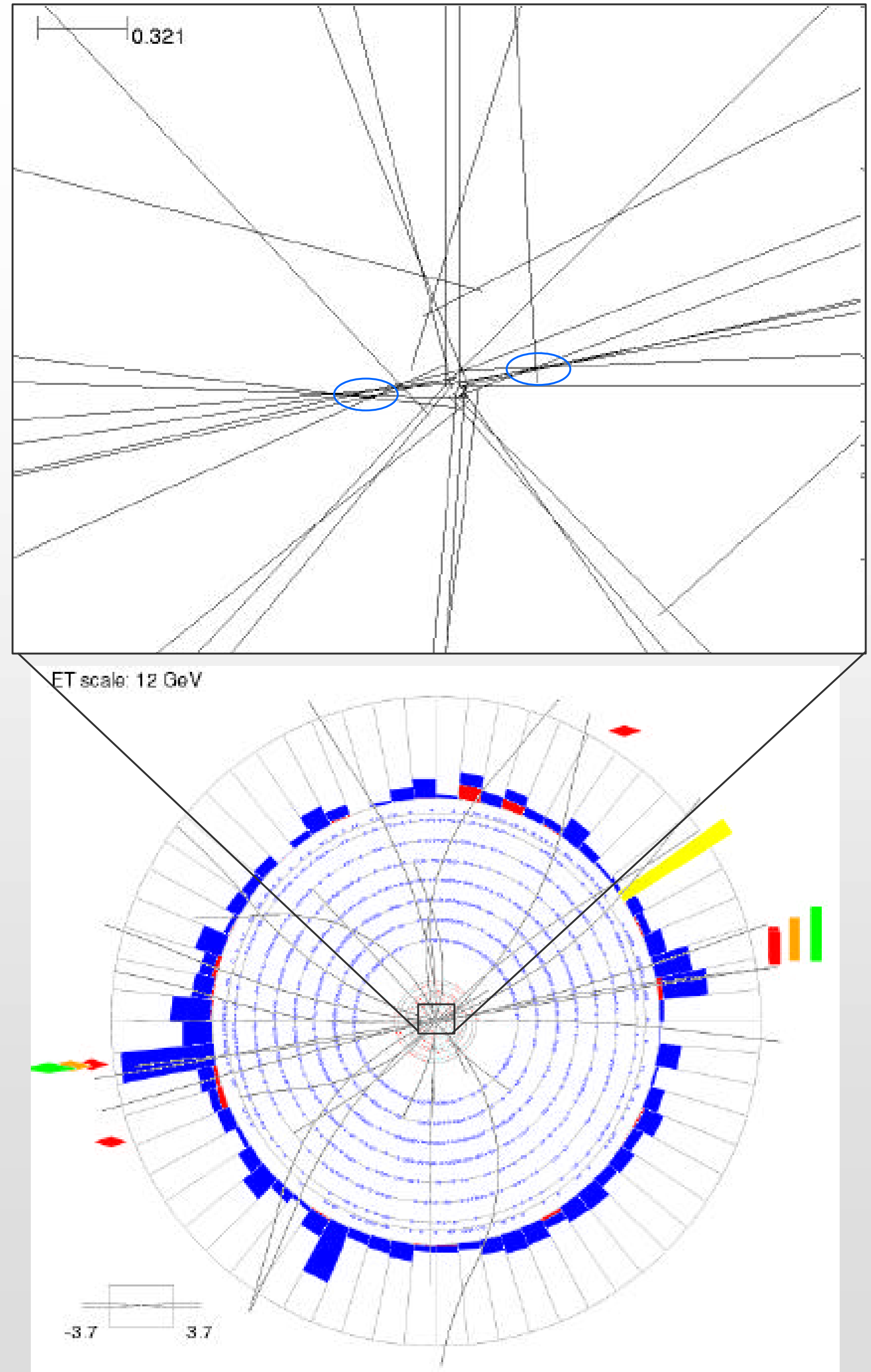
Z -> bb

Andy Haas
University of Washington

All DZero Meeting
May 23, 2003

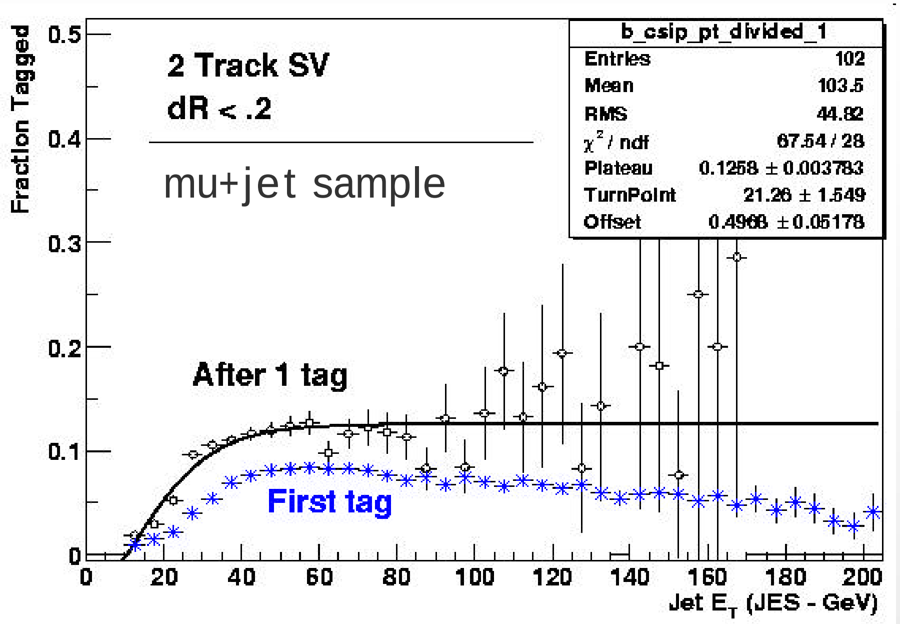


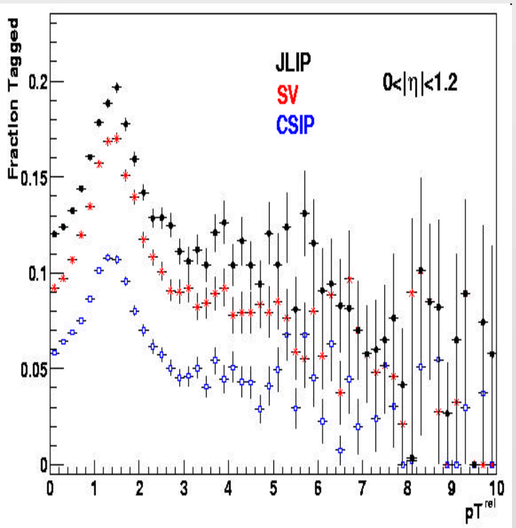
Outline

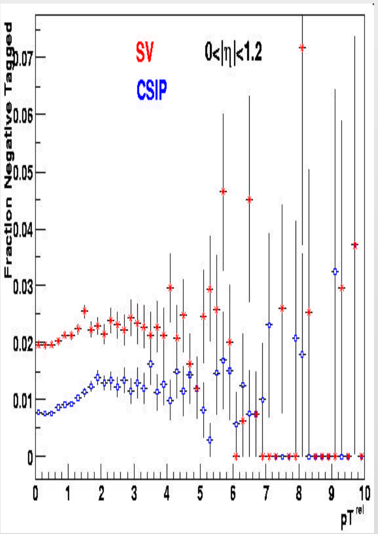
- Brief summary of data processing and b-tagging in data
- Z->bb background in the Higgs multi-jet analysis
  - Qualitative signs in the multi-jet data sample
- The importance of Z->bb for RunII physics
- Some facts, history, and previous studies / analyses
- Quantifying the signal in muon-jet events
  - Details of the data sample / discussion on triggering
  - Various methods of background subtraction
  - Comparison with Monte Carlo
- Testing the signal by varying cuts, jet corrections, triggers, b-taggers, and fitting methods

Data Processing / b-Tagging

- Process skimmed TMB datasets and TMB's from Monte Carlo
  - p13.08.00
  - higgs_multijet framework package
  - Custom rootuple (~400b/event)
  - b-tagging from TMB using latest d0root_analysis code (for primary and secondary vertices)
- Study various taggers
  - JLIP, CSIP, and SV
- Study the *kinematics* of b-tagging
  - vs. E_T and eta
  - Try to discriminate between different turn-on's for fake tags and real tags
- Look at S/B and efficiency of various tagging cuts



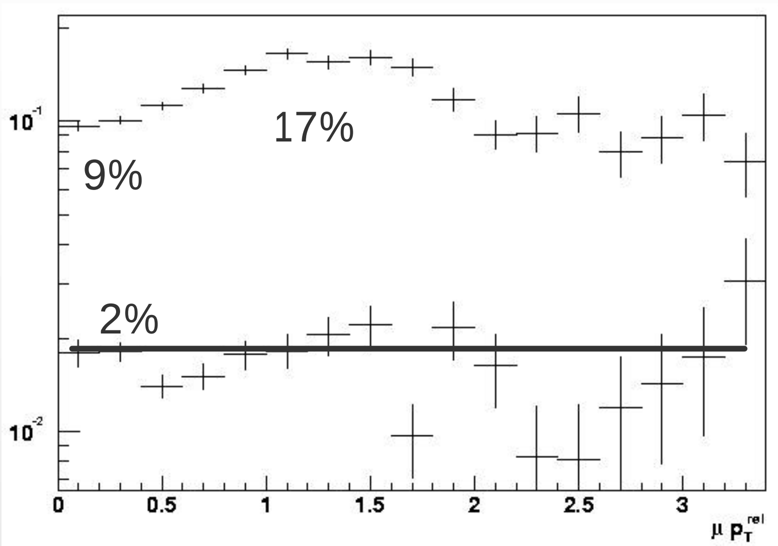




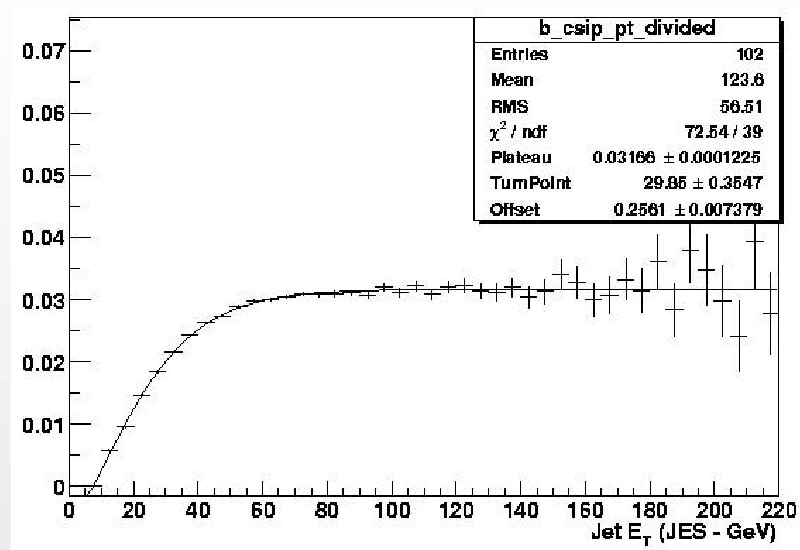
b-Tagging Tight / Loose

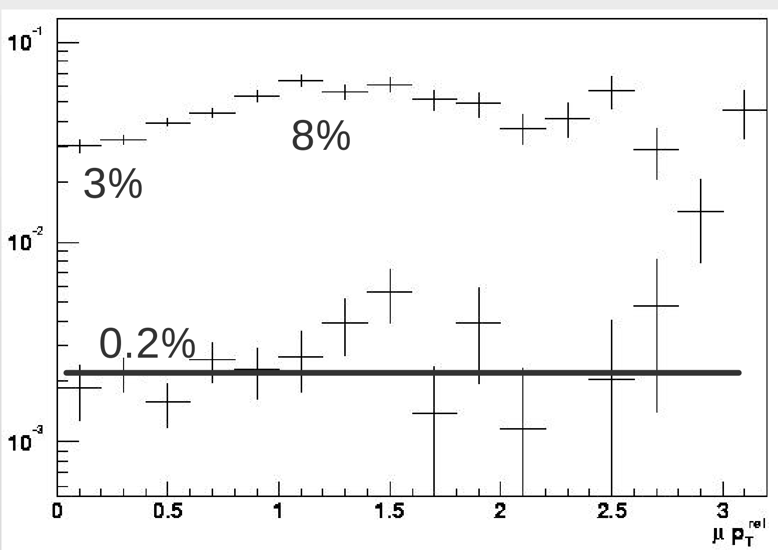
Muon-tagged Jets (p_T>40 GeV, |eta|<1.2)

All Jets (|eta|<1.2)

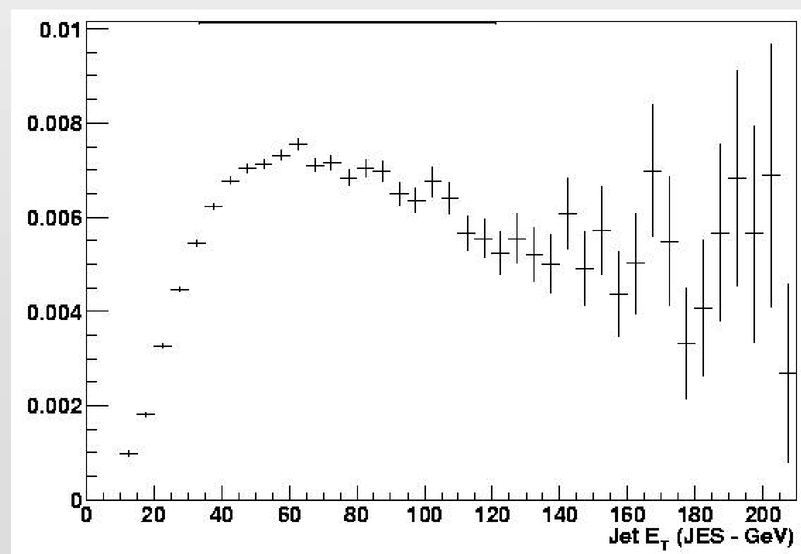


2 Track SV



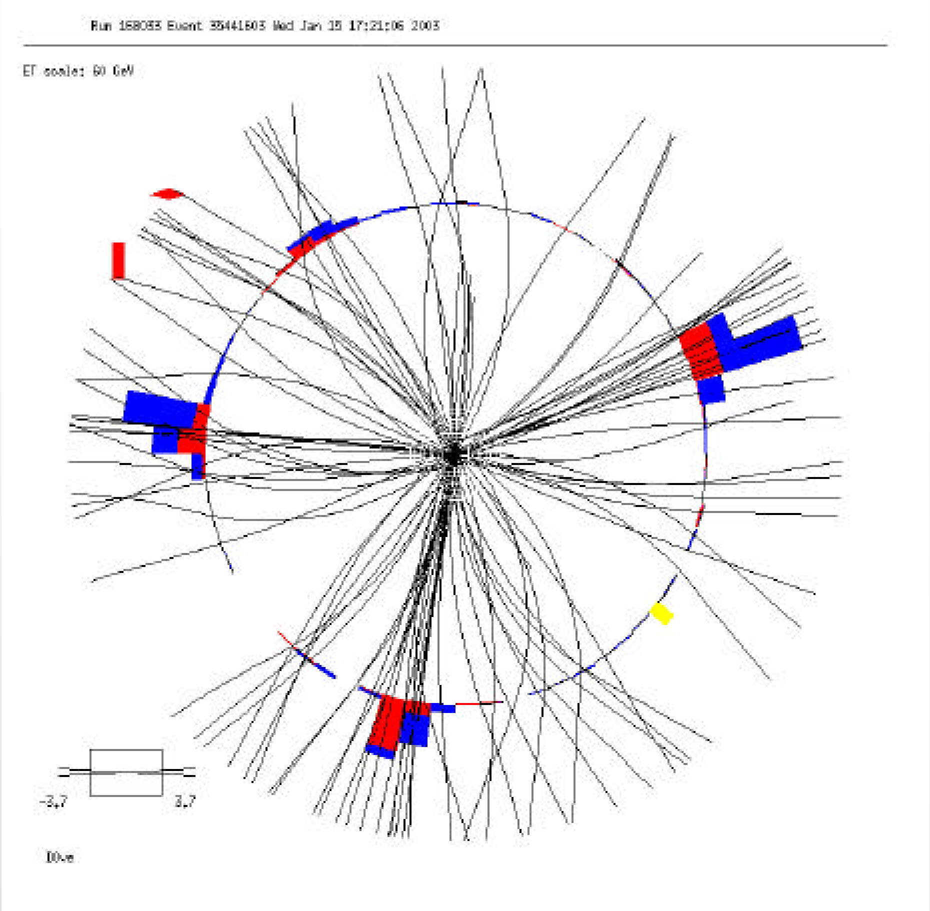


3 Track SV



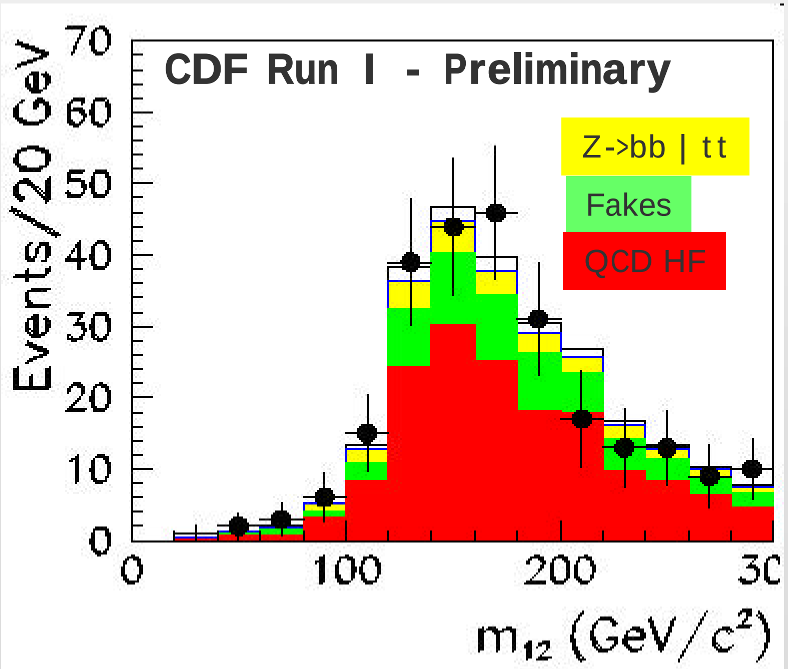
Multi-jet Data Sample

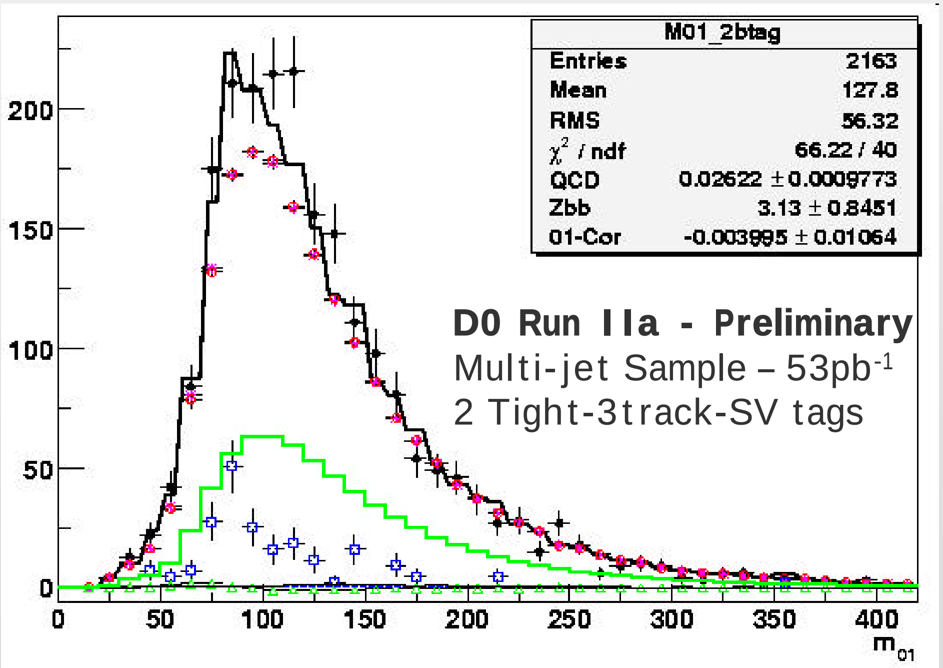
- Trigger:
  - CJT(4,5)
  - L2J(3,8), L2HT(50)
  - L3J(3,15)
- Skimmed:
  - Recoed with p13.04-06
  - 3 JCCB jets > 15 GeV
  - 1 JCCB jet > 20 GeV
  - 53 pb^{-1} and 11.3 million events



Multi-jet m_{01}

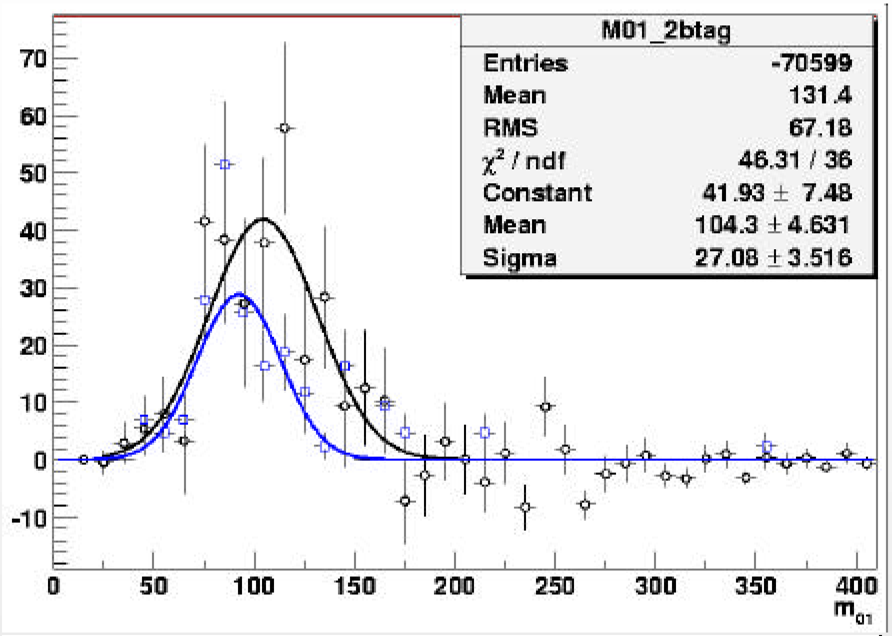
- Fit the double-tagged invariant mass spectrum to a sum of backgrounds:
  - QCD – Heavy Flavor
    - Shape estimated from b-tagging kinematic corrections to the single-tagged sample
    - Normalization left a free parameter (bbjj cross-section poorly known…)
  - Fake double-tag fraction estimated from fake-tag rates and kinematics
  - Z->bb and tt backgrounds are included from Monte Carlo normalized to the relatively well-known cross-sections
- Know to expect Z->bb background from CDF Run I analysis
- We verify the qualitative observation of Z->bb background

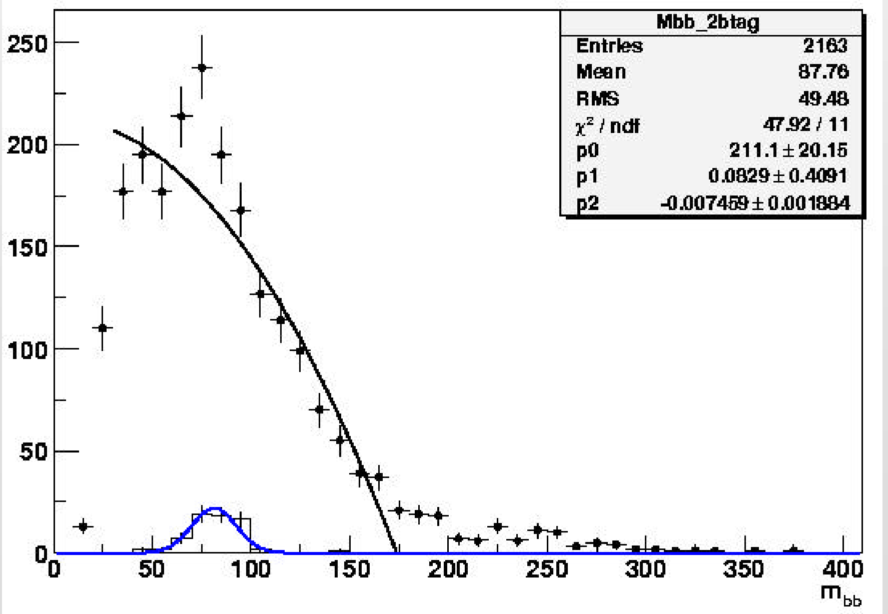




QCD HF Subtraction

- Z -> bb peak is very wide
  - Not always plotting the invariant mass of the jets which came from the Z





- Maybe clearer to plot the invariant mass of the leading two b-tagged jets
- See evidence of a peak corresponding to where we expect Z->bb to be
- Background is much more difficult to estimate (use a parameterization)
  - High-statistics bb+jet Monte Carlo would be necessary

Why do we need Z->bb ?

The Z->bb peak is a 'standard candle' for b-jets

- Measure b-jet energy scale
  - Reduce final Run IIa uncertainty on m_t from 2.7 -> 1.6 GeV
- Measure b-jet energy resolution
  - Crucial for placing limits on Higgs production (with H->bb)
- Measure b-jet tagging efficiency, fake rate
  - And kinematic dependence… vs. p_T , vs. eta…

Observing and measuring the Z->bb decay proves that tools used to look for the Higgs really work!

Z->bb Facts and Previous Studies

- Cross-sections:
  - QCD, pT>20 GeV = 35,000,000 nb
  - bb, pT>20 GeV = 100,000 nb
  - Z->bb (4.594 x Z->ee) = 1180 nb
- DZero had a hint of a signal during Run I
  - Lack of silicon b-tagging led to low efficiency…
  - Word of mouth only…
- CDF saw 3.23 sigma, 91+-30 +-19 events, in 103+-7 pb^{-1}
  - Trigger: 7.5 GeV pT central muon
  - hep-ex/9806022 – never published…
- Run IIa DZero Z->bb Observability Study
  - Assumed STT trigger
  - D0Note 3604

CDF Run I

- Trigger on
  7.5 GeV pT central muon
- No match to jet (dR)
- Require exactly 2 jets
- 2 SVX tags
- d-phi > 3.0
- Sum of non-jet E_T < 10 GeV

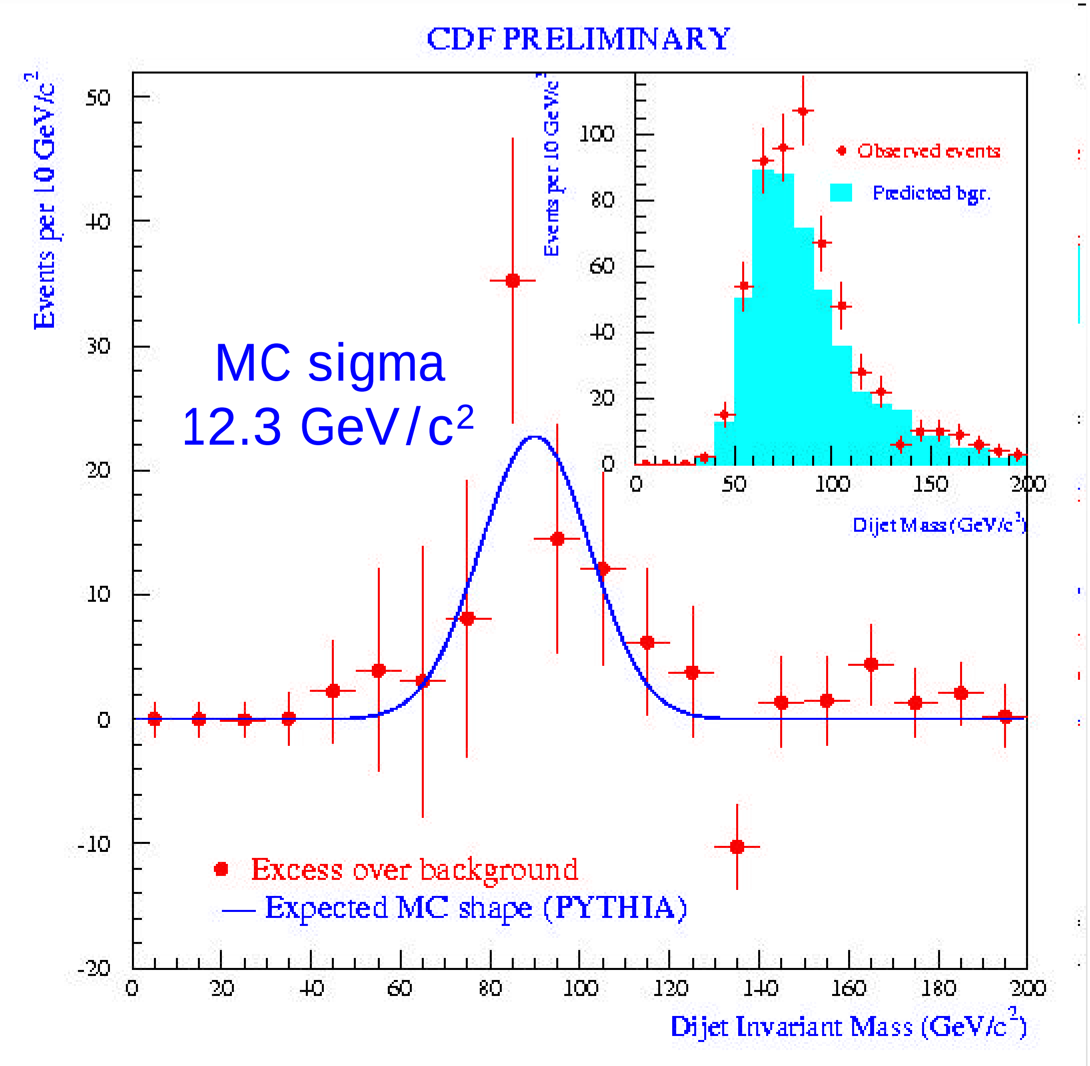
- Extrapolate 1-tagged distribution from non-signal region (m<70 and m>110 GeV/c^2) into signal region

- Z->cc and W->cs found to be negligible

- Binned observe 69+-20 events
  Expect 41+-19 events from MC

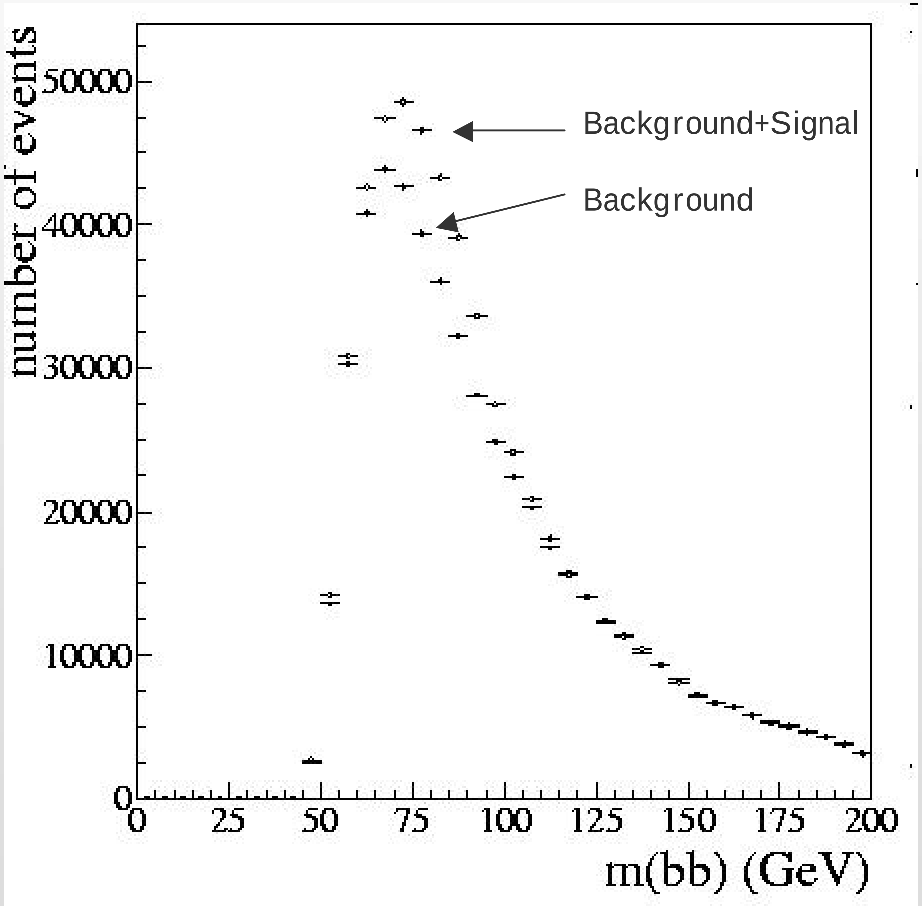
- No comparison to bb MC rates

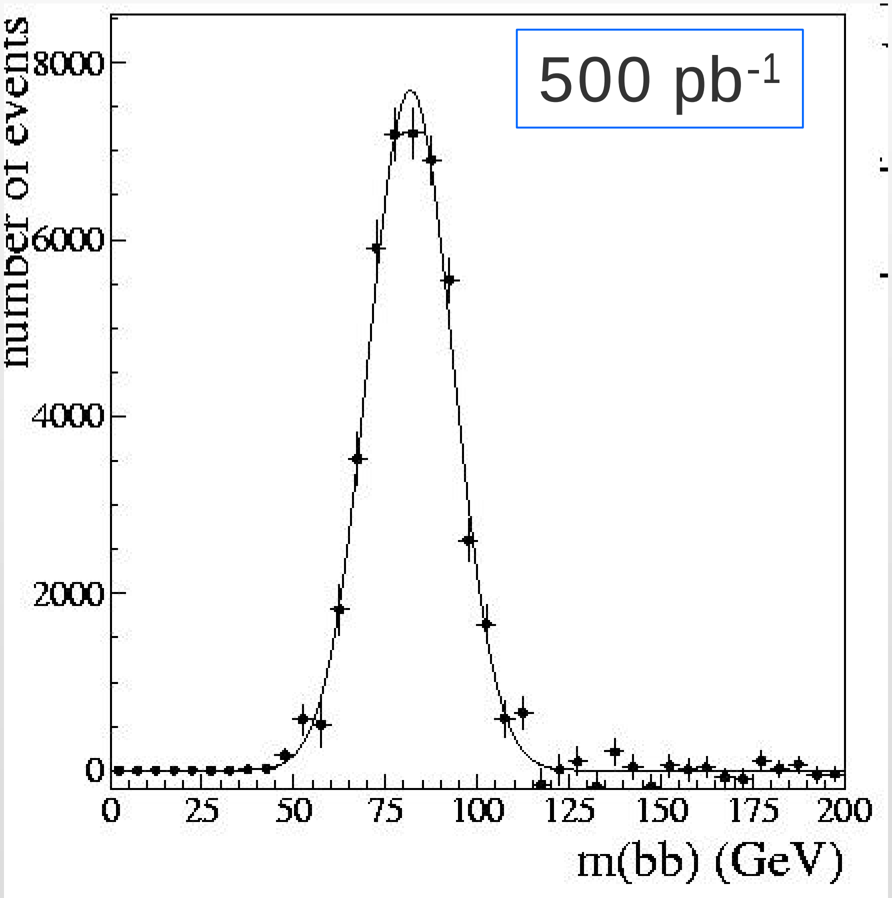
- Unbinned get 90+-30+-19 events, 3.23 sigma



DZero Run IIa STT Study

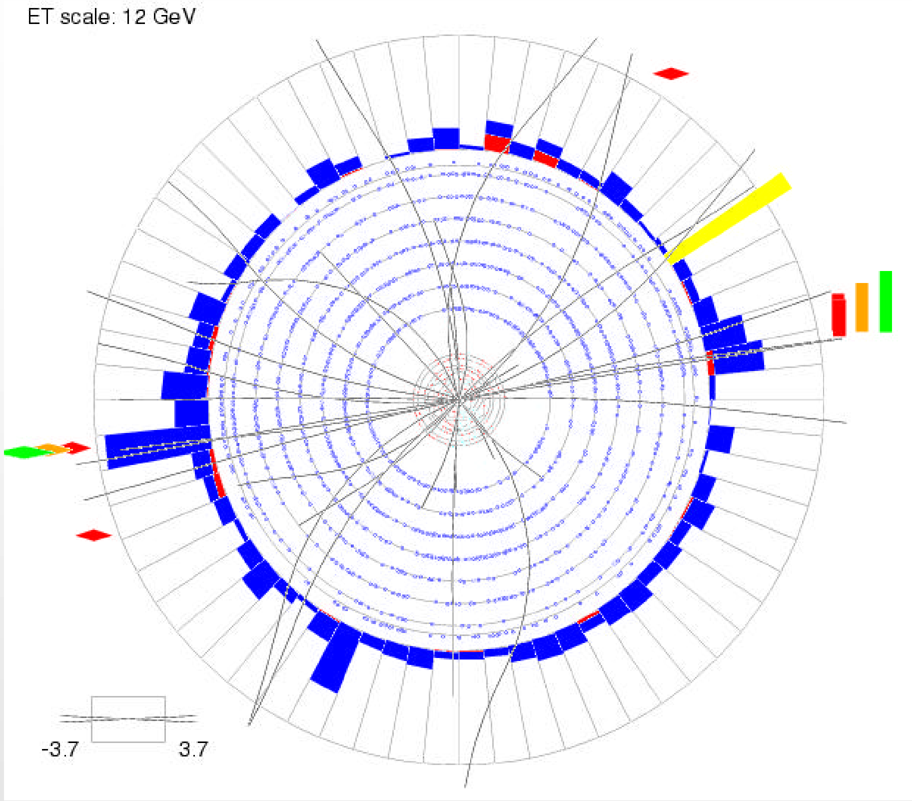
- D0 Note 3604 – M. Narain, U. Heintz
  - 50% tagging – achievable soon
  - 11 GeV/c^2 resolution - reasonable
  - 20% trigger efficiency – need the STT for this!





The Mu+jet Sample

- Skimmed by T. Golling for Top
- Requires:
  - JCCB Jet with E_T>15 GeV
  - Global muon < 0.7 dR of a jet
  - No p_T cut on muon! Excellent.
  - ~50 pb^{-1} and 1.2 million events



Triggering

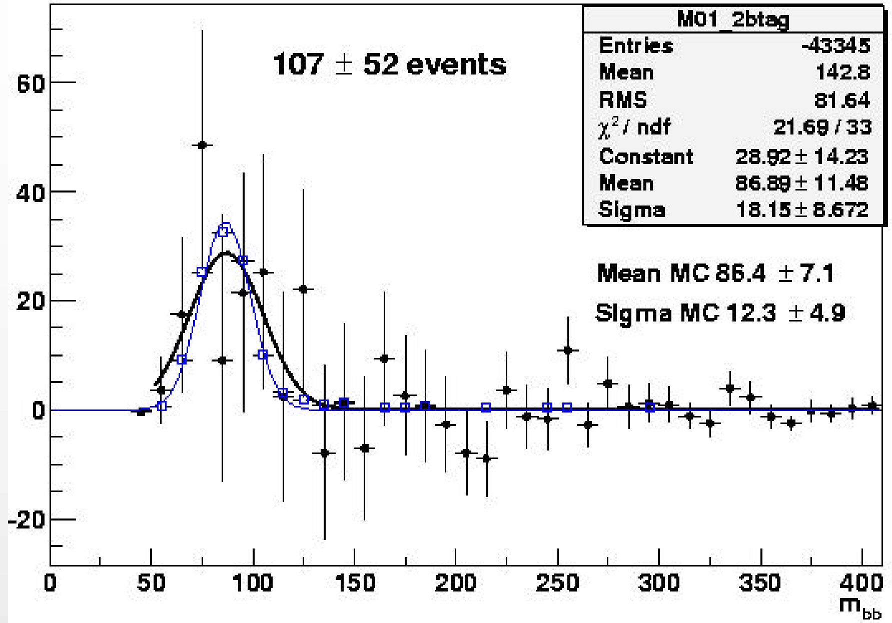
- Was a good trigger for Z->bb in place?
  - Not bad… but we could do better… even before the STT is available
- Want high b-jet content, no E_T bias !
- Most data passing cuts (2 jets > 25 GeV, with full JES) fired
  - MU_2TRK3_L2M0 (68%)
  - MU_JT20_L2M0 (77%)
- The 20 GeV jet trigger has large bias towards high jet E_T
- The 2-track trigger was prescaled due to low rejection

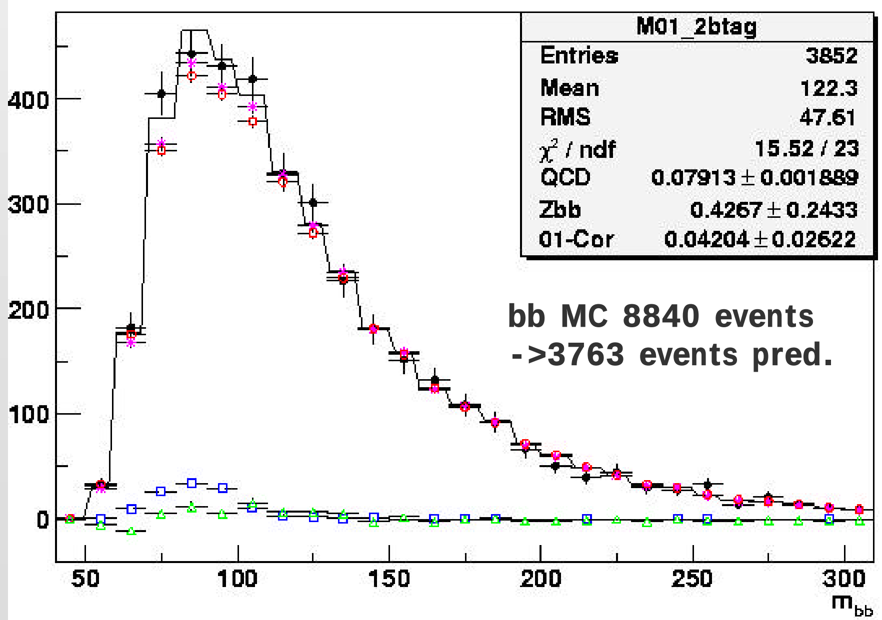
- Level 1
  - Clearly want a muon
  - Possibly match A-stub to CTT track?
- Level 2
  - Clean up muon
- Level 3
  - Very low (10 GeV) jet requirement
  - Loose b-tagging AND/OR Match jet to muon

```
20428 times (6.207798 percent) for trigger  -> 2MU_A_L2M0
22620 times (6.873917 percent) for trigger  -> 4JT10
22862 times (6.947458 percent) for trigger  -> MU_W_L2M0_TRK3
22898 times (6.958398 percent) for trigger  -> 2MU_C_2L2_TRKLO
30719 times (9.335096 percent) for trigger  -> MU_TAU10_L2L0
35975 times (10.932324 percent) for trigger  -> MU_W_L2M0_TRK10
39478 times (11.996840 percent) for trigger  -> MU_A_EM10
45257 times (13.753001 percent) for trigger  -> MU_EM_L2M0_L3L0
53137 times (16.147628 percent) for trigger  -> MU_JT20_L2M0_BID
56375 times (17.131613 percent) for trigger  -> MU_W_L2M5_TRK10
70604 times (21.455617 percent) for trigger  -> MU_TAU10_2T_L2M0
73401 times (22.305588 percent) for trigger  -> mu1ptxatxx_CEM5
79366 times (24.118273 percent) for trigger  -> MU_W_L2M0_2TRK3
98745 times (30.007293 percent) for trigger  -> MU_EM_L2M5
101725 times (30.912876 percent) for trigger  -> MU_JT25_L2M0
117771 times (35.789042 percent) for trigger  -> MU_EM_L2M0
189888 times (57.704440 percent) for trigger  -> MU_TAU10_L2M0
222468 times (67.605069 percent) for trigger  -> MU_2TRK3_L2M0
252616 times (76.766645 percent) for trigger  -> MU_JT20_L2M0
```

Evidence in mu+jet data

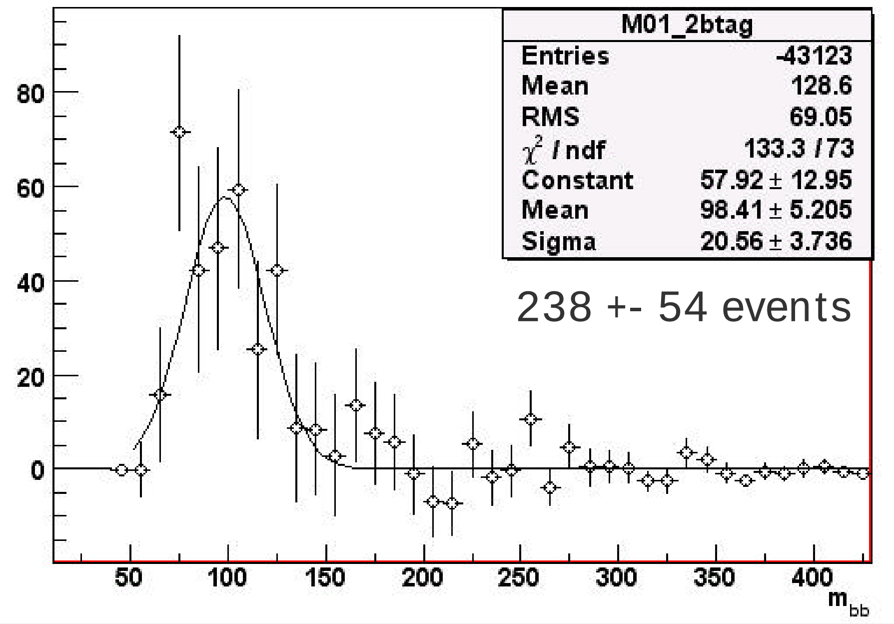
- p_T > 25 GeV, full JES corrections
- |eta| of jets < 2.0 (fiducial tags)
- d-phi > 2.5 (loose back-to-back)
- Loose 2-track SV b-tagging
- Fit double-tagged sample (3852 events) to background + signal
- Subtract background and fit to Gaus.
- Compare to Z->bb Monte Carlo
- Use the CDF Run I method:
  - Estimate background from the single-tagged sample
  - Also correct for E_T turn-on of the second b-tag (using 0-1 tag method: explained later…)
- Compare with bb Monte Carlo
  - Pythia direct bb MC (26,000 evts)
    - p_T>20 GeV, |eta| < 3
  - Same trigger/skimming efficiency for bb as Z->bb, ~40%
  - Shape and position of peak agrees
  - Peak in data is wider by ~50%

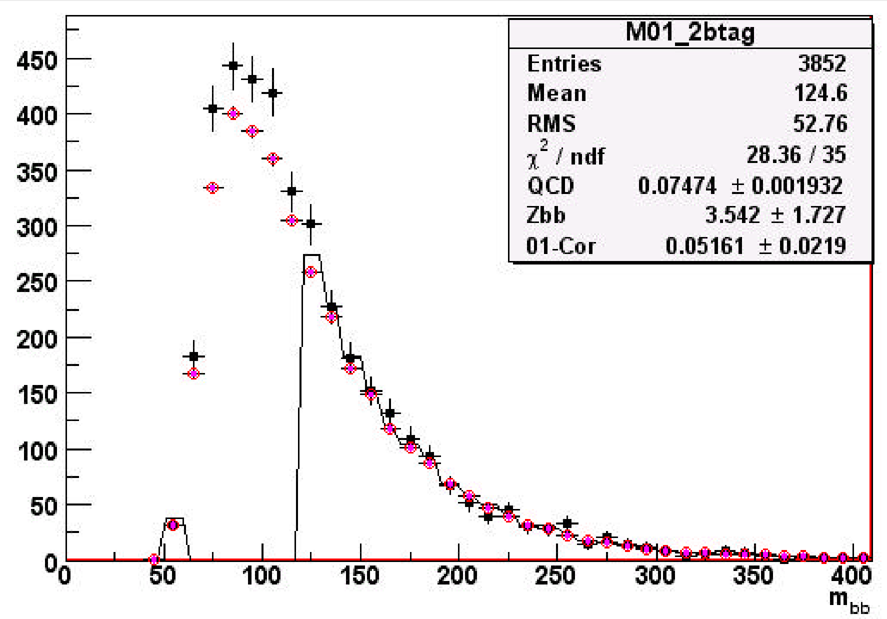




Background Estimation Methods

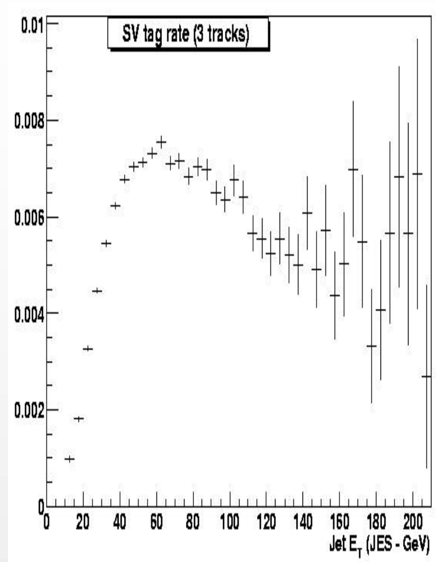
- Z->bb peak is very close to the peak of background
  - Great care must be taken to determine the *shape* of the background!
- Simplest method:
  - Used in the CDF Run I pre-print and assumed valid in D0 Run II MC study
  - Assume shapes of 1-tagged sample and 2-tagged sample are the same outside the 'signal region'
  - Normalize 1-tagged distribution to the 2-tagged distribution outside the 'signal region'
  - Might not be bad if di-jet mass resolution is small, so 'signal region' is small
  - Ignores bias in shape due to turn-on of 2nd b-tag requirement
  - Gives a result with ~2 times as many Z->bb events as expected

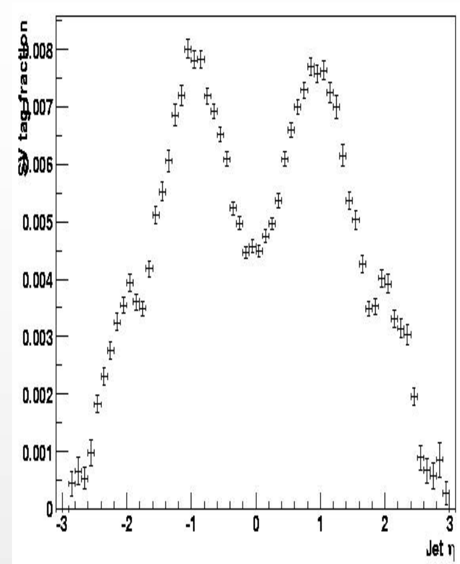


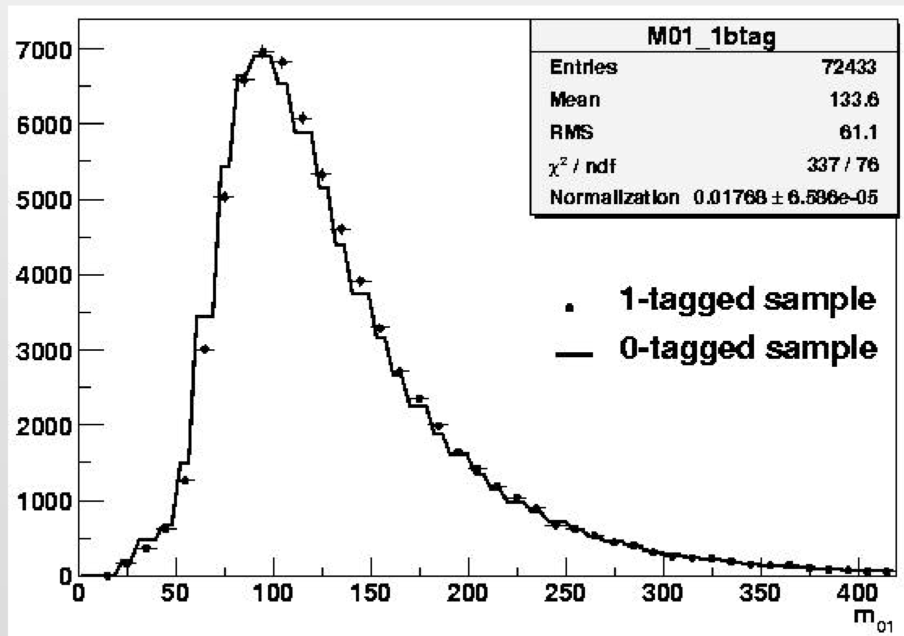


Correcting for b-tagging bias

- Try to account for b-tagging efficiency and fake-rate as function of E_T and eta
  - This will certainly change the shape of the 2-tagged invariant mass spectrum
- Results are very similar when:
  - Apply the correction to the 1-tagged sample
  - Apply the correction *twice* to the 0-tagged sample
- Another method is to directly observe the change in shape from the 0-tagged sample to the 1-tagged sample
  - Then apply the same correction to the 2-tagged sample

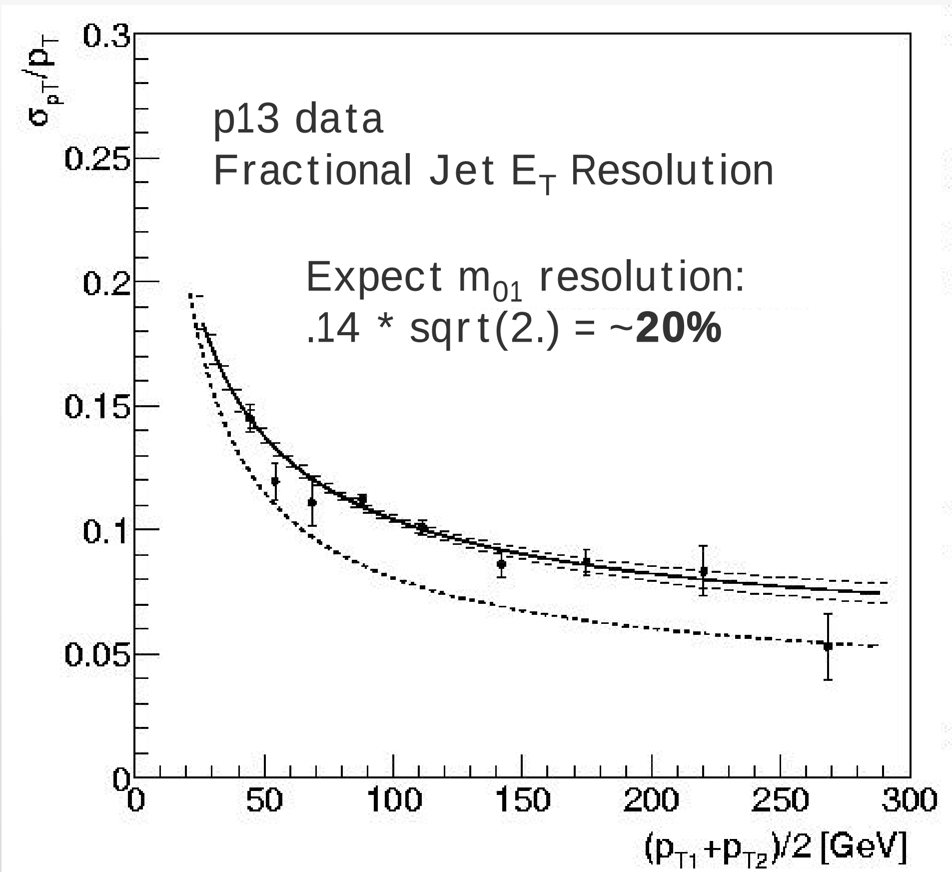


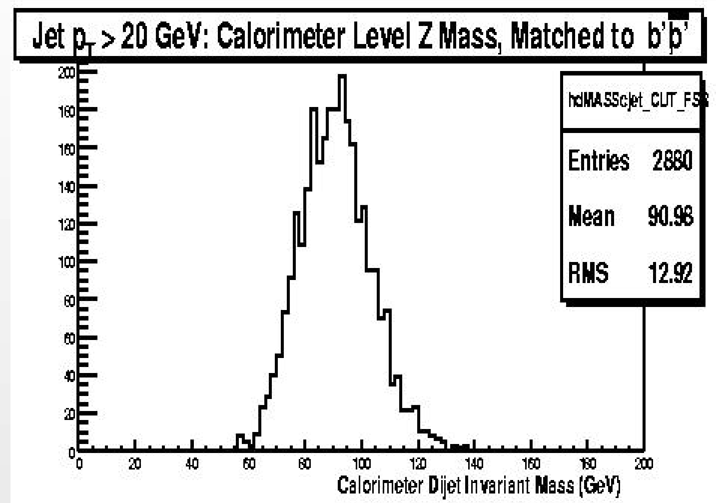


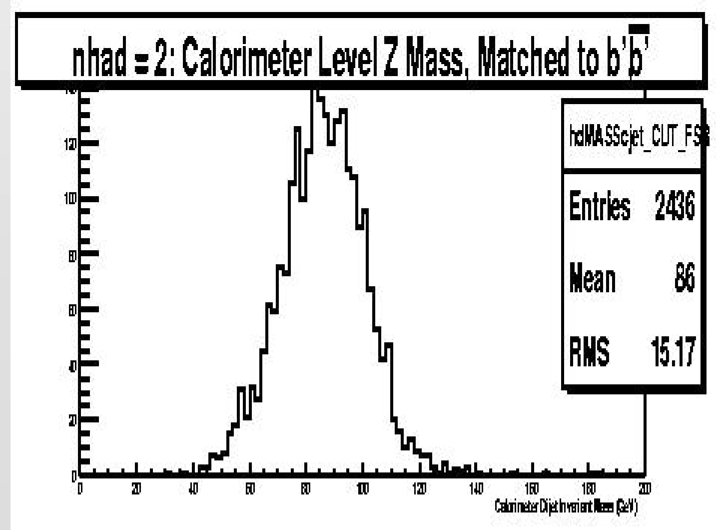


Di-jet Resolution

- D0 Note 4136 – A. Jenkins, A. Goussiou
  - Nice MC study of dependence of resolution of various parameters
- JES Group
  - Certified resolution in p13 data





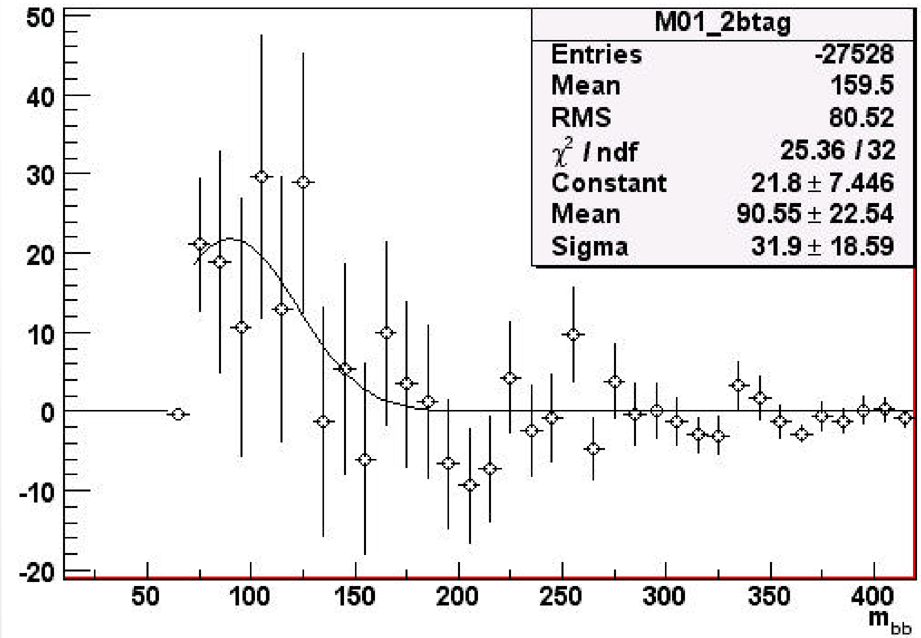


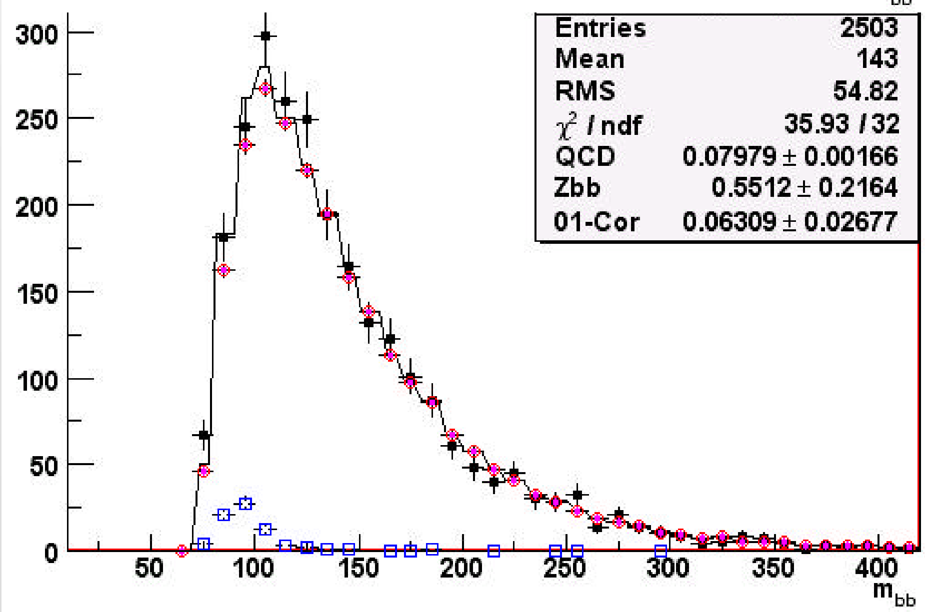
Tests of the Signal

- JES
  - Full corrections
  - No muon corrections
  - No corrections
- Vary E_T, d-phi, and eta cuts
- Triggers
  - Select only well-understood trigger
  - Try ‘fake trigger’ turn-on
- Try other b-taggers
  - And other b-tagger cuts and parameters
- Leave MC out of fit – compare Chi^2 Probability
- Shift Z mass and observe data’s peak behavior

Tighter Jet E_T Cuts

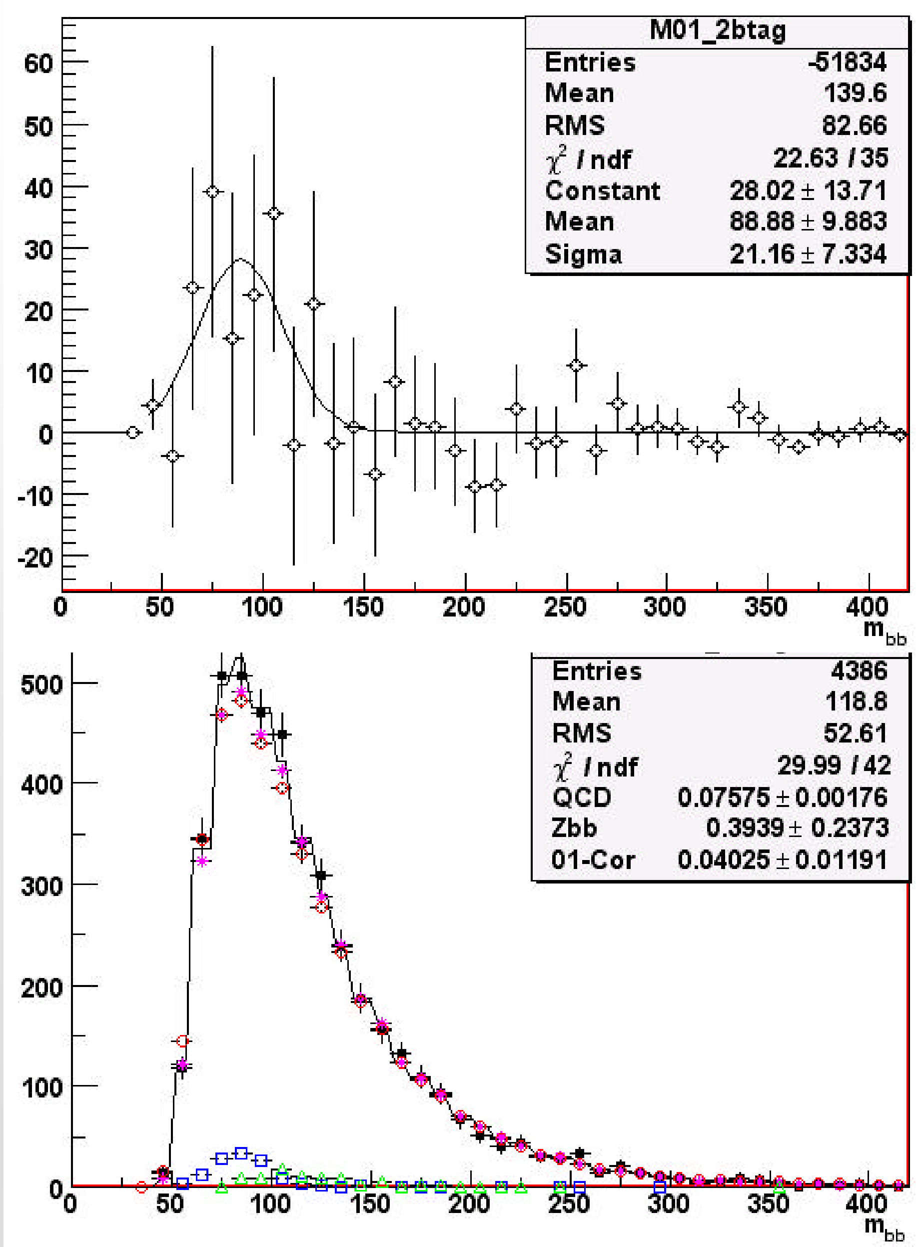
- Tighten Jet E_T cut to 35 GeV (JES, full corrections)
- Efficiency of Z->bb increases from 43% to 55%
- Still consistent with bb MC
  - Predicts 2010 bb events, based on Z->bb peak size
  - Observe 2503 events





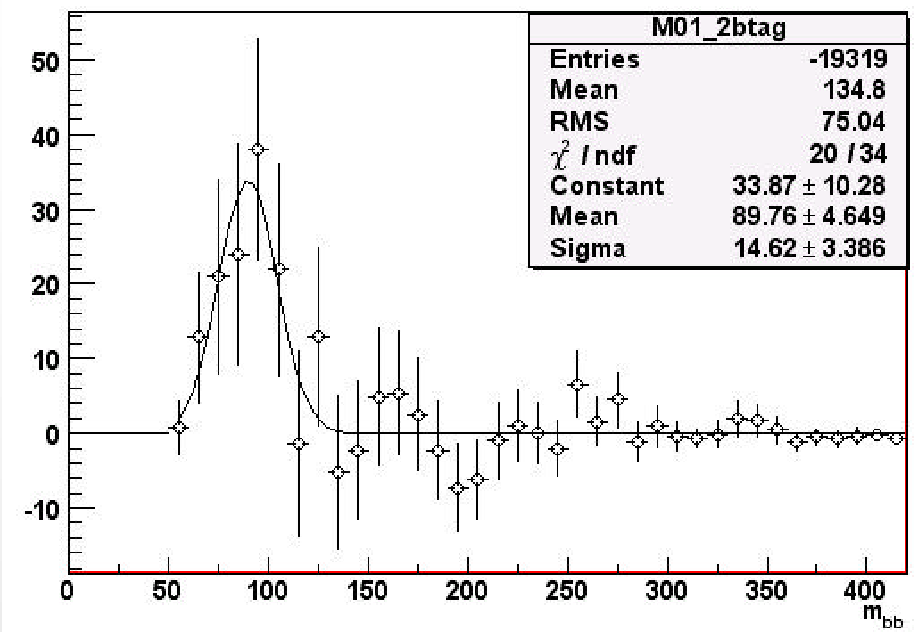
Looser Jet E_T Cuts

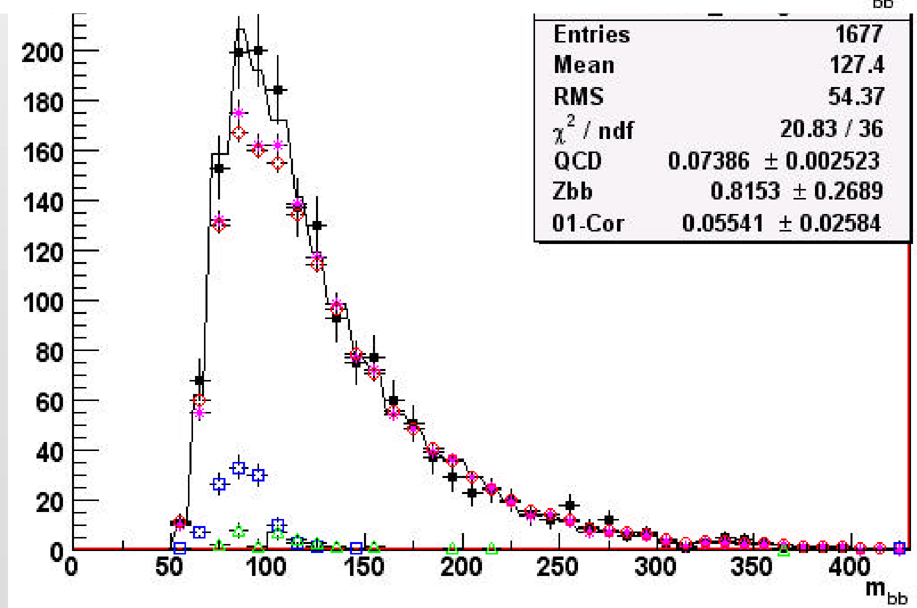
- Loosen Jet E_T cut to 15 GeV (JES, full corrections)
- Efficiency of Z->bb decreases from 43% to 39%
- Still consistent with bb MC
  - Predicts 4334 bb events, based on Z->bb peak size
  - Observe 4386 events



Tighter Delta-Phi Cuts

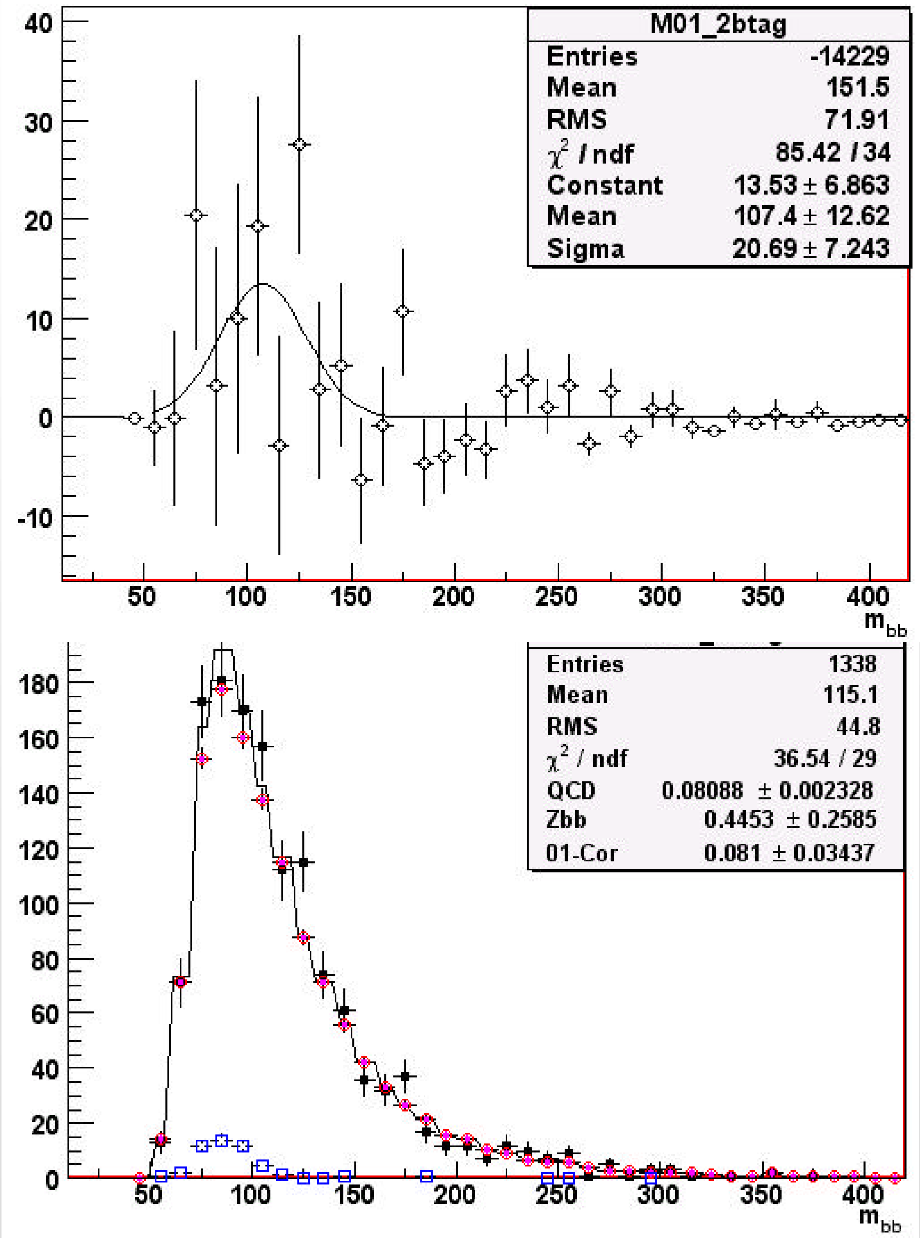
- About the same amount of signal, but background is reduced
- bb MC predicts 3389 events x .81 from Z->bb is 2745 events
  - We observe only 1677 events
  - ~60% too much signal
  - Within errors / statistics ?





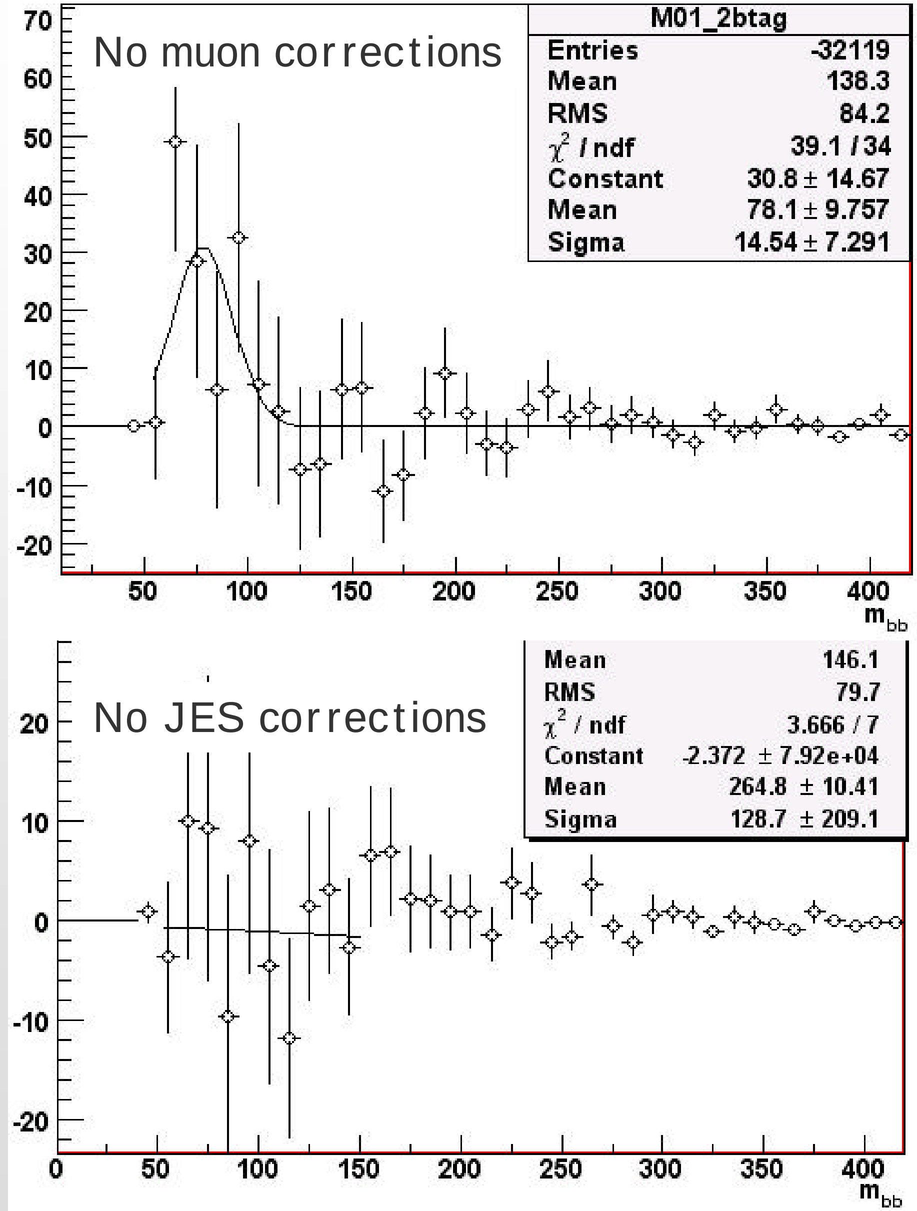
Tighter Jet Eta Cuts

- Decrease cut on jets to |eta| < 1.0
- Signal / Background goes from 108/3852 to 54/1338 (2.8% to 4.0%)
- bb MC predicts 1459 events, based on Z->bb peak height
- Observe 1338 events



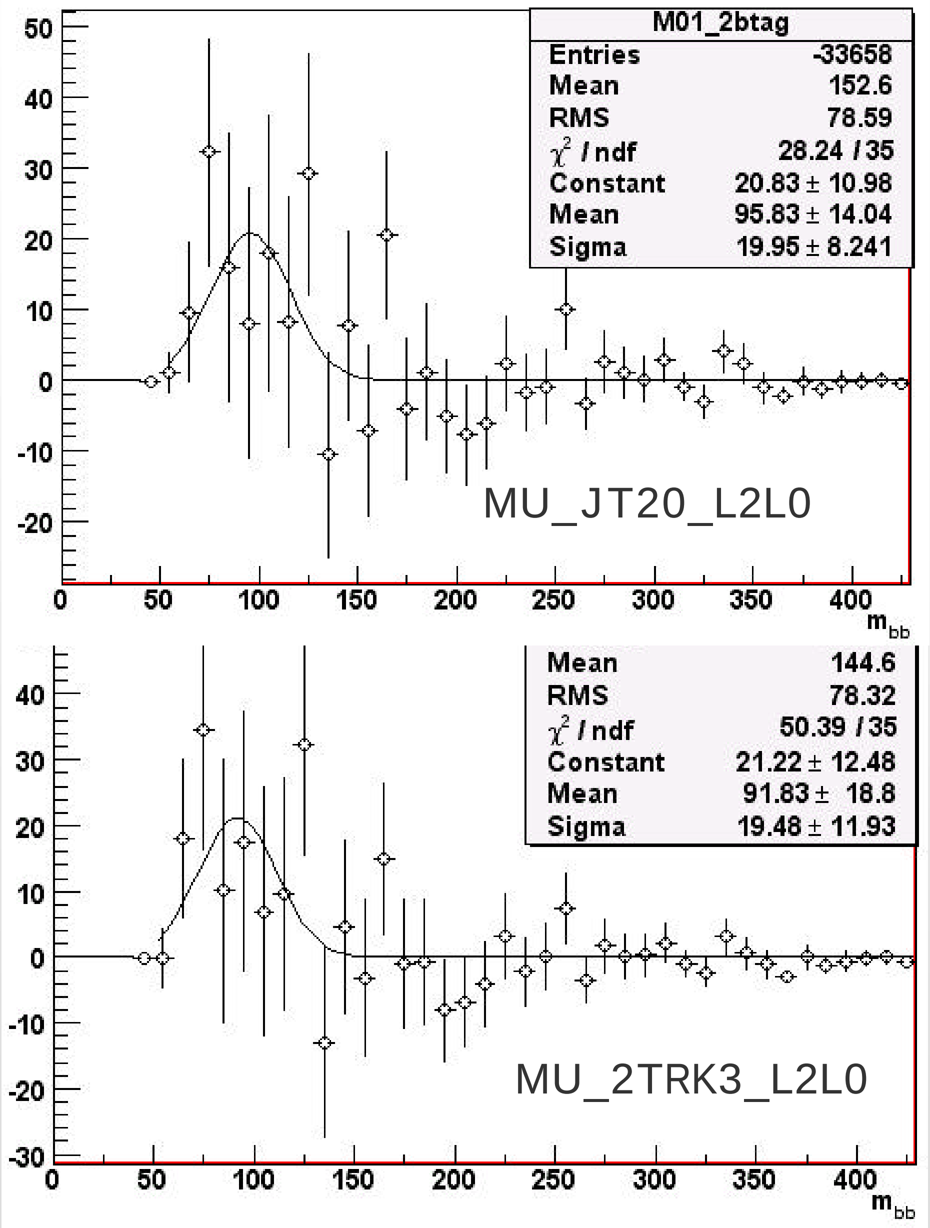
Jet Energy Corrections

- Try without JES corrections:
  - With no muon correction
    - Peak width not much different
    - Peak height not much different
    - Shape less Gaussian: Chi^2 goes from 22 -> 39
  - With no corrections at all
    - Peak is gone !



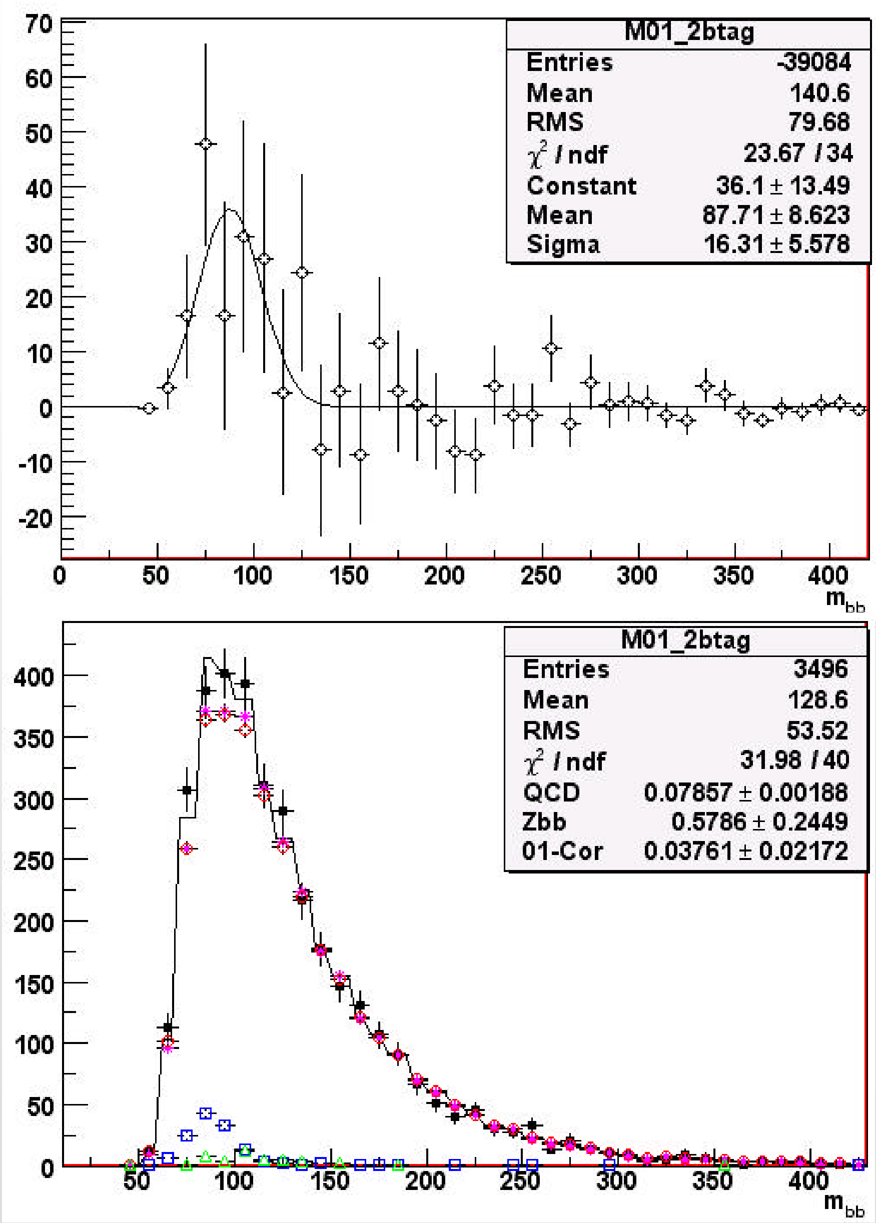
Trigger Selection

- Try just selecting single triggers:
  - MU_JT20_L2L0
  - MU_2TRK3_L2L0
- Invariant mass spectra still do not turn on lower, due to skimming cuts
- There are less events in each case (about 70%)
- But signal shapes, sizes, and positions are stable



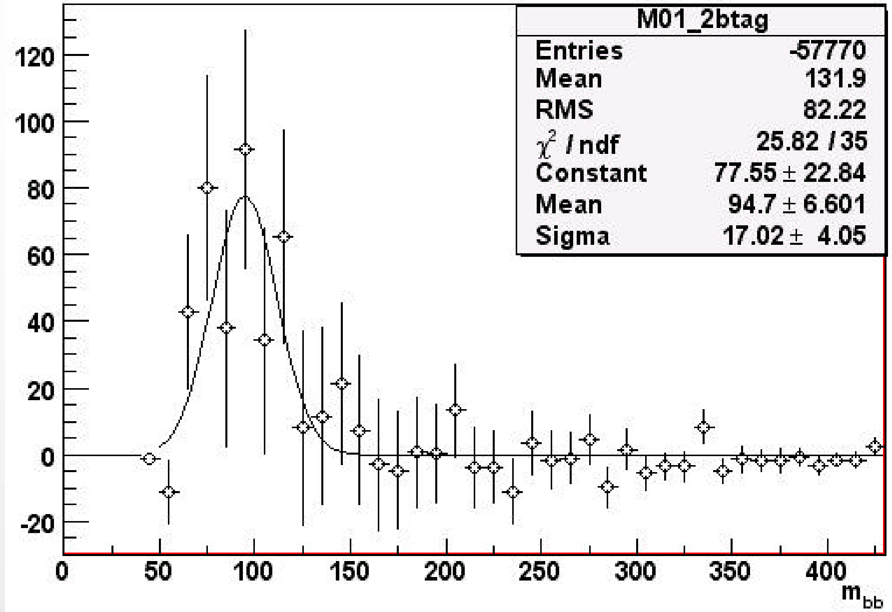
Fake Trigger Turn-on

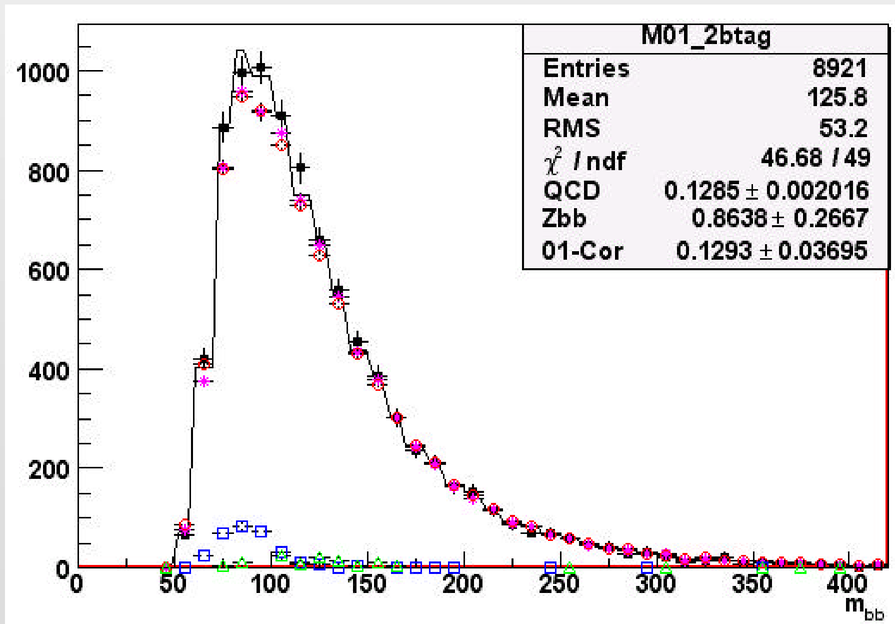
- Try adding a fake trigger turn-on artificially:
  - TanH function with a
    - mean of 20 GeV uncorrected
    - width of 5 GeV
- Harder jet E_T cut seems to select slightly more Z->bb
  - Efficiency goes to 58%
  - bb MC gives 6429 events x .58 = 3729 events
  - Observe 3496 events
- No other strange effects seen from artificial trigger turn-on



JLIP Tagger

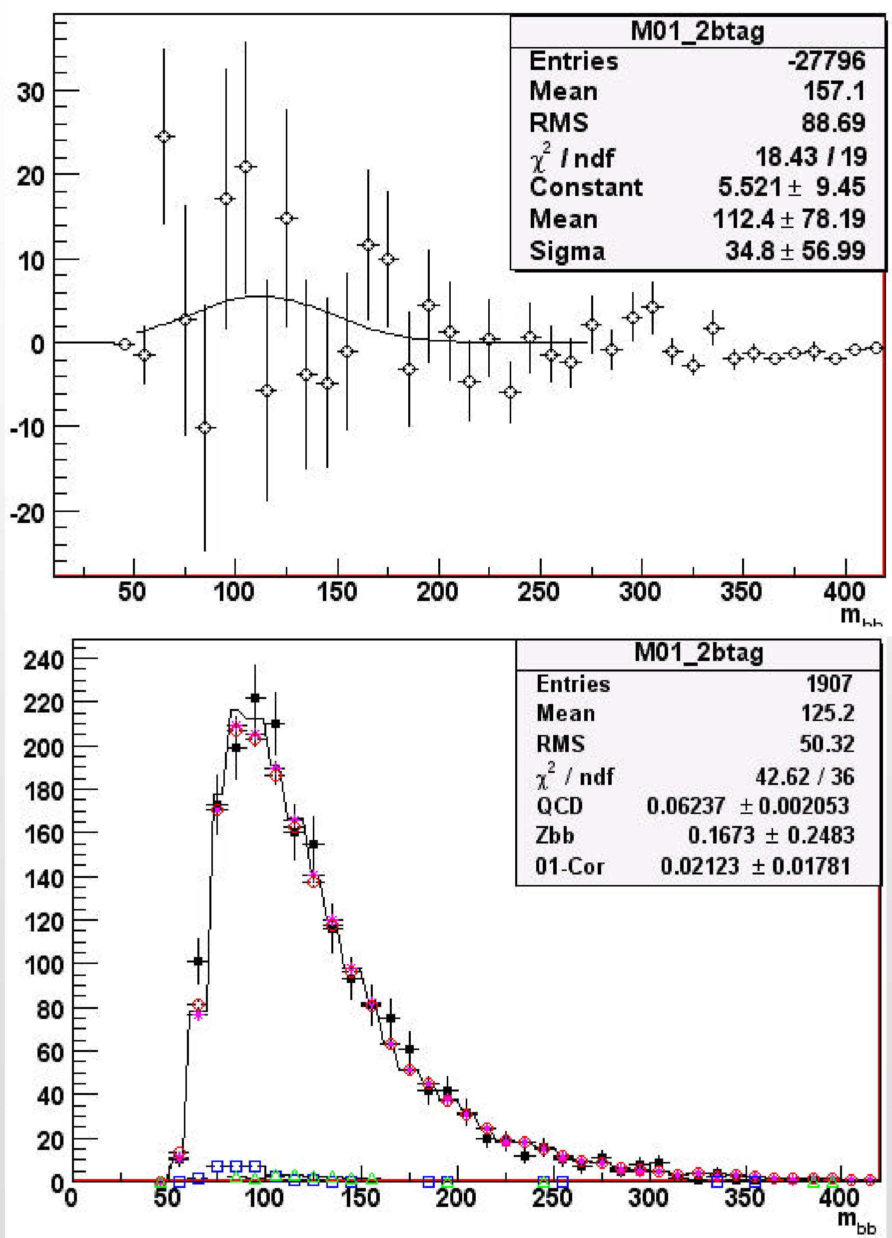
- Observe more events with the JLIP tagger
  - Known to be more efficient, but higher fake rate, the way it is used here
- Peak shape and position are consistent
- Predict 11,430 events from bb MC, and observe 8921
  - But we may be letting in more background?
- Mini-conclusion
  - A looser tagger is better, since we require 2 tags, and already have a muon tag





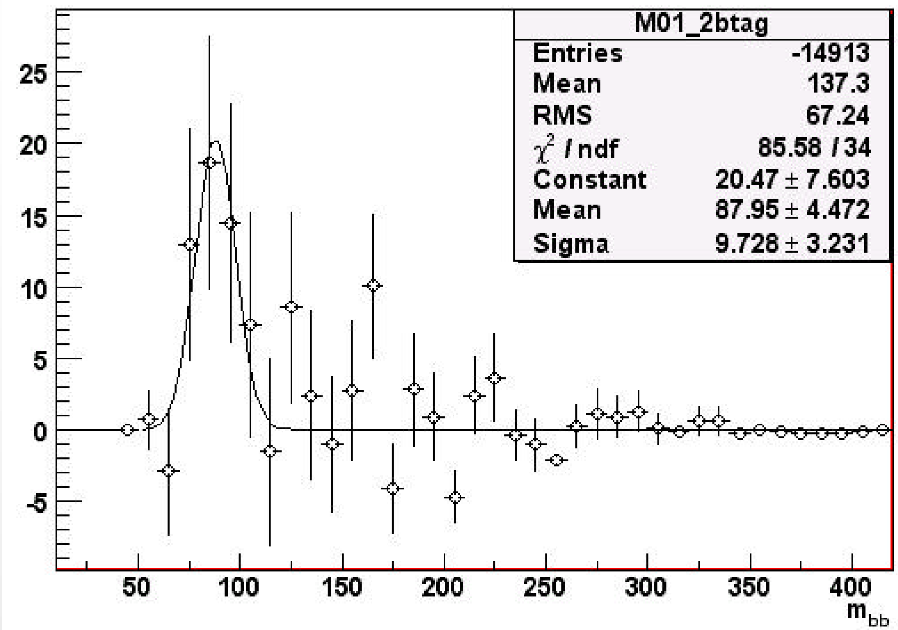
CSIP Tagger

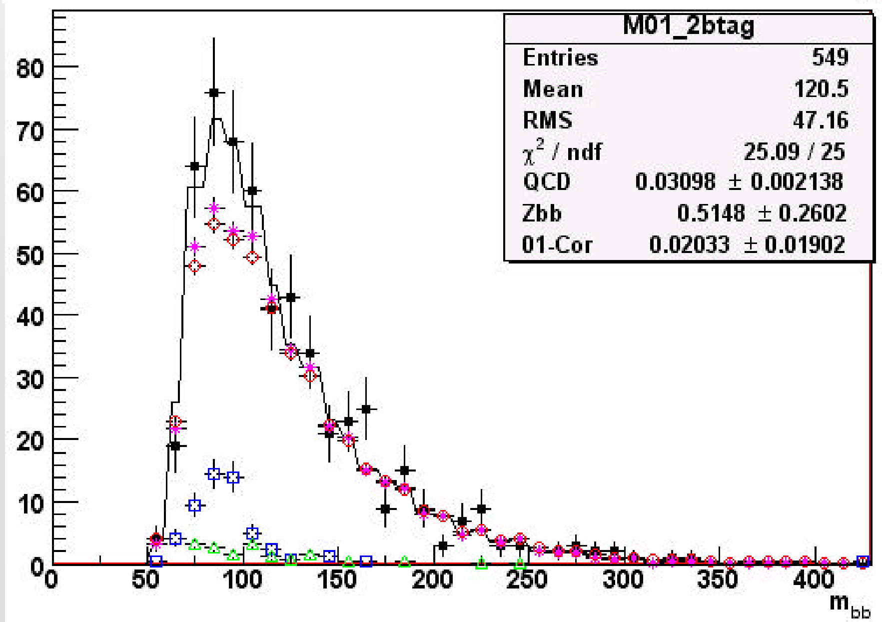
- Signal has mostly disappeared?
- There is some yet unsolved problem with my application of the CSIP tagger in the forward region



Tighter SV b-Tags

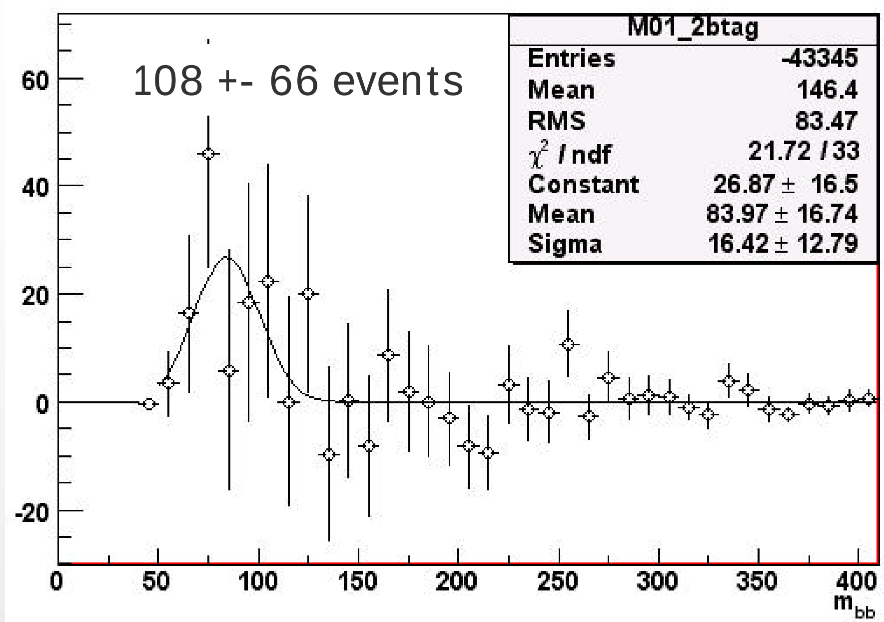
- Require 3 tracks on each Secondary Vertex b-tag
- These are very tight tags
- Signal fraction increases slightly, but we have low statistics now, especially for the peak fit, which gives a tiny width
- bb MC gives 2535 events
  - x .51 = 1293
  - About ~2 times more than observed bb…

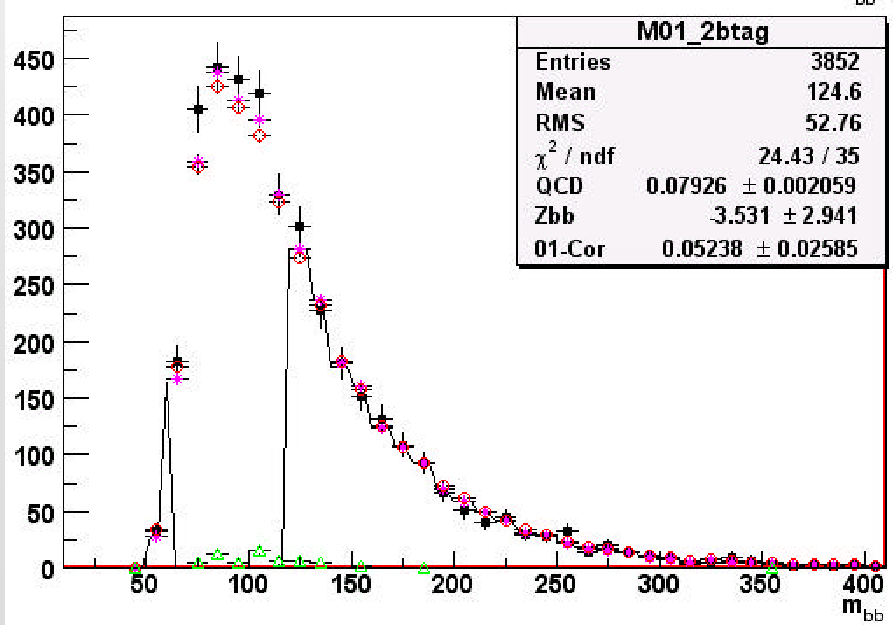




Fitting Without Monte Carlo

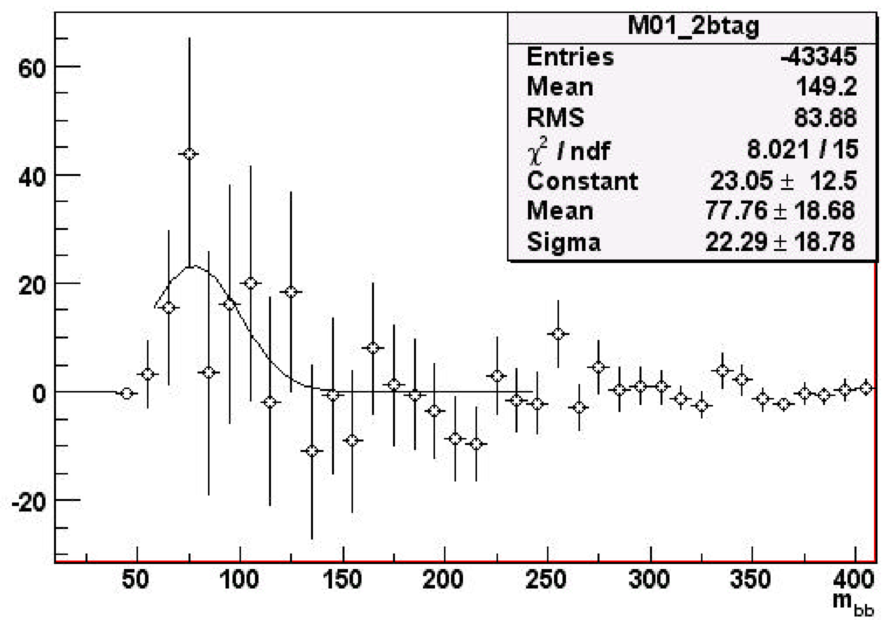
- Remove the Z->bb MC from the fit, and just ignore the 'signal region'
- Chi-square / NDF goes from 15/23 to 24/35 (.67 to .69)
- Number of events remains the same as when including Z->bb Monte Carlo
- Position and width of peak is stable with previous result

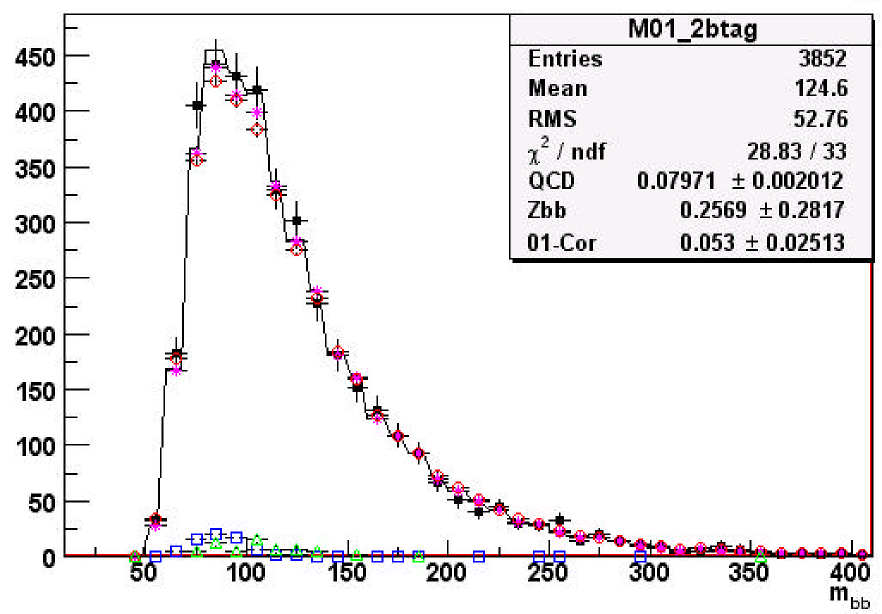




Shifting Z Monte Carlo Mass Positive

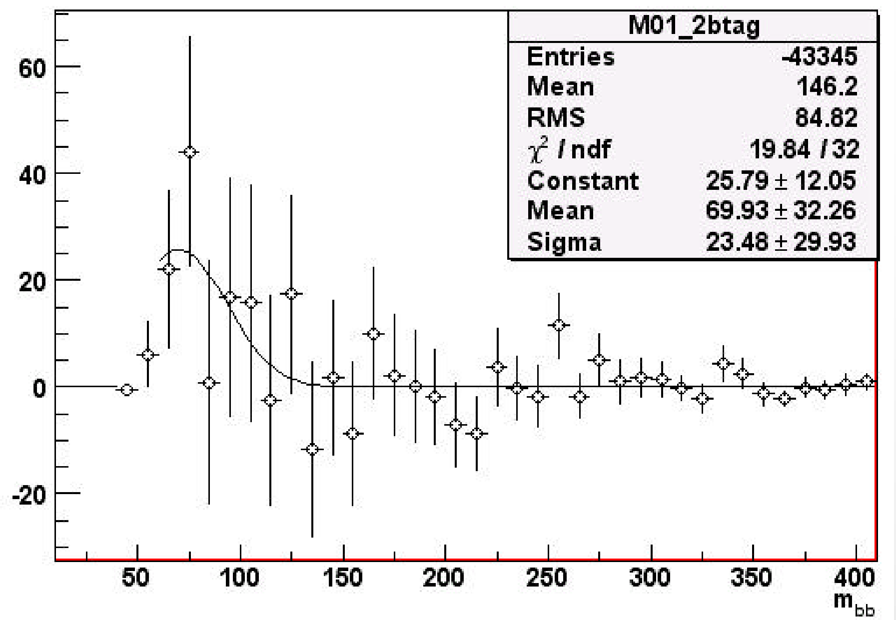
- Z mass peak in MC shifted by +10 GeV/c^2
- Chi-squared goes from 15/23 to 29/33 (.67 to .87)
- Peak in data goes away a bit, but doesn't follow shifted MC

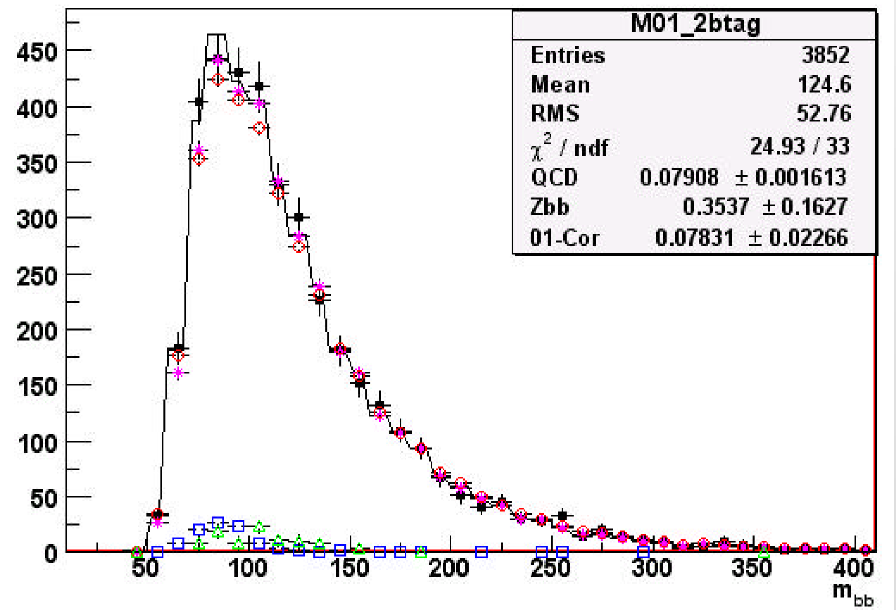




Shifting Z Monte Carlo Mass Negative

- Z mass peak in MC shifted by -10 GeV/c^2
- Chi-squared goes from 15/23 to 25/33 (.67 to .76)
- Peak in data goes away a bit, but doesn't follow shifted MC





Conclusions

- Evidence has been presented for the presence of Z->bb signal
  - Qualitatively in the multi-jet sample
  - Quantitatively in the mu+jet sample
- Results are consistent with
  - Previous CDF Run I observation
  - Expected rates from Monte Carlo (Z->bb and bb)
  - Measured jet energy resolutions in p13 data
  - Variations in fitting technique, cuts, tagging, and triggers
- Triggering is what makes Z->bb hard to see
  - Unprescaled muon trigger with no explicit p_T cut makes observation possible (CDF Run I had a 7.5 GeV p_T muon cut)
- Measuring the Z->bb peak is very important
  - Work aimed at understanding the signal is beneficial to the more general understanding of b-tagging, triggering, and jet energy resolution

This study depended on these peoples' helpful conversations, insight, and expertise. Thanks for everything!

Avto, Andre, Levan, J. Hobbs, and Gordon