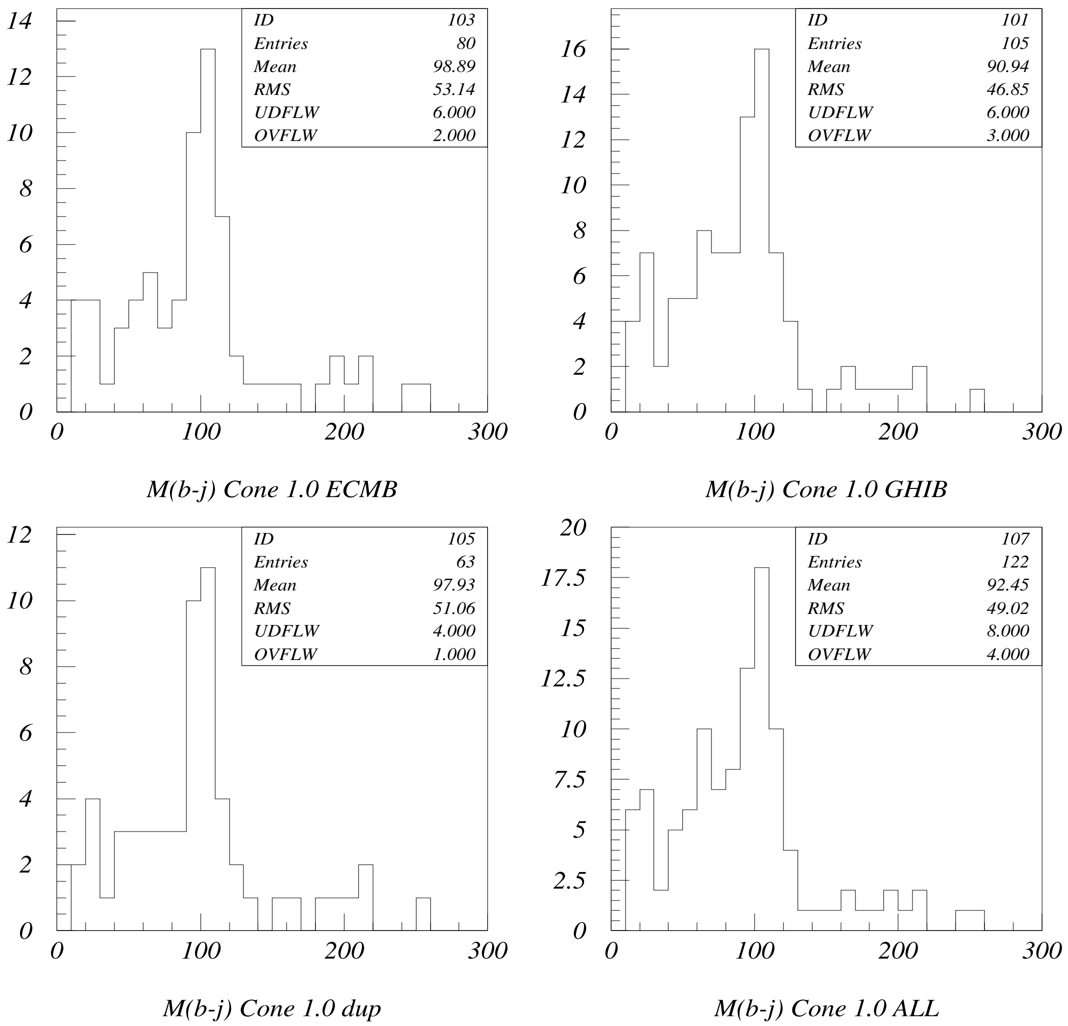


Figure 17: The cone of 1.0 dijet mass plot for the different triggers. Upper left, W_NOTRK (ECMB tapes); upper right, photon trigger (GHIB tapes); lower left the events coming in on both triggers; lower right, sum of the two trigger paths with duplicates removed.

7 The Second Jet

Since we see a small number of double SECVTX tagged events (1 event), we asked if this is consistent with a $b\bar{b}$ state. In a $Z\gamma$, $Z \to b\bar{b}$ Monte Carlo, we find 17% of events with a tag have a second tag. This does not include any scaling to account for a difference between data and Monte Carlo tagging efficiencies as in the top analysis. This implies that if the excess is on the order of 18-20 events, we would expect about 3.2 events to be double tagged.

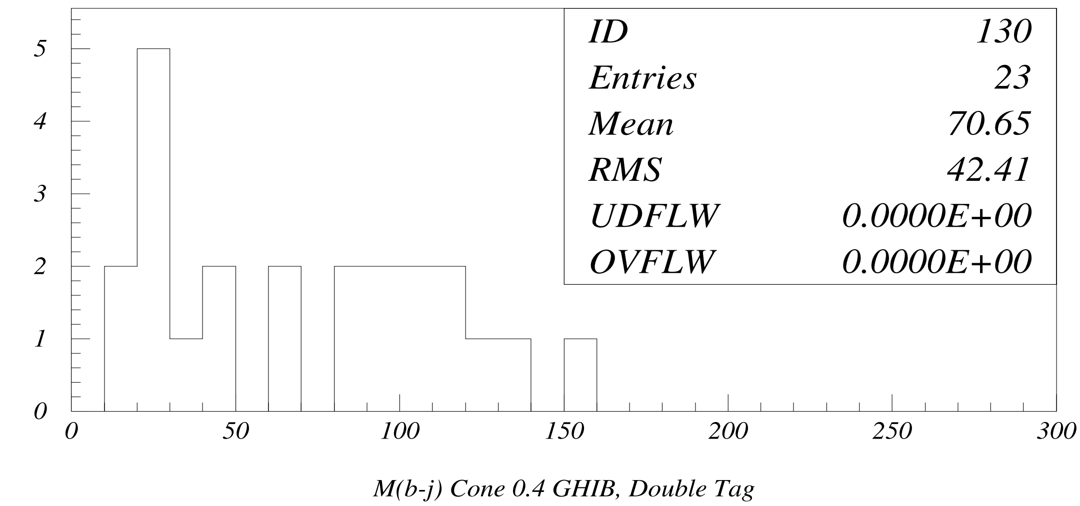
Another estimate we can make is based on the top tagging rate. In the 79-event sample there are 45 taggable jets chosen as the second (non-tagged) jet, a rate of 56%. In the 30-event signal region there are 17 taggable jets, about the same rate. If the excess is 18-20 events, we expect approximately 11 taggable second jets. The rate of tagging taggable jets when they are b's is 33%[1]. Multiplying this rate times the number of taggable jets, we expect about 3.5 double-tagged events if the excess were a $b\bar{b}$ state. There is about a 14% probability of 3.5 events fluctuating to 1 or 0.

Note that there is not only a lack of double tags in the signal region but apparently in the background too. We would expect a significant fraction of all these events to contain two b's from, for example, gluon splitting. If, as in the W+jets samples [2], about half the b's come from $b\bar{b}$ sources, we would expect $45 \times .5 \times .33 = 7$ double tags in the whole sample.

We also looked at the JETPROB distribution of the 2nd jet. We again used a $Z\gamma$ Monte Carlo to find the expected JETPROB distributions for b and generic jets. The data JETPROB distribution is fit with the Monte Carlo distributions. We estimate that the b-fraction of the peak region is above 50%, while the b-fraction of the side bands is below 50%. Details of the method are in Appendix D.

In addition, when we remove the $\not{E}_t$ cut and plot the dijet mass for events which are double tagged we find no evidence of peaking. See figure 18.

It is not clear if the second jet is a b or not.



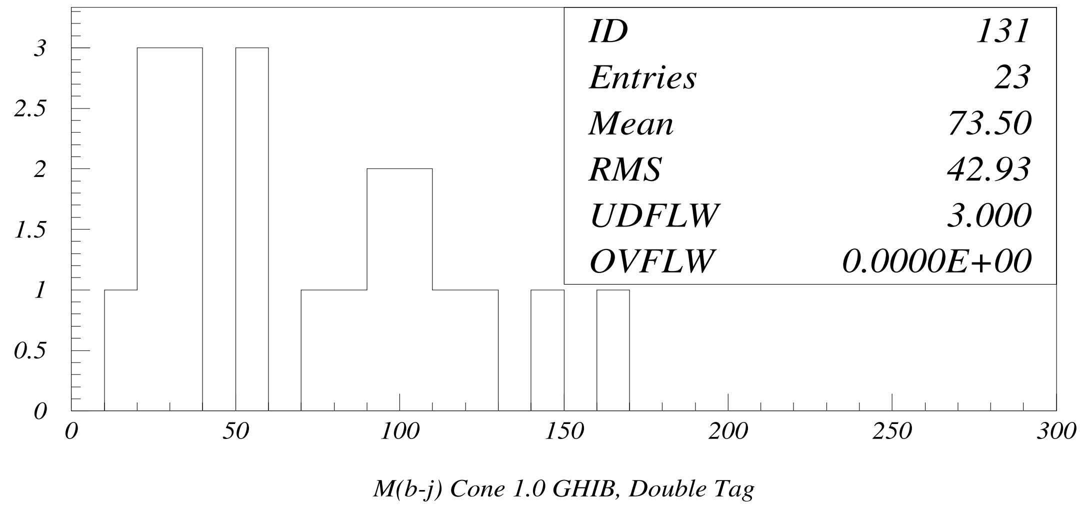


Figure 18: The dijet mass for GHIB events with two SECVTX tags but with no $\not\!\!E_t$ cut. The upper plot is made with jets with a cone of 0.4, the lower is for cone 1.0 jets.

8 The Effect of the $\not\!\!E_t$ Cut

In this section we investigate whether the $\not\!\!E_t$ cut could sculpt the dijet mass peak. Figure 19 shows the signal emerging from the inclusive photons as the $\not\!\!E_t$ cut is increased. Figure 20 is the same, but with cones of 0.4.

We also begin to look at the possibility that the $\not\!\!E_t$ cut sculpts a peak from Standard Model background i.e, events without a b-tag. Figure 21 shows the effect of the $\not\!\!E_t$ cut on the mass distribution. The data is the GHIB sample which was made with no $\not\!\!E_t$ or tagging requirement. The sequence of histograms is the dijet mass in events with two or more jets, for increasing values of the $\not\!\!E_t$ cut. The plot with points is the average dijet mass plotted *vs* the $\not\!\!E_t$ cut. Clearly the $\not\!\!E_t$ cut pushes the mass distribution to higher masses. This is probably because the jets must be energetic in order to fluctuate low and produce a significant amount of $\not\!\!E_t$. There does not appear to be a distinct peak forming but it does show that the $\not\!\!E_t$ cut tends to increase the dijet mass.

The last plot in this figure is the probability of the event passing the $\not\!\!E_t$ cut plotted *vs* the dijet mass. In the region of the peak the probability is approximately 7-10%.

To see if the signal is seeking calorimeter cracks, we can look at the detector η for these events. Figure 22 shows the η distribution for the cone 1.0 and cone 0.4 b and second jets. The solid-line histograms are events from the GHIB sample with a tag, but with no $\not\!\!E_t$ requirement. For these events, the bjet η is peaked in the central, as would be expected because of the tagging requirement. The second jet η is distributed evenly. The dashed histogram is for the signal events ($\not\!\!E_t > 25$). The bjet remains reasonably distributed but the second jet, in the cone 1.0 case, now seems to prefer the $\eta = 1.1$ cracks. The shaded histograms are the events in the signal region of the cone 1.0 dijet mass plot.

A significant complication to finding this signal is that bjets tend to have true $\not\!\!E_t$ due to energetic neutrino's. If we have a sample with no $\not\!\!E_t$ in the production mechanism, then select events with a significant amount of $\not\!\!E_t$, we probably have significantly increased the b purity. To an unknown extent, the $\not\!\!E_t$ cut is an indirect SLT requirement. The method 1 background does not take this effect into account so it could underpredict the background tags in the high-$\not\!\!E_t$ sample.

The upper histogram in figure 23 shows the $\not\!\!E_t$ distribution of 5% of the GHIB events with the Method 1 background overlayed. The background is overpredicted in the low-$\not\!\!E_t$ region but agrees in the high-$\not\!\!E_t$, signal region. (In the whole high-$\not\!\!E_t$ sample the Method 1 background underpredicts the tags, the difference is probably statistics.) We expect the low-$\not\!\!E_t$ region to be underpredicted due to the excess of charge 2/3 quarks in S.M. photon-plus-jet events. The high-$\not\!\!E_t$ region may agree better because the b purity is higher there.

The lower plot in figure 23 shows the probability of finding an SLT plotted *vs* $\not\!\!E_t$. The probability increases significantly as the $\not\!\!E_t$ increases.

In terms of the excess over Method 1 background, this effect would look just like a signal. A Method 2 background estimate should be able to separate this effect from a signal.

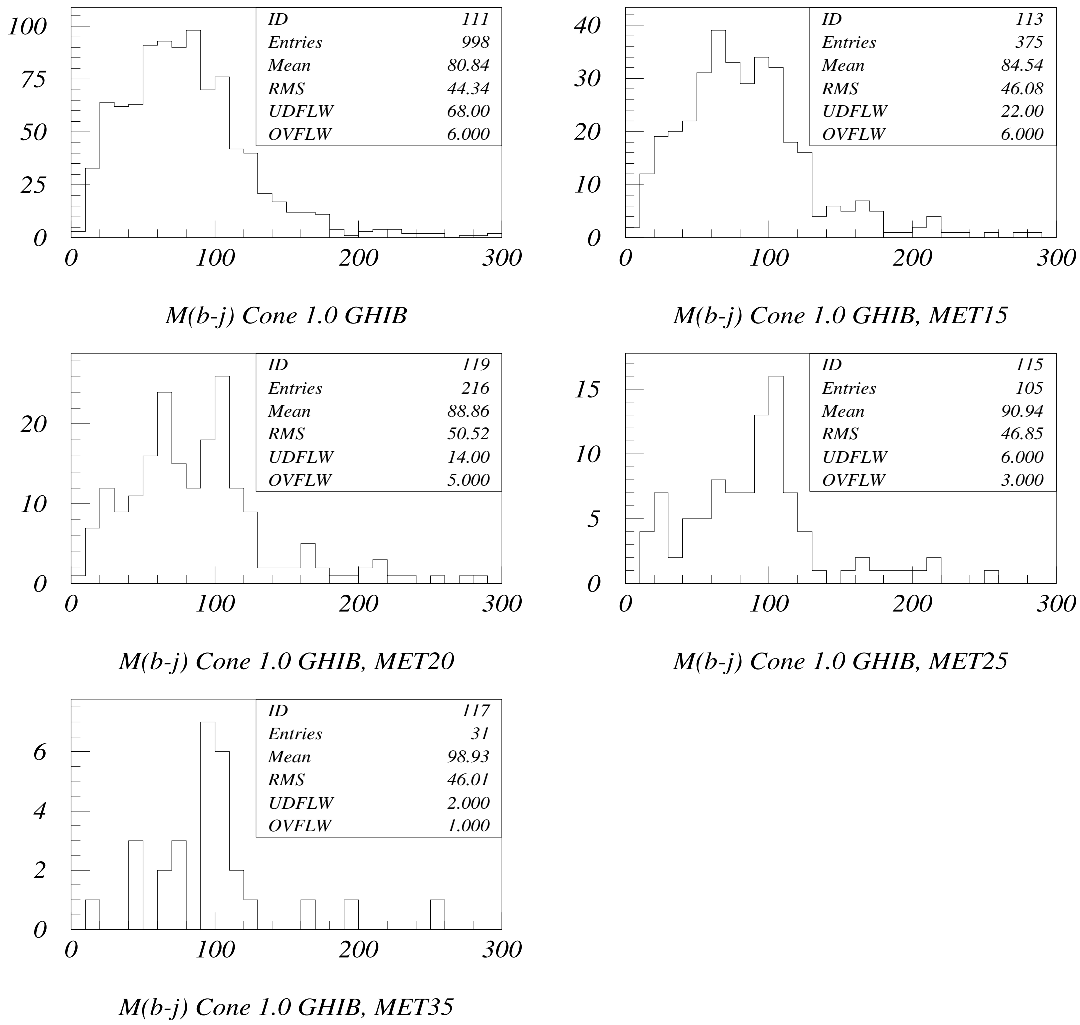


Figure 19: The cone of 1.0 dijet mass plot for photons coming in on the inclusive photon trigger. The series of plots are for increasing $\not{E}_t$ cut.

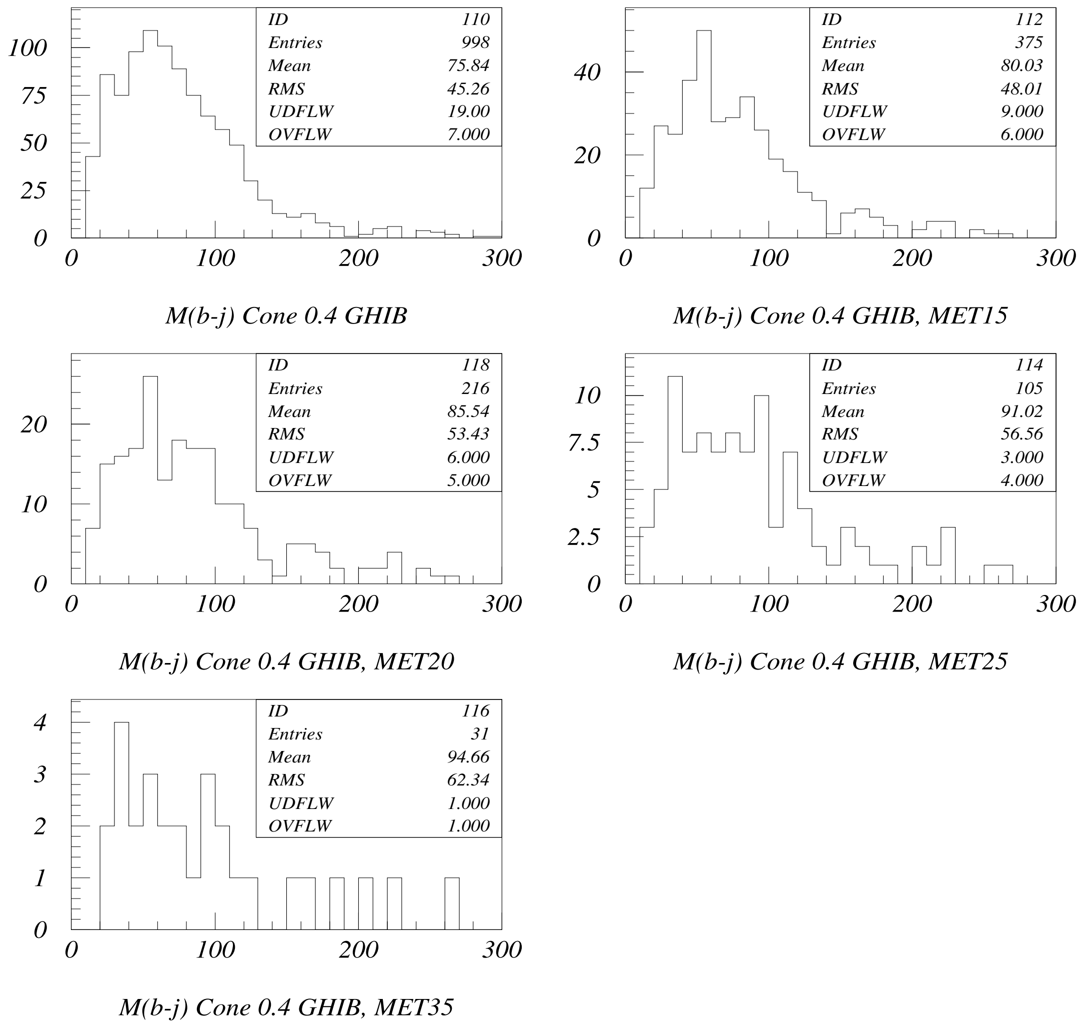


Figure 20: The cone of 0.4 dijet mass plot for photons coming in on the inclusive photon trigger. The series of plots are for increasing $\not\!\!E_t$ cut.

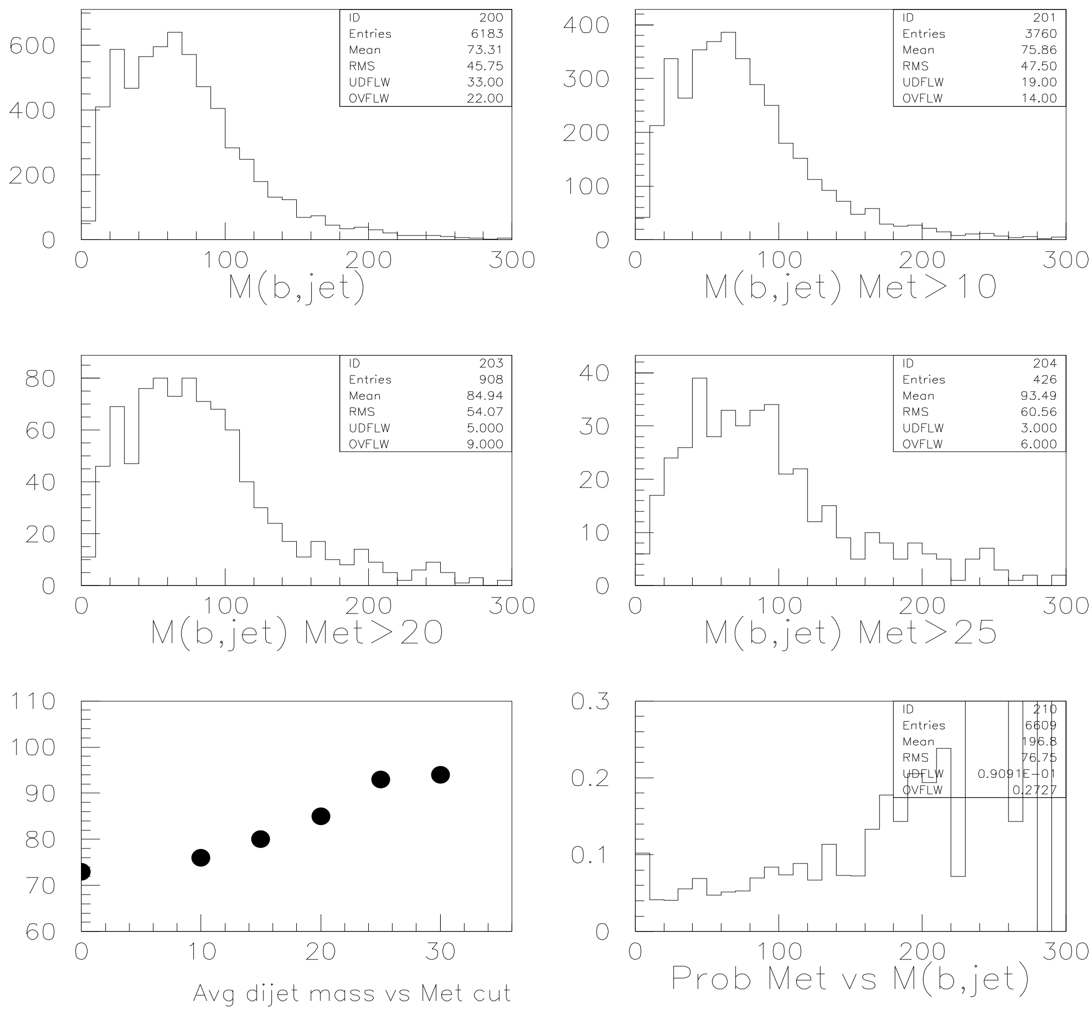


Figure 21: the dijet mass calculated with cone 1.0 jets for a photon sample starting with no $\not{E}_t$ and no tagging cut. The sequence of the upper four plots, left to right, is $\not{E}_t$ >0, 10, 20, and 25 GeV. The plot with circles is the average dijet mass plotted *vs* the $\not{E}_t$ cut. The last plot is the probability that an event passes the $\not{E}_t > 25$ cut *vs* the dijet mass.

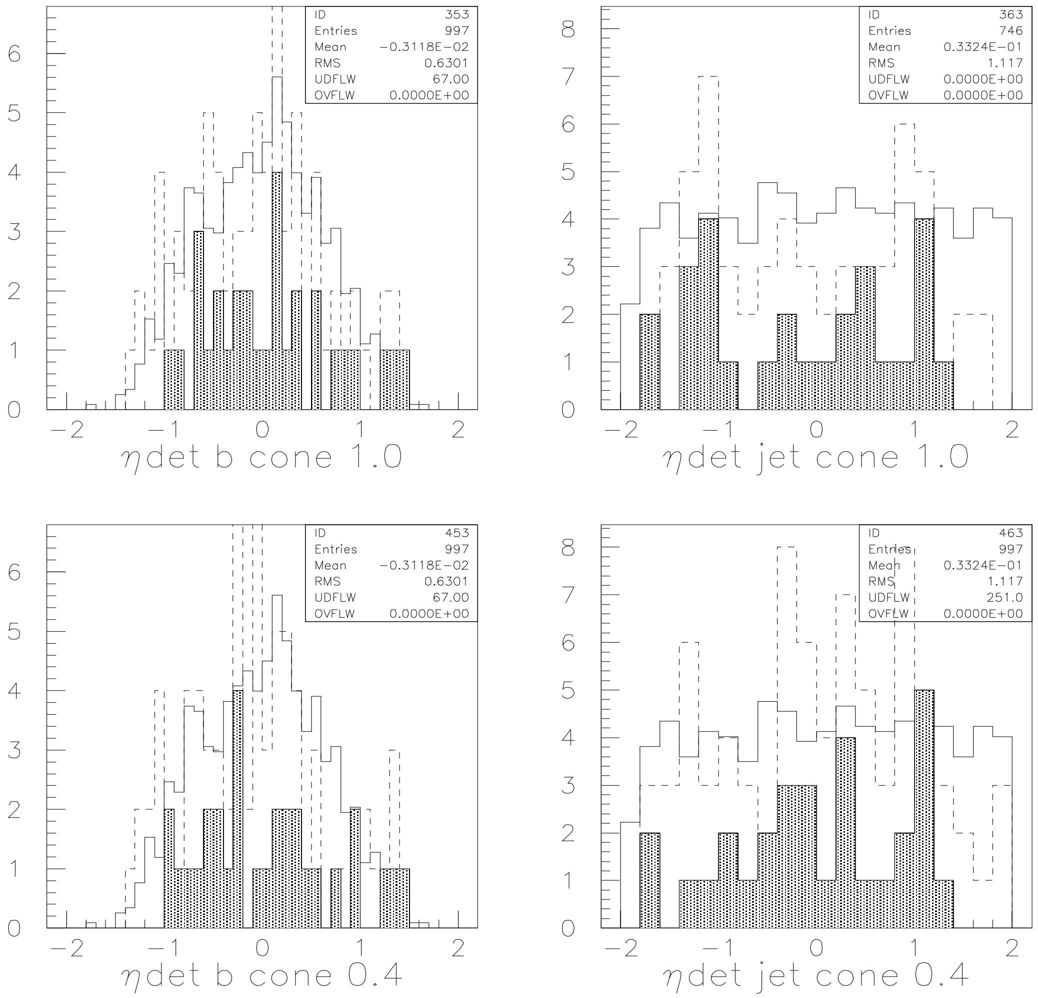


Figure 22: Detector η distributions. The histograms are the tagged GHIB sample (no $\not\!\!E_t$ cut), the dashed histograms are the signal events, (tagged, $\not\!\!E_t > 25$), and the shaded region are the events in the peak. Left column is the bjet, right is the second jet, top row is cone 1.0, bottom row is cone 0.4.

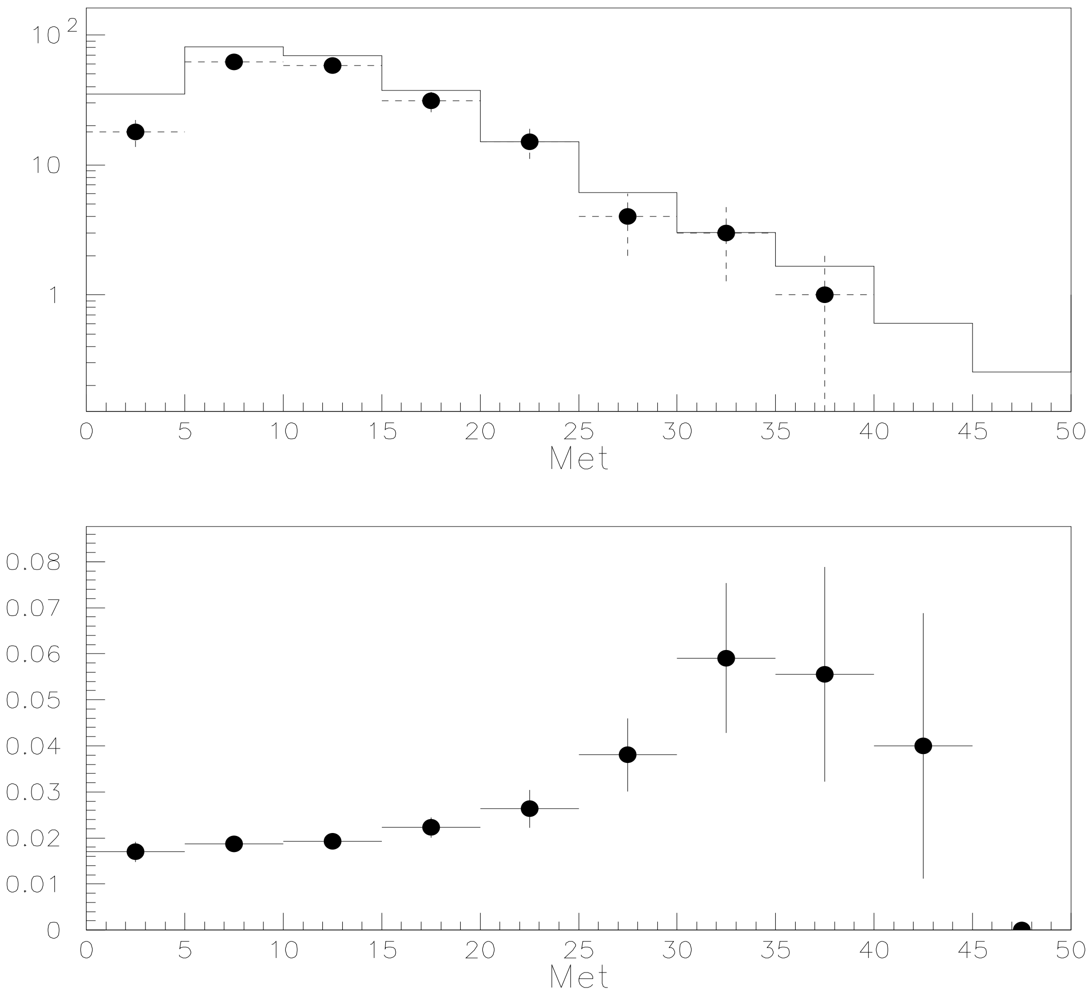


Figure 23: The upper plot is the $\not{E}_t$ for tags (points) and Method 1 background (histogram). The sample is about 5% of the GHIB photons. The lower plot is the probability of finding an SLT in the event plotted *vs* $\not{E}_t$.

9 Backgrounds

In this section we look at possible backgrounds. We try to answer the question of whether these backgrounds could create a peak in our data.

9.1 B-tag backgrounds

In the dijet mass plots with the Method 1 background overlay, the prediction matches the non-signal regions well. There is no evidence of tagging causing any peaking. See figure 2 and 5.

9.2 W and Z Backgrounds

A source of Standard Model background with this signature could be $W\gamma$ or $Z\gamma$ production. The W or Z would have to decay to jets or τ's in a manner giving a peak in the dijet mass distribution near 105 GeV. The $\not\!\!E_t$ would have to be accidental, or possibly from energetic ν's.

To begin to answer how many $W\gamma$ and $Z\gamma$ events we expect in our plots, we look in the 1A and 1B inclusive leptons for the e and μ $W\gamma$ decay modes. The strategy is to use the Monte Carlo to measure the relative acceptances and efficiencies of $\ell\gamma$ compared to $bj\gamma$, then use this to scale the number of $\ell\gamma$ events observed in the data to estimate the $W\gamma \rightarrow bj\gamma$ background in our sample.

$$N_{exp} = N_{\ell\gamma,obs} \frac{A(bj\gamma)}{A(\ell\gamma)} \times relative\ luminosity$$

We generated 1100 pb^{-1} of $W\gamma$ with Pythia/QFL, allowing the W to decay in all modes. We require the W lepton to pass TOPFND cuts and have $E_t > 25$, the event to have $\not\!\!E_t > 25$, and the photon to pass loose photon cuts and have $E_t > 25$. We find 68 events in the Monte Carlo.

Figure 24 shows the dijet mass distribution for the Monte Carlo. We find 853 events passing photon cuts and containing two jets. 113 events or 13% pass the $\not\!\!E_t$ cut. From figure 21 we saw that about 7-10% of photon-2jet events pass the $\not\!\!E_t$ cut, which is in reasonable agreement with the probability to fake the $\not\!\!E_t$. There is no apparent shift in the mass peak when the $\not\!\!E_t$ cut is applied. We find only two Monte Carlo events are tagged (both pass the $\not\!\!E_t$ cut).

In the 1A+1B data we find 44 $\ell\gamma\not\!\!E_t$ events. Since the b, jet, γ sample is 1B only, we have a relative luminosity factor of about 0.8. Plugging the results into the formula, we expect $44 \times (2/68) \times 0.8 = 1 \pm 1$ $W\gamma$ events in our sample. We expect the $W\gamma$ background to peak well below our signal region, but that prediction depends on the absolute jet energy scale.

For the $Z\gamma$ mode, we generated 2100 pb^{-1} with Pythia, allowing the Z to decay in all modes. We require the first lepton to pass the tight TOPFND cuts, $E_t > 25$, and the second to pass the loose TOPFND cuts, $E_t > 25$, and the photon to pass the loose cuts, $E_t > 25$. We find 55 events in the Monte Carlo.

Figure 25 shows the dijet mass distribution for the Monte Carlo. We find 2302 events passing photon cuts and containing two jets. 311 events or 14% pass the $\not{E}_t$ cut. We find 25 events passing the $\not{E}_t$ cut and SECVTX tagged.

In the 1A+1B data we find 5 $\ell\ell\gamma$ events. Since the b, jet, γ sample is 1B only, we have a relative luminosity factor of about 0.8. Plugging the results into the formula, we expect $5 \times (25/55) \times 0.8 = 1.8 \pm 1$ $Z\gamma$ events in our sample.

9.3 Photons from Cosmic Rays

One possibility is that the signal is Standard Model $b\bar{b}$ production with a cosmic ray causing a fake photon and thereby the $\not{E}_t$. In scanning the events, we find no evidence for spurious muon stubs. The events have already been required to have no out-of-time energy.

We have examined the CPR and CES responses and believe they are similar to S.M. dominated samples such as γ+jets before $\not{E}_t$ and tagging. This study cannot be compared with the photon group's measurements of CPR rates for π°/γ separation. Here we are only using the detector's response to see if these photons look more like typical photons or cosmics. Table 5 shows that a generic photon sample, the tagged events, and the signal region events are similar. The generic photons are made from the inclusive photon tapes, requiring a photon, $E_t > 25$, passing loose cuts and balanced by a single jet. The tagged sample is the 134 events with photon, tag and $\not{E}_t$. The signal region is the events with a dijet mass in the 90-120 GeV range. The cosmics sample is events with a loose photon, $\not{E}_t$, no primary vertex, and two or more muon stubs not attached to tracks.

Sample	fraction with CPR hit (%)	If CPR hit, fraction with CPR-CES match (%)
γ	91	76
γ, tag, $\not{E}_t$	85 ± 3	79 ± 4
Signal mass region	80 ± 7	62 ± 10
cosmics	89	44

Table 5: A comparison of the signal events and a generic photon sample. Second column is the the fraction of events with a CPR hit in the same wedge. The third column is the fraction of events with a CPR hit that also match the CES hit in ϕ.

We compare the number of events with a CPR cluster. We find the tagged events are consistent with generic photons, although they have fewer CPR clusters at the $1.5 - 2.0$ σ level.

We then compare the CPR $r \times \phi$ to the CES $r \times \phi$. If they agree within 5cm, we call it a match. Because cosmics come in from all angles, they have more random CPR-CES positions. Again we find the tagged events are consistent with generic photons, although they have worse CPR agreement at the 1.5 σ level.

We can use the CES-CPR matching rates to estimate the number of cosmics in our sample. The signal sample is consistent with all photons, $0 \pm 10\%$ cosmics. The signal region result is $40 \pm 30\%$ cosmics.

9.4 Photons from W's

$W \rightarrow e\nu$ can fake a photon if the track is lost. To estimate how many of these events we expect in our sample, we can look at the second leg of a $Z \rightarrow ee$ peak. The first leg is required to pass good electron cuts. With no requirements on the second ELES cluster (except that it is central), we find 2994 ± 72 events in the Z peak. If we require the second cluster passes our loose photon cuts, we find 83 ± 17 events are still in the peak. This implies that the probability of an electron being accepted as a photon is $2.8 \pm 0.6\%$. There are less than 50 $W \rightarrow e\nu$ events with a SECVTX tag so we would expect less than 1.5 W events in our sample.

9.5 Photon from Fakes

It is well known that prompt photons can be faked from π^0, η or Kaons. Inclusive photon studies have shown that there is roughly a 60% prompt photon purity in our E_t range. See Figure 26 in Appendix B which shows what happens when we use tighter cuts on the photons to increase purity.

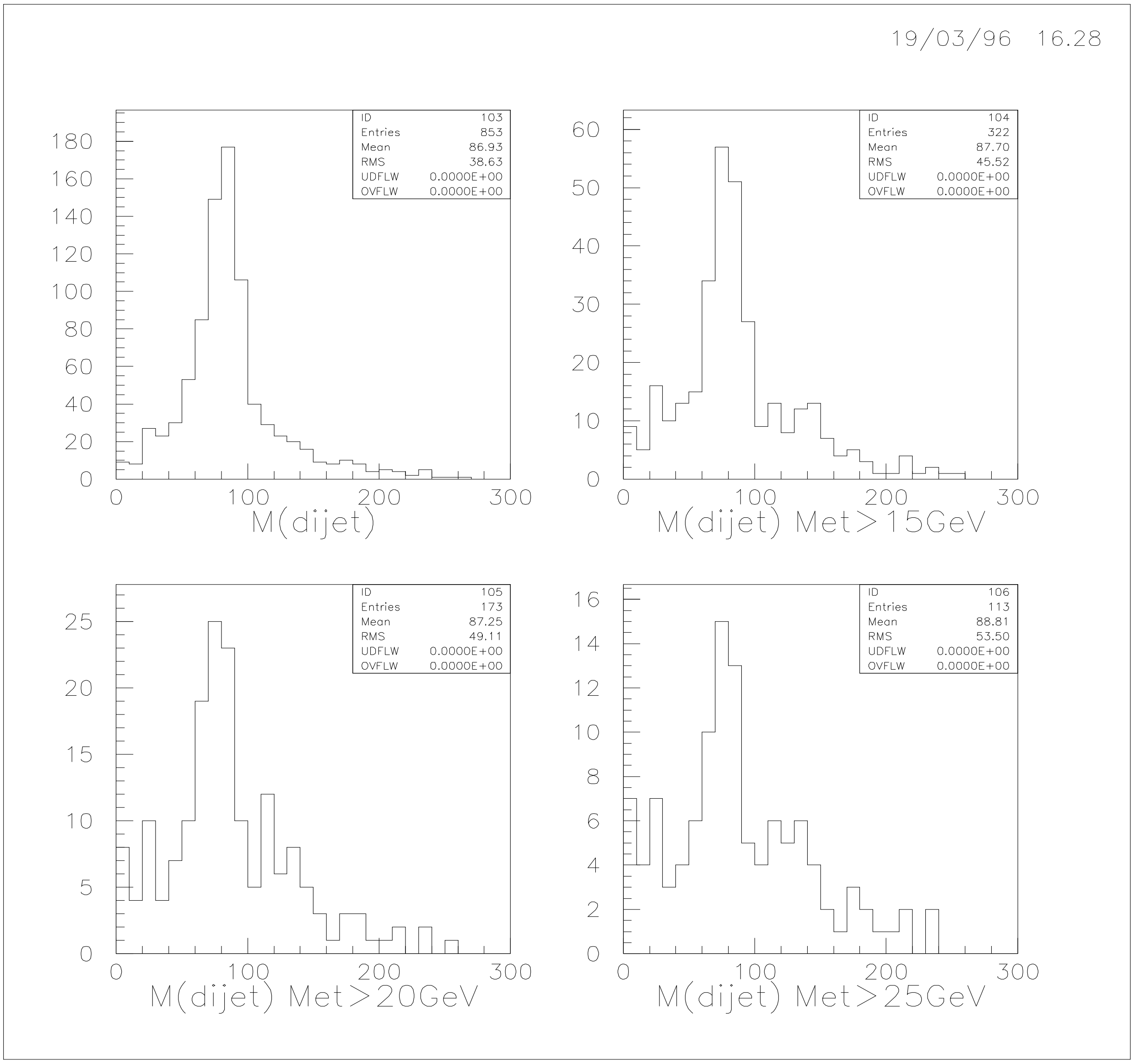


Figure 24: The dijet mass for $W\gamma$ Pythia Monte Carlo for various $\not\!\!E_t$ cuts. Only two events have a SECVTX tag. Jet cones of 0.4 are used.

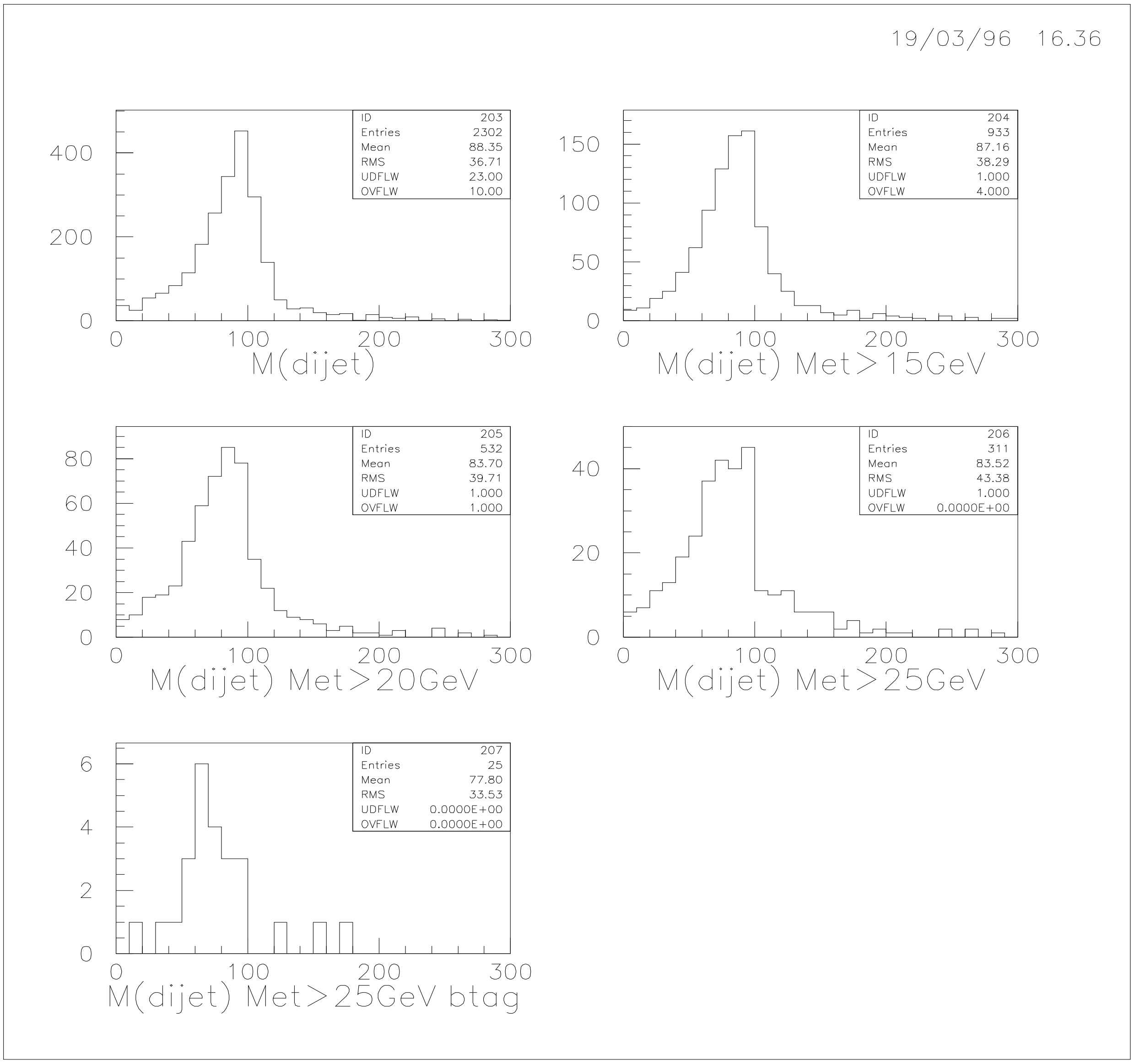


Figure 25: The dijet mass for $Z\gamma$ Pythia/QFL Monte Carlo for varying $\not\!\!E_t$ cut. The last plot also requires a SECVTX tag. Jet cones of 0.4 are used. (The Z mass using cones of 1.0 is shown in figure 14.)

10 Conclusions

We see an excess in the number of b-tags in events with a photon, $\not{E}_t$, and a jet over the top-specific Method 1 background estimate. The excess is localized in the dijet mass distribution, and is narrow. The signal prefers higher photon E_t and $\not{E}_t$.

This is an unusual data set in the collection of objects it contains. Each one needs to be checked for fakes. Each also needs to be studied for how it might contribute to a fake signal. We comment on each one here.

In standard photon analyses, there is typically 60-70% π°'s in the photon samples with $E_t > 25$, which we have yet to consider.

The $\not{E}_t$ is potentially troubling. There is a large S.M. background in photon and jets, what is the effect of selecting a subsample with large $\not{E}_t$? It will select mis-measured jets, could it create a peak? We have looked at the untagged sample and there is no obvious peak created by a $\not{E}_t$ cut, although the mean dijet mass does increase (towards our peak) with increasing $\not{E}_t$ cut.

The $\not{E}_t$ is always opposite the photon. This would be expected in a S.M. sample where the photon is measured accurately and the recoiling jets tend to be measured low.

The most troubling aspect of the peak is the fact that it prefers cone 1.0 jets and almost disappears in the cone 0.4 dijet mass plot. If the peak contains S.M. background with very mismeasured jets is this what we would expect to see? Scanning the events that move in and out of the peak reveals abnormalities but not a clear pattern of abnormalities. Sometimes the cone 1.0 is justified and sometimes not. Using JTC96S corrections instead of JTC90S makes the peak significantly broader and more consistent with either a dijet peak or a sculpted distribution.

The btag has a reasonable $c\tau$ (See figure 27 in Appendix B) distribution and SECVTX is well-understood. However, the Method 1 background may not accurately represent the background in a sample with $\not{E}_t$ since b's tend to produce $\not{E}_t$ through neutrinos. The second jet has conflicting evidence for whether or not it is a another b.

Referring back to figure 1, we see most of the tagging excess in the one and three jet events which we do not understand yet. In addition we have shown that the signal tends to have a larger photon E_t and $\not{E}_t$ but the tagging excess does not appear to be in those regions.

Possibly some form of the following may be going on: the $\not{E}_t$ cut requires the jets to be fairly big and mismeasured to produce the $\not{E}_t$. This will give a reasonable mass spectrum. The cone of 1.0 then takes the subset of events with scattered energy and promotes them from the low end of the dijet spectrum to the middle (where there may have already been a fluctuation high). This could be correlated with something as mundane as additional vertices providing the scattered energy.

Of course, it could also be new physics - an object that decays to two jets, possibly related to the similar possible excess in $W + 2$jets.

11 Appendix A- Photon Cuts

Loose Photon Cuts		
Central		
P_t of "electron"	$<$	5 GeV
CES $(\chi_s^2 + \chi_w^2)/2$	$<$	20
Lshare	$<$	0.2
Had/EM	$<$	0.055 + 0.00045E
E_T in a cone of 0.4/ E_T	$<$	0.15
P_t of tracks in a cone of 0.4	$<$	10 GeV

Tight Photon Cuts		
Central		
3-D tracks pointing at the EM cluster	$=$	0
CES $(\chi_s^2 + \chi_w^2)/2$	$<$	20
Lshare	$<$	0.2
Had/EM	$<$	0.055 + 0.00045E
E_T in a cone of 0.4/ E_T	$<$	0.15
P_t of tracks in a cone of 0.4	$<$	2 GeV
No second CES strip cluster with E	$>$	1 GeV in the same wedge

Table 6: Loose and tight photon cuts. We use the loose cuts here.

12 Appendix B- Basic Checks

Figure 26 shows the cone 0.4 and cone 1.0 mass distributions when the tight photon cuts in Appendix A are used. A few more peak events are lost compared to the overall loss, but the losses are not inconsistent within statistics. Figure 27 shows that the $c\tau$ distribution of the tags agrees well with a top Monte Carlo.

Real b jets should have a natural mass on the order of the b quark mass. Figure 28 shows the b mass distribution using a cone of 0.4 and the effect of requiring the mass to be near the b quark mass.

One possible way to find anomalous mass peaks in jets is to mismeasure the jet energy. The guard energy for a jet is the fraction of energy in a tower that borders the 90° or 30° cracks; it should flag jets that might lose significant energy. Figure 29 shows guard energy distribution and the effect of requiring low guard energy.

Figure 30 shows some kinematics of the events: the P_t of the b+jet system, the b+jet mass plotted *vs* the the $\Delta\phi$ between the γ and the b+jet system, the P_t of the γ plotted *vs* the P_t of the b+jet system, the P_t of the b plotted *vs* the P_t of the jet.

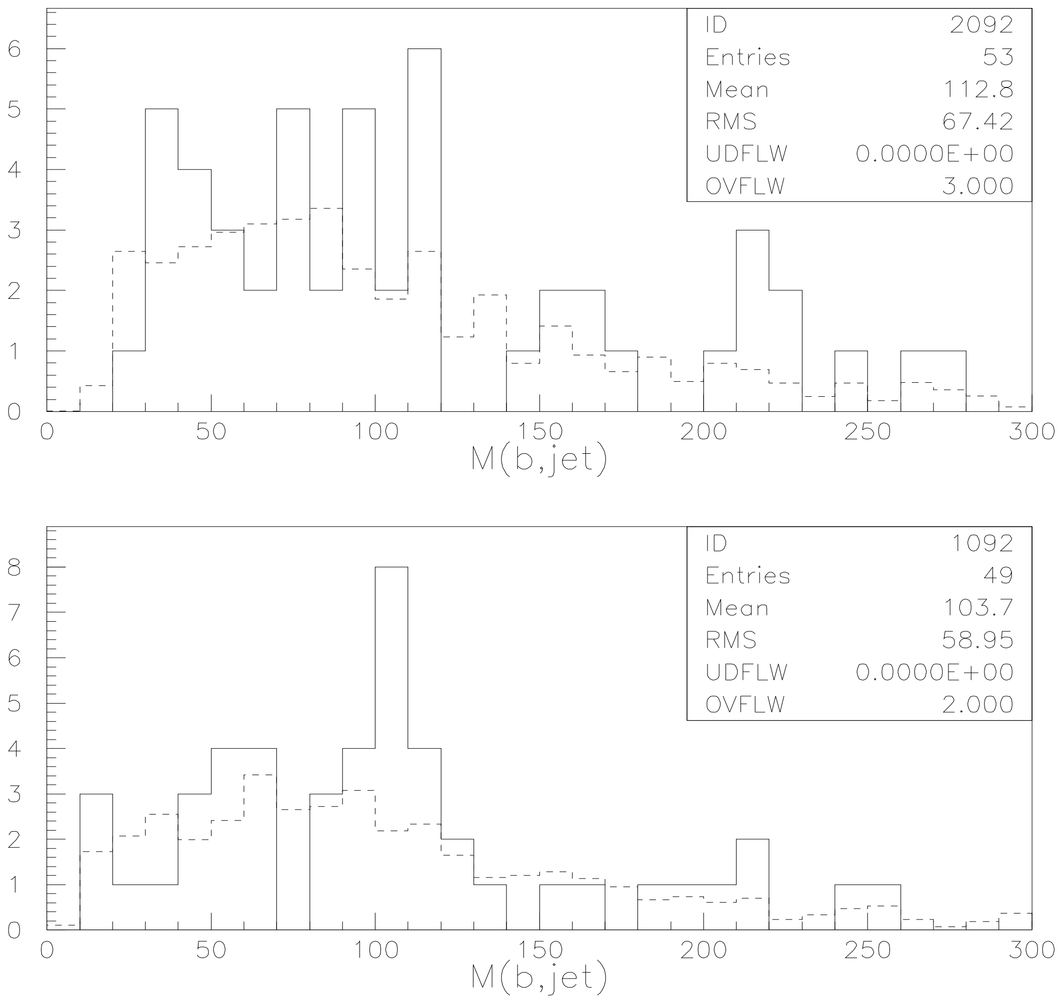


Figure 26: The cone 0.4 (top) and cone 1.0 (bottom) $b + jet$ mass plots with tight cuts on the photon.

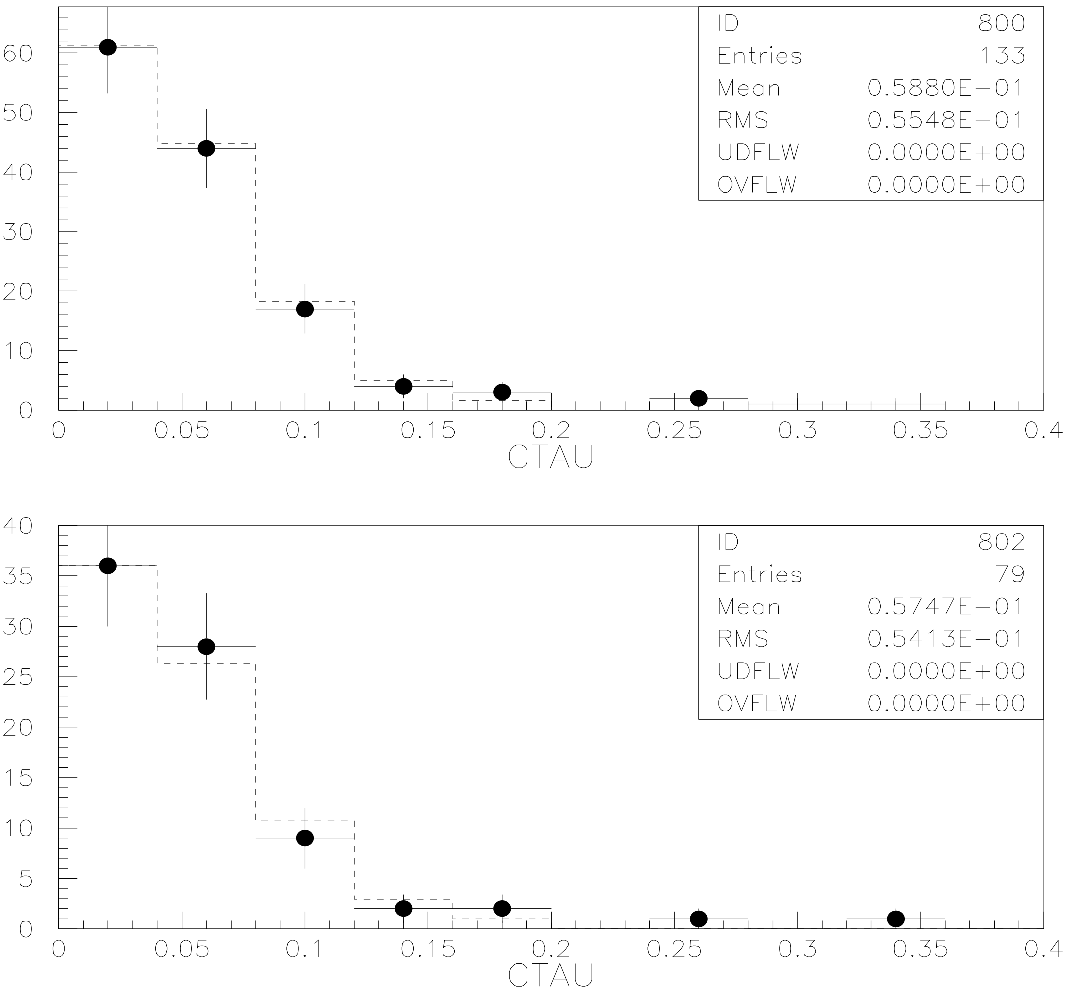


Figure 27: The $c\tau$ distribution of tags in the 133 tagged events (upper) and the 79 2-or-more-jet tagged events (lower). The overlay is from a top 175 Monte Carlo.

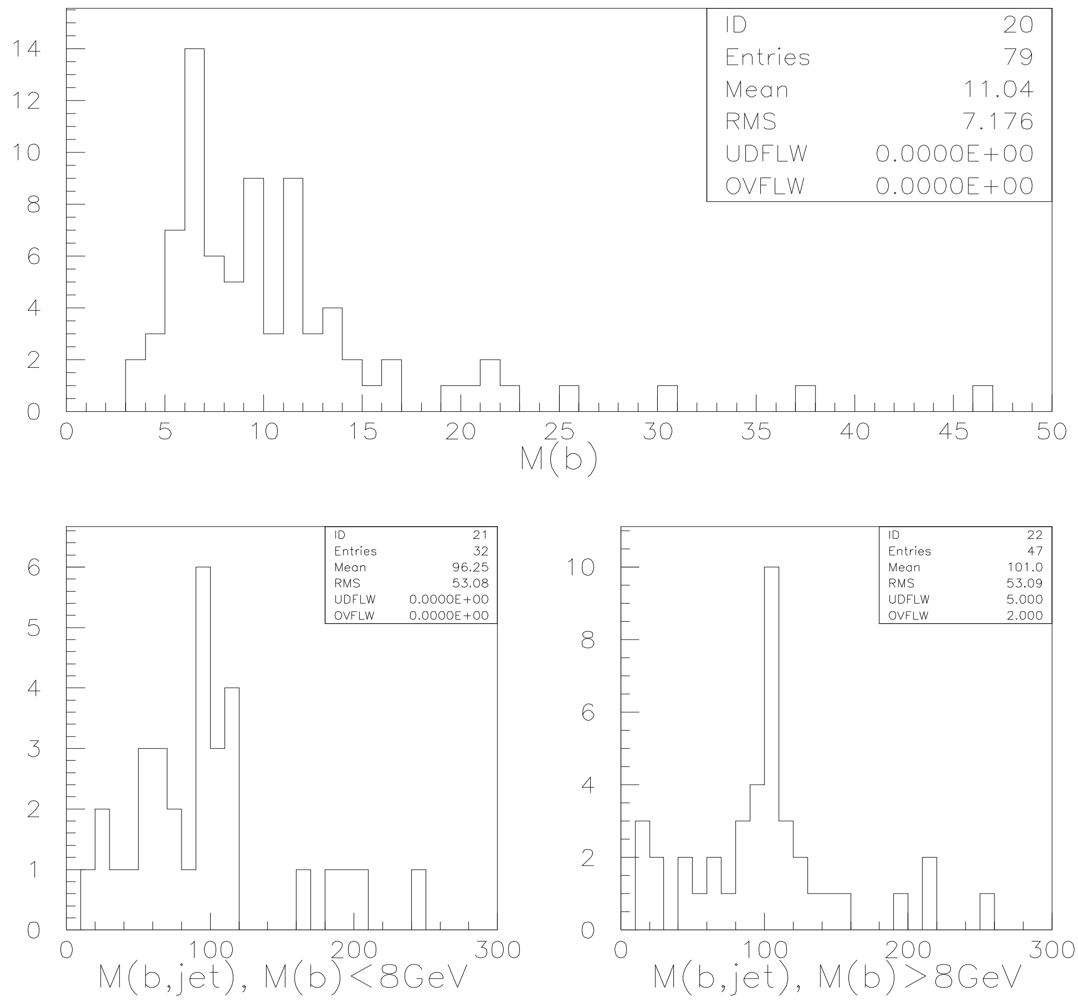


Figure 28: Top plot is the mass of the cone 0.4 b jet. The lower two histograms show the cone 1.0 dijet mass distribution for low and high b jet mass.

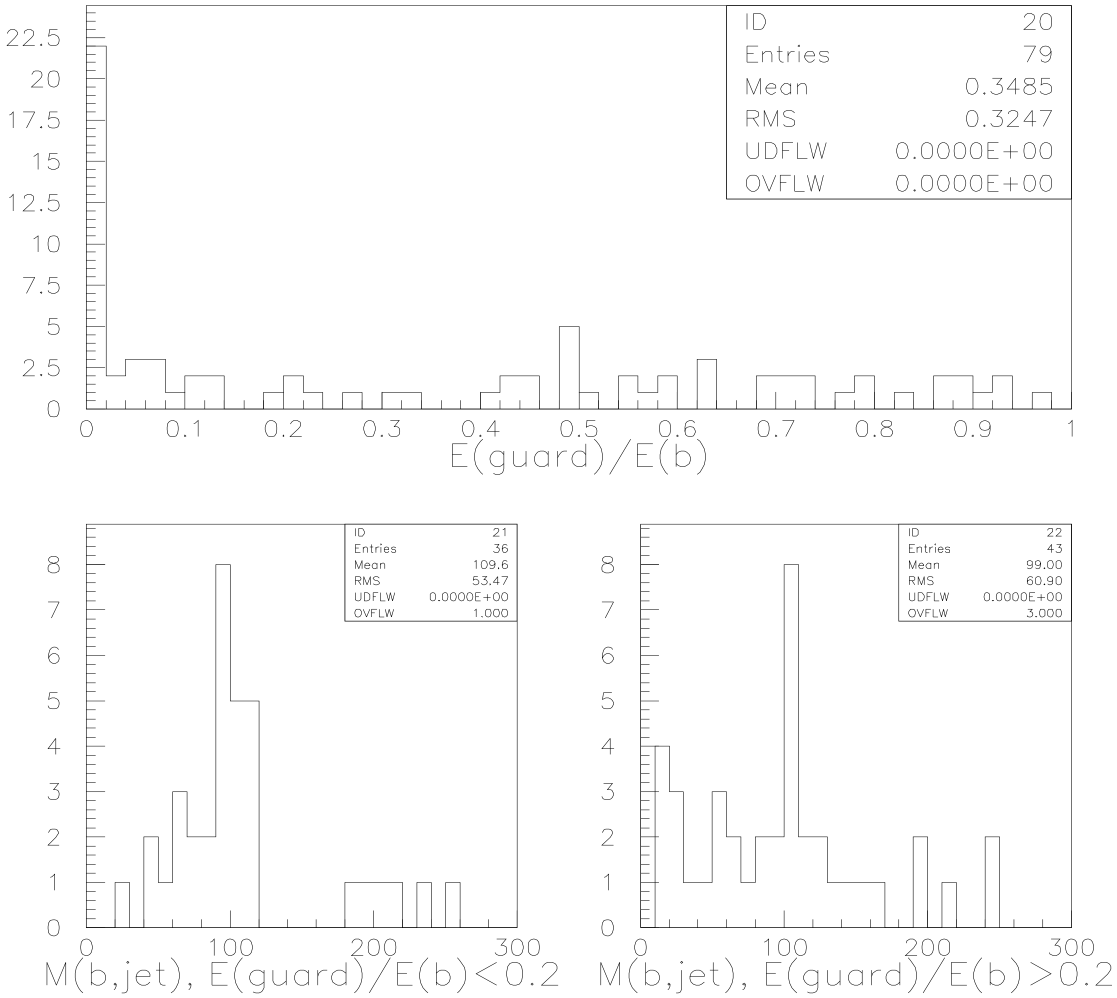


Figure 29: Top plot is the fraction of the cone 0.4 b jet energy in a guard region (major calorimeter cracks). The lower two histograms show the cone 1.0 dijet mass distribution for high and low guard energy fraction.

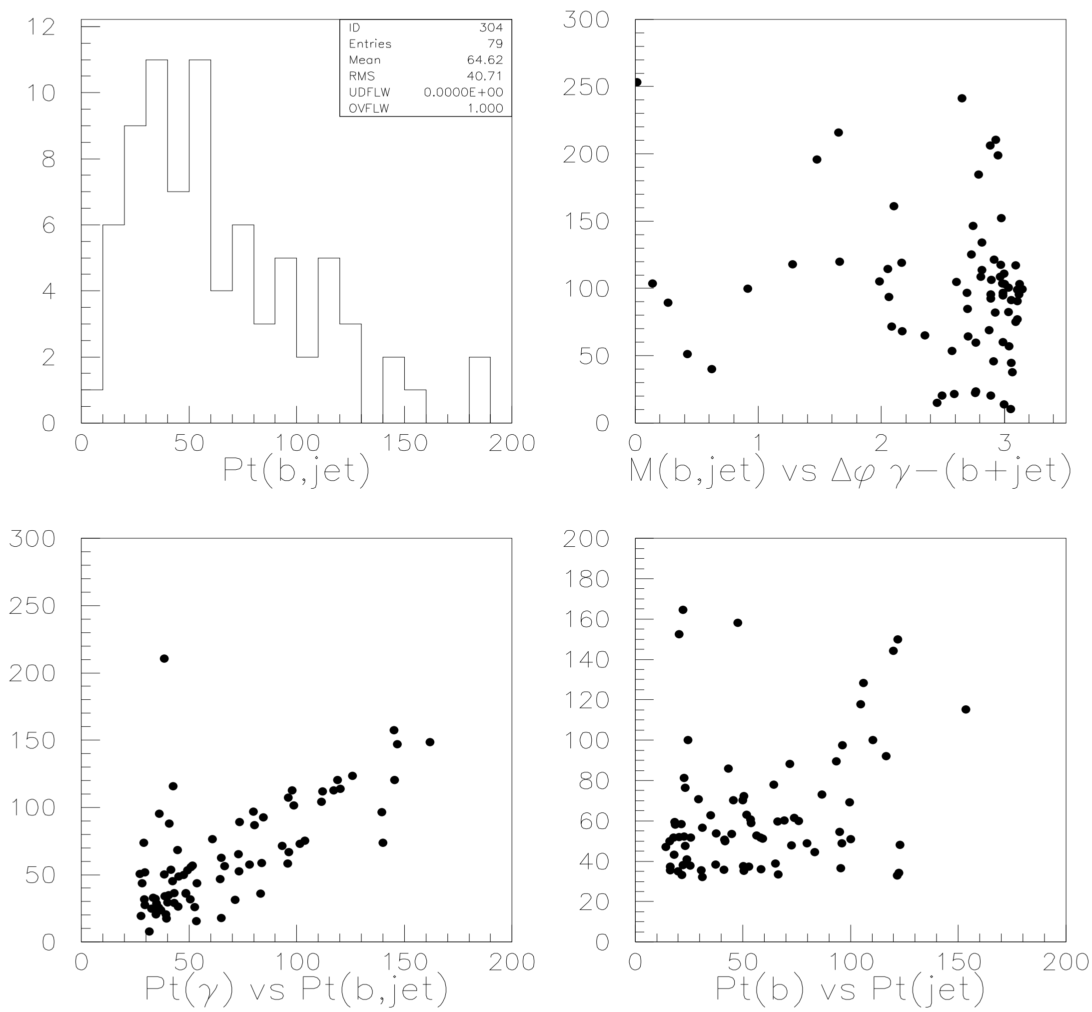


Figure 30: Kinematics of the tagged events. Upper left, the P_t of the b+jet system; upper right, the b+jet mass plotted *vs* the the $\Delta\phi$ between the γ and the b+jet system; lower left, the P_t of the γ plotted *vs* the P_t of the b+jet system; lower right, the P_t of the b plotted *vs* the P_t of the jet. Jets with a cone of 0.4 and AA corrections are used.

13 Appendix C- Trigger Studies

The photon plus btag analysis has been done using a photon plus $\not{E}_t$ sample stripped from the stream B inclusive electron tapes (ECMB). Cross checks to this analysis have been done using an inclusive photon sample stripped from the stream B inclusive photon tapes (GHIB). Here we describe the trigger path of these samples.

Photons on ECMB tapes come predominantly (98.5%) from the L3 W no track trigger: ELEA_CEM_25GEV_W_NOTRK. This trigger requires $\not{E}_t > 25$ GeV, $E_T > 25$ GeV, $HAD/EM < 0.05$, $LSHR < 0.2$, $E_{border} < 2.5$ GeV and $\chi^2_{str} < 20$. The remaining events come in on miscellaneous electron triggers (*i.e.* they are $e + \gamma$ events). Note that the HAD/EM and Border tower cuts are tighter than those used for our offline photon requirements and the $\chi^2_{str} < 20$ will remove a few events that would pass the $(\chi^2_{str} + \chi^2_{wir})/2 < 20.$ cut used offline.

The GHIB tapes contain photons from 3 L3 streams:

GHIB L3 - PHOB_CEM3_23_ISOL* requires L2 - CEM_23_ISO_XCES*

GLOB L3 - PHOB_CEM4_10_ISOL*,

GNIB L3 - PHOB_CEM*_*_NOISOL, PHOA_PHO_JET

At Level 3 all of these triggers prerequisite a corresponding L2 trigger and require a cluster passing the E_T cut and fiducial cuts: $|x| < 21.$cm and $9. < |z| < 217.$cm ($|x| < 17.5$cm and $14. < |z| < 217.$cm for 10 GeV triggers). For our purposes we can ignore the GLOB stream since we are cutting at 25 GeV in our skim. For these trigger studies we worked on a small subset of the stripped events concentrating on the 16905 events stripped from tape CCR482 (a total of 740K events were stripped from all tapes).

At L2 there are only 3 unprescaled triggers that feed our samples:
CEM_23_ISO_XCES_V*, CEM_50_V* and MET_20_CEM_16_XCES.

Table 7 lists the dominant L2 and L3 triggers feeding these two data sets and their contributions. For the L2 triggers two lists are presented: first the inclusive contributions from the 3 unprescaled triggers, then an exclusive list starting with the CEM_23_ISO_XCES_V* trigger and then listing the additional contributions from other L2 triggers.

The primary L3 trigger for the ECMB sample is the W_NOTRACK as expected. However, only 75% of the events pass the PHOB_CEM3_23_ISOL*. Clearly there is an inefficiency involved in the combination of the L3 fiducial cuts and the NN + XCES requirements of the L2 for the photon triggers. Applying the fiducial cuts used in the photon trigger to the ECMB events results in a 91% efficiency which is consistent with the ratio of the L2 to the L3 23 GeV photon triggers. This implies that the combination of Neural Net isolation and XCES have at least an 18% inefficiency for our loose photons. Most of the remaining 18% of ECMB events come from CEM_50 and the MET_20_CEM_16 triggers.

On the GHIB tapes almost all of the events are fed by the 23 GeV ISO trigger. With the MET requirement, the 50 GeV NOISO trigger contributes significantly but only 57% of the events pass the W Notrack trigger. Figure 31 shows the distributions of HAD/EM, E_{border}, χ^2_{strip} and χ^2_{wire} for photons in GHIB events with $\not{E}_t > 25$ GeV. The photons were required to pass the loose photon cuts. The vertical lines show the cuts applied at L3 in the

Trigger	Fraction of Events ECMB	GHIB	GHIB w/ $\not{E}_t > 25$
Level 3:			
ELEA_CEM_25GEV_W_NOTRK	0.985	0.053	0.571
PHOB_CEM3_23_ISOL*	0.75	0.966	0.864
PHOB_CEM5_50_NOISOL*	0.291	0.101	0.392
Level 2:			
CEM23_ISO_XCES*	0.823	0.967	0.868
CEM50*	0.281	0.079	0.340
MET_20_CEM_16_XCES	0.903	0.325	0.839
Level 2 (exclusively):			
CEM23_ISO_XCES*	0.823	0.967	0.861
CEM50*	0.053	0.017	0.108
Other Photon	0.025	0.004	0.006
MET_20_CEM_16_XCES	0.084	0.004	0.017
All other	0.015	0.008	0.003

Table 7: Fraction of events passing various trigger requirements for the two photon skims. For Level 3 the fractions are inclusive while for Level 2 they are exclusive starting from the top of the list.

W_NOTRACK trigger. Table 8 lists the fraction of events passing these three requirements. Almost all of the loss comes from the E_{border} cut. Figure 31d shows the efficiency as a function of photon E_T.

Figure 32 shows the same cuts and efficiency for a sample of electrons from $Z \to ee$. The Z's were selected requiring a central electron passing the tight R analysis cuts. The second leg was then required to pass all the loose photon cuts except those on the tracks. A summary of the efficiencies appears in Table 8. The same efficiency trend appears in the electrons although with less dramatic effect. Clearly the efficiency has a significant E_T dependence. The fact that the electrons are more efficient probably indicates extra junk in the GHIB sample which is being removed by the W_NOTRK cuts.

	GHIB		$Z \to ee$	
L3 Cut	Individual	Cumulative	Individual	Cumulative
$HAD/EM < 0.05$	0.92	0.92	0.95	0.95
$E_{border} < 2.5$GeV	0.74	0.67	0.89	0.85
$\chi^2_{strip} < 20$	0.96	0.66	0.99	0.84

Table 8: Fraction of GHIB events passing W_NOTRACK L3 cuts. The events were required to pass loose photon cuts and have $\not{E}_t > 25$ GeV. The last two columns are for electrons from $Z \to ee$ passing all the loose photon cuts (except on tracks).

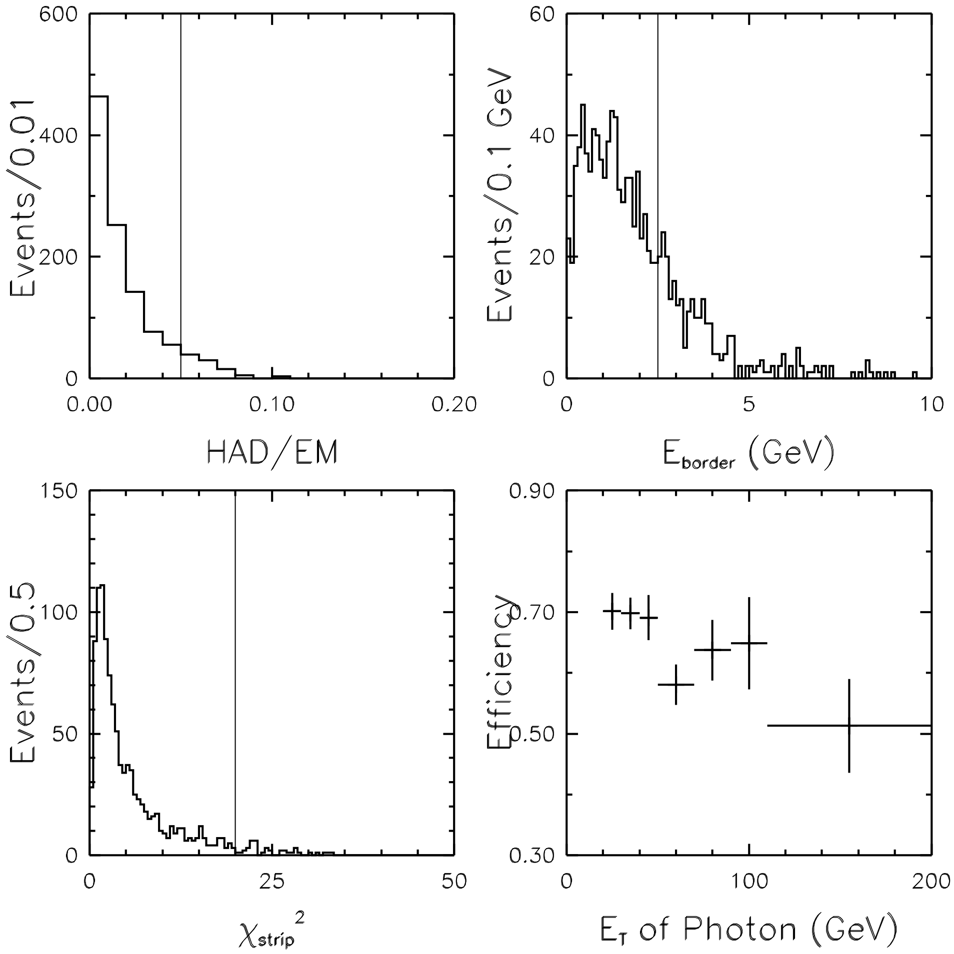


Figure 31: Distributions of a) HAD/EM, b) E_{border}, and c) χ^2_{strip} for loose photons from the GHIB sample. The vertical lines indicate the location of cuts for the W_NOTRK trigger. The Efficiency as a function of E_T is shown in d).

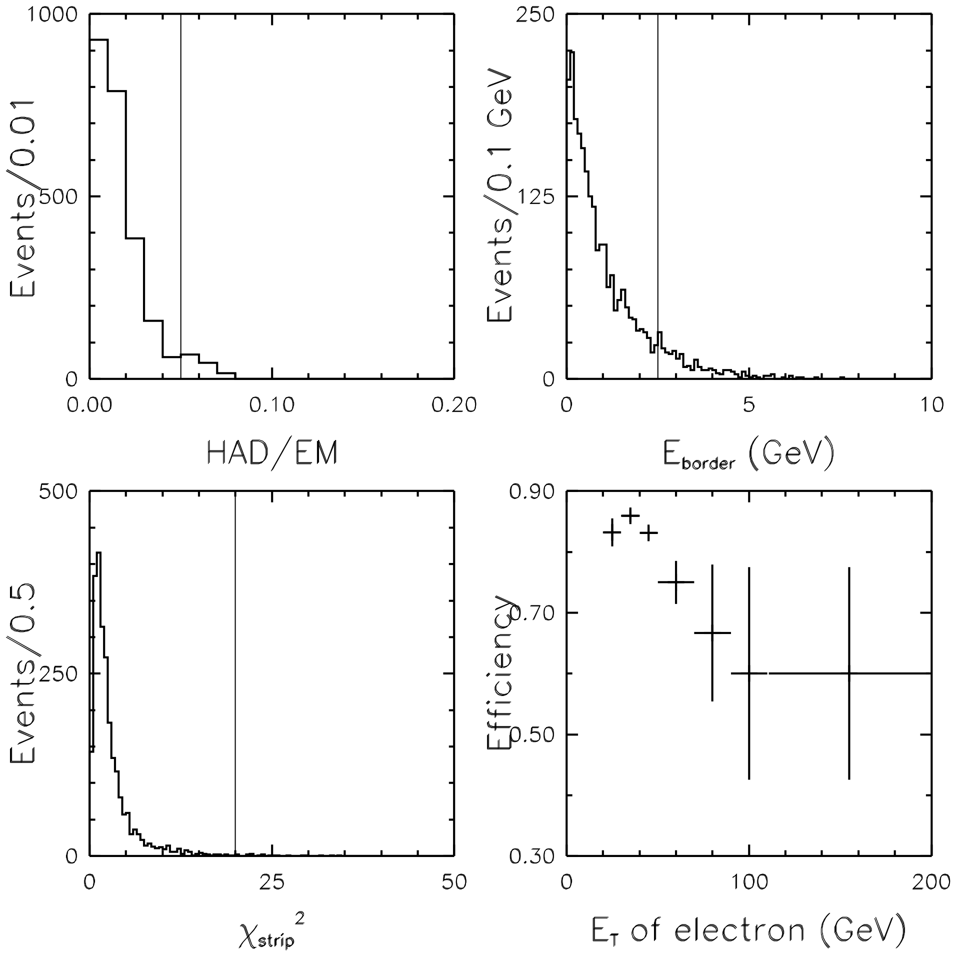


Figure 32: Distributions of a) HAD/EM, b) E_{border}, and c) χ^2_{strip} for electrons passing loose photon cuts (except track cuts) from a sample of $Z \rightarrow ee$. The vertical lines indicate the location of cuts for the W_NOTRK trigger. The Efficiency as a function of E_T is shown in d).

14 Appendix D - JETPROB Fit to the Second Jet

One way to estimate the $b\bar{b}$ content is to look at the JETPROB distribution for the second jet. Others have studied this method extensively; at this point we are just making an estimate of the b-content of the second jet when the first jet is tagged.

To do this we use the JETPROB distribution for the second jet and compare it with MC shapes expected for b-jets and generic jets. Figures 33 and 34 shows the JETPROB distribution for the data as well as for the expected shapes from γZ for the cases where $Z \rightarrow u\bar{u}$, $Z \rightarrow c\bar{c}$, and $Z \rightarrow b\bar{b}$. In the $b\bar{b}$ case we require one jet to be tagged and plot the JETPROB of the other jet. For the $u\bar{u}$ and $c\bar{c}$ cases, instead of requiring the first jet to be a fake, we simply look at the jet with the smaller SECVTX Method 1 probability to simulate the jet which is less likely to be tagged.

From the figure, there are basically three regions: The signal region (The first bin), the intermediate region, and the untaggable region (the last bin). Since the last bin includes jets outside the SVX fiducial region we ignore it for the purposes of this calculation In addition we make the simplifying assumption that the shapes for $u\bar{u}$ and $c\bar{c}$ are the same[1]. Finally, we estimate the amount of b-quarks in the 2nd jet using the following equations:

$$N_{Data}(FirstBin) = N_b(FirstBin) + N_G(FirstBin)$$

$$N_{Data}(OtherBins) = N_b(OtherBins) + N_G(OtherBins)$$

Where N_{Data} is the number of events in the Data, N_b is the number of b-quark jets, and N_G is the number of generic jets. From the Figures we find the following:

$$\frac{N_G(FirstBin)}{N_G(OtherBins)} = \frac{16}{161} = 0.10 \pm 0.03$$

$$\frac{N_b(FirstBin)}{N_b(OtherBins)} = \frac{36}{35} = 1.03 \pm 0.24$$

For the peak region we have:

$$\frac{N_{Data}(FirstBin)}{N_{Data}(OtherBins)} = \frac{4}{10} = 0.4 \pm 0.24$$

Outside the peak region we have:

$$\frac{N_{Data}(FirstBin)}{N_{Data}(OtherBins)} = \frac{4}{14} = 0.29 \pm 0.16$$

We now have two equations and two unknowns. Plugging in and solving we find that inside the peak region we estimate 6.7 b-jets and 7.3 generic jets (46.8±20%). Outside the peak region we estimate that the b-content is 31.6±20%. We note that doing a likelihood fit to all 9 bins, we find that the ratio's change to 59±17% and 24±15% respectively. Both these numbers are consistent with non-zero b-content at the 1.6σ level which is not terribly

[1]We note that this is not a particularly good assumption, since a charm is twice as likely to be in the first bin as an up quark. Specifically $\frac{N_c(FirstBin)}{N_c(OtherBins)} = \frac{32}{134} = 0.24 \pm 0.04$

surprising since one expects that a good fraction of the SM b's produced in association with a γ are from $g\gamma$ production where there was gluon splitting via $g \rightarrow b\bar{b}$. However, there is roughly a 1σ difference between the peak and sideband regions. We again note that the absolute values of the percent b-content is an overestimate since some of the b-jets come from charm quarks.

If we take these numbers seriously, we can do a background subtraction in the peak region and estimate the b-content of our 'signal.' Specifically, we assume that the b-content of the events in the last bin of the JETPROB distribution is the same as the rest of the data. Thus, if 59% of the 30 events in the peak are from b's, we estimate that there are 17.7 b quarks. We estimate the background in the peak region in two ways: 1) From the sideband arguments there appear to be about 9 events from background. 2) Method I predicts about 12 events underneath the peak. We use for this estimate 10$\pm$ 3. Thus, we expect that 2.4 b's from background processes in the peak (10 Events $\times$24% b's in the background). So, we expect that the b-fraction of events from the 'signal' is.

$$b_{Fraction} = \frac{17.7 - 2.4}{30 - 10} = \frac{15.3}{20} = 77\%.$$

The SECVTX double tagging rate would imply that the excess probably is not a $b\bar{b}$ state, while the JETPROB result would imply that it probably is a $b\bar{b}$ state. With the statistical errors at this time, it is hard to make a statement.

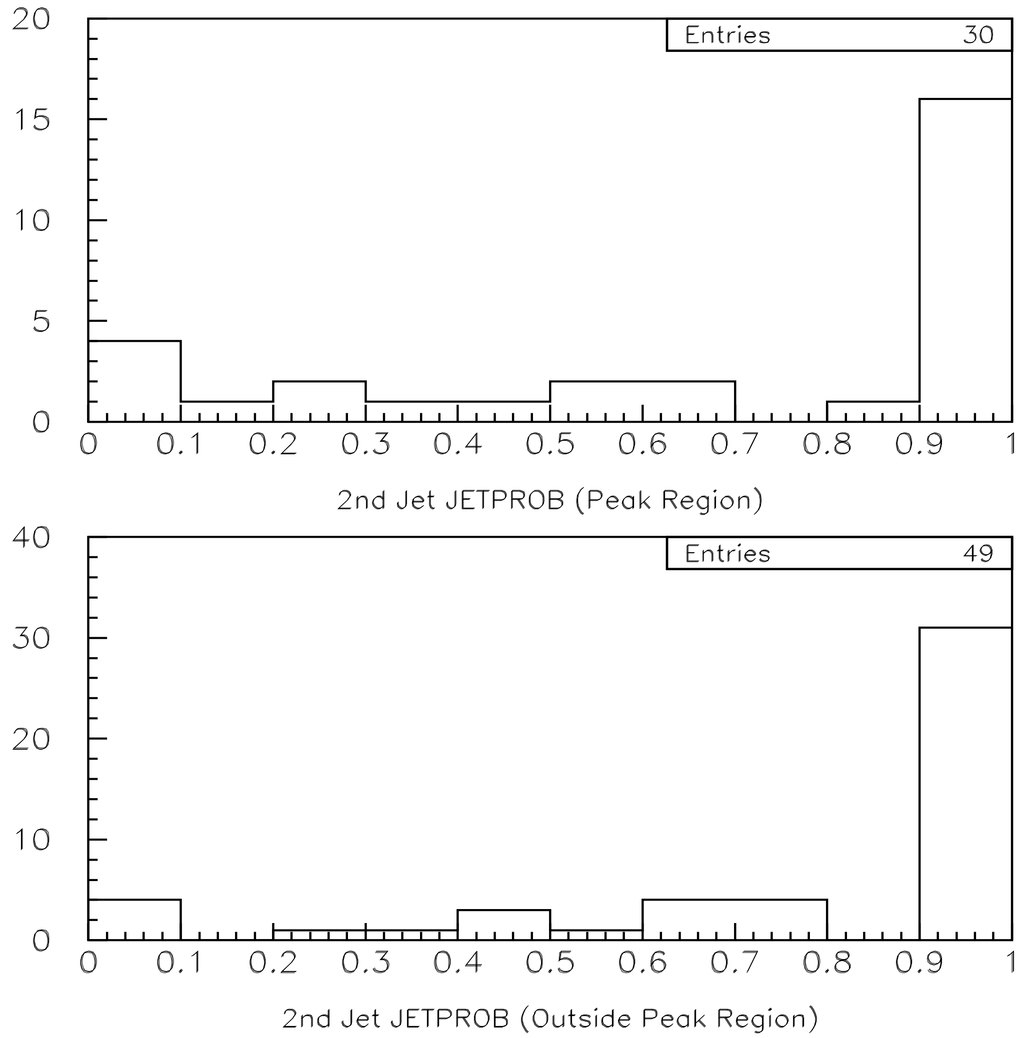


Figure 33: The JETPROB distribution for the second (leading, non-tagged) jet in data events. The upper histogram is for events in the signal region (105 ± 15) and the lower is for events outside the signal region.

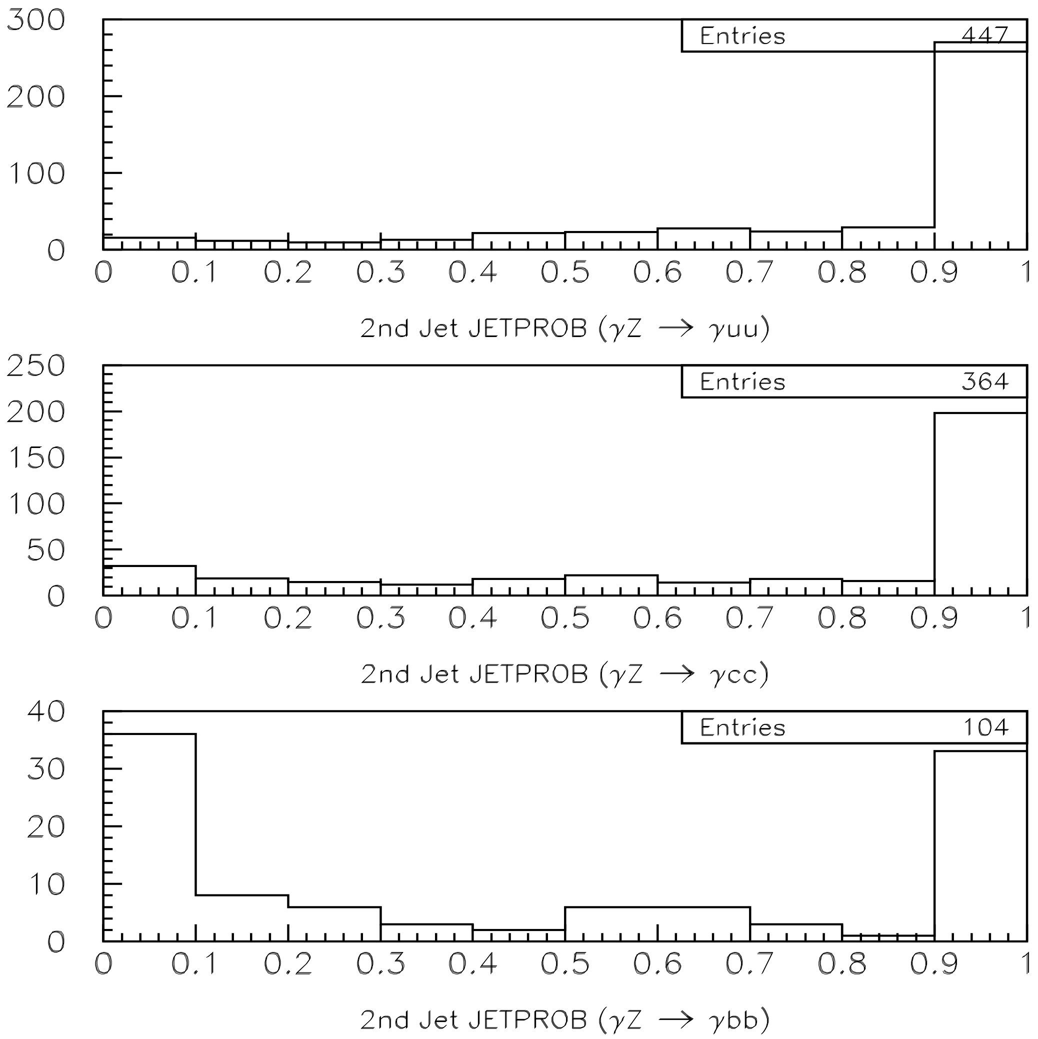


Figure 34: The JETPROB distribution for the second jet in Monte Carlo events. For the $b\bar{b}$ case we use actual tagged events and plot the JETPROB for the next leading jet. For the $u\bar{u}$ and $c\bar{c}$ events, we chose the jet with the higher SECVTX background probability as the "tagged" jet

References

[1] W. Yao, *et al*, *Top Search in Lepton+Jets with SECVTX*, CDF/ANAL/TOP/CDFR/-2989.

[2] Btag group, *Update of SECVTX Tagging in 1B data*, CDF/DOC/SEC_VTX/CDFR/-3178.