

W+jets background model from data

an update

Orin Harris & John BackusMayes

University of Washington

Single top group meeting

1-10-2007

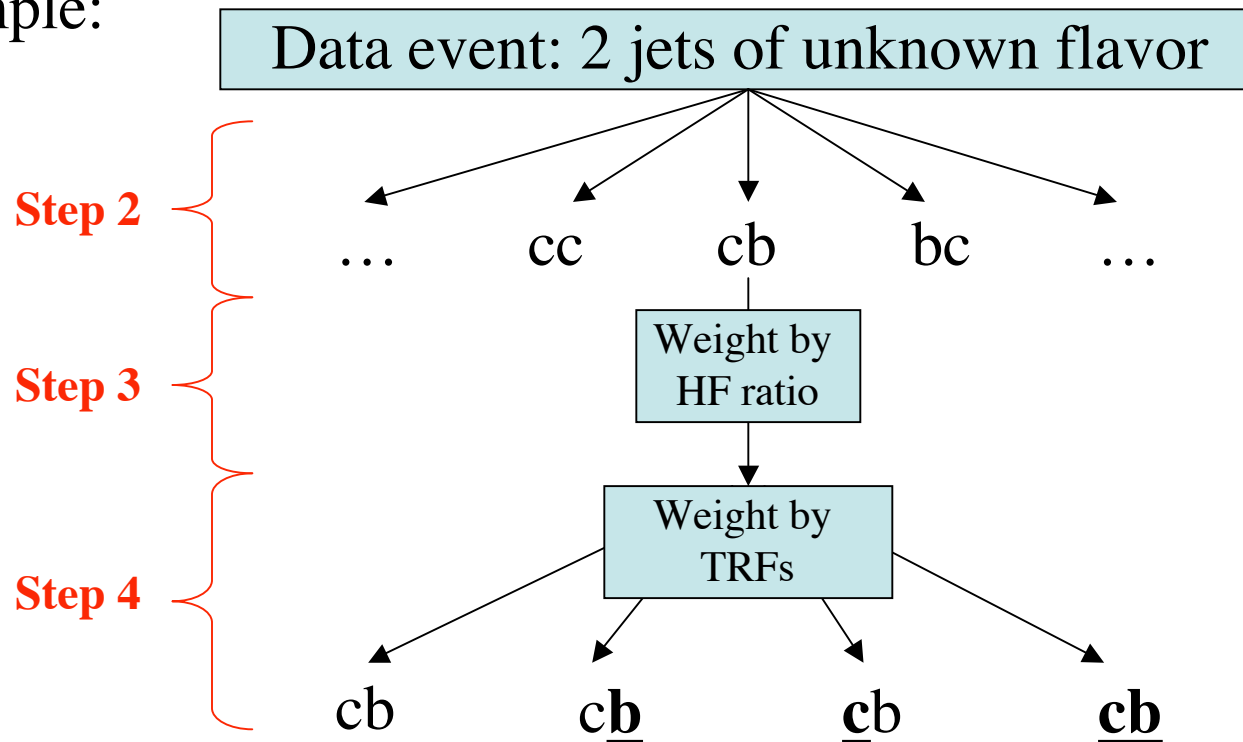
Outline

- Method for creating W+jets background from data
- The model's sensitivity to HF ratios
- How and why the model is inaccurate
- Flavor dependence on the angle between jets
(could it be used to selectively apply jet flavors to data?)
- A look at W+light jets from data
- Conclusion

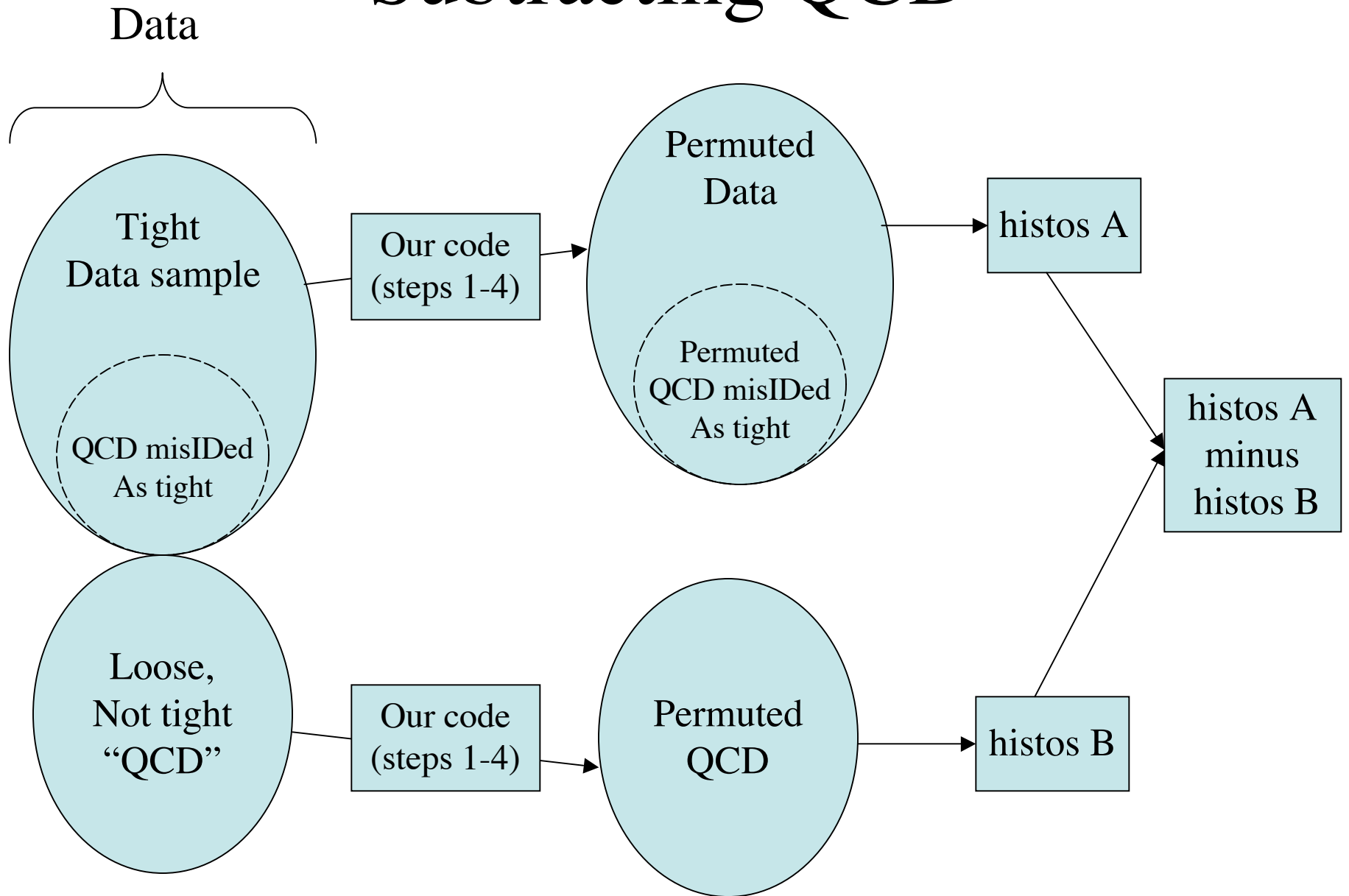
Method: data masquerades as MC

1. Use b-tagging processor to apply TRFs to pre-tag single top data skim
2. Split each event into a set of identical events, one event for each possible permutation of jet flavors {light,c,b}
3. Weight each permutation by its corresponding HF fraction
4. Pass each of these to the b-tag permuter to apply TRFs
5. Subtract QCD (next page)

For example:



Subtracting QCD



MC-based HF ratios: count # of events in each flavor-combination bin in W+jets MC skim

2 jet bin
(67.0914% of data)

bb: 1.45286%
cb: 0.176551%
cc: 2.61726%
jb: 2.92003%
jc: 12.2888%
jj: 80.5445%

NLO MCFM:
Wbb/Wjj = 1.3%
(v00-08-07: 1.8%)
(v00-07-04: 1.2%)

3 jet bin
(25.4291% of data)

bbb: 0.0237041%
cbb: 0.0605243%
ccb: 0.0273977%
ccc: 0.289969%
jbb: 2.07993%
jcb: 0.426397%
jcc: 4.88401%
jjb: 3.7245%
jjc: 15.0079%
jjj: 73.4757%

4 jet bin
(7.4795% of data)

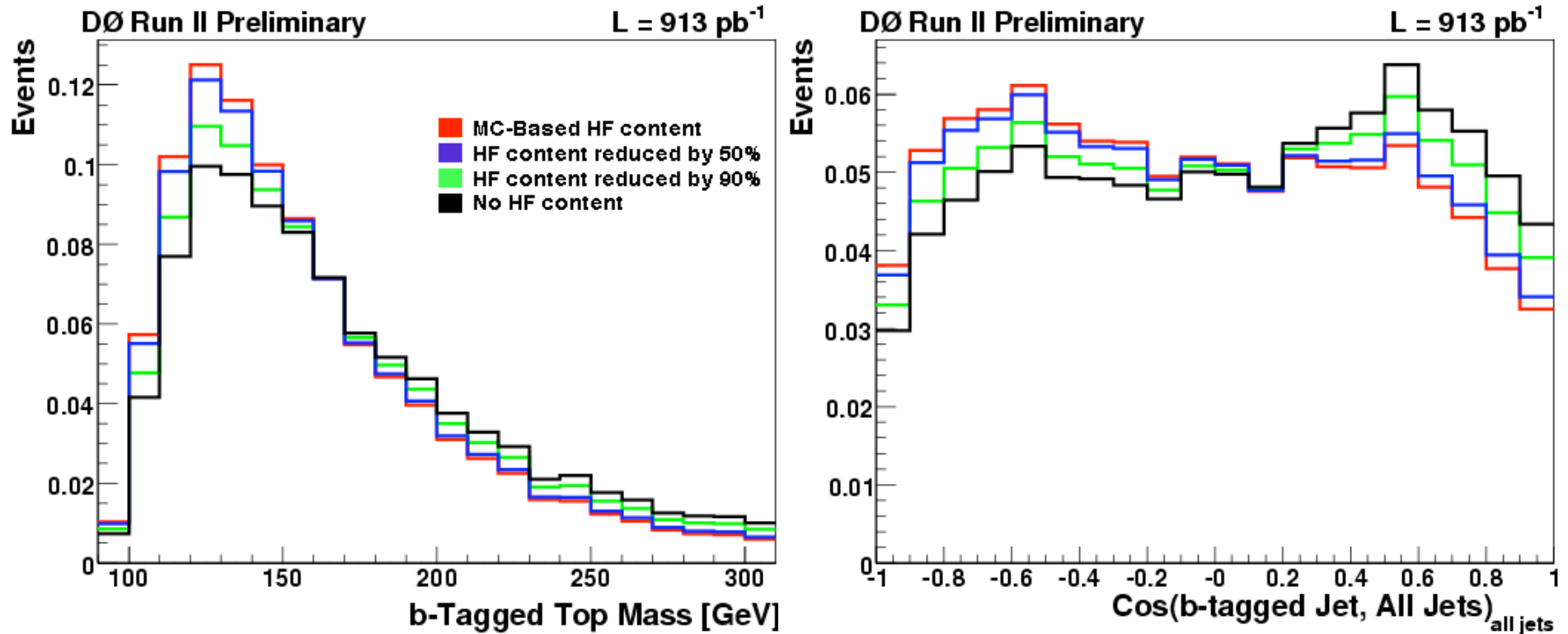
bbbb: 0.0022481%
cbbb: 0%
ccbb: 0.00114672%
cccb: 0%
cccc: 0.0482936%
jbbb: 0.0371421%
jcbb: 0.130879%
jccb: 0.0482936%
jccc: 0.570356%
jjbb: 2.55949%
jjcb: 0.572607%
jjcc: 6.71521%
jjjb: 4.51496%
jjjc: 13.7427%
jjjj: 71.0633%

The next few slides attempt to show:

- Our model shows some sensitivity to HF content
- But for a given HF content, our model is not sensitive to HF ratios
- The deviation of our model from MC corresponds to the HF content of our model
- This implies that our model when applied to W+HF deviates from MC, and is insensitive to HF ratios

Sensitivity to HF content

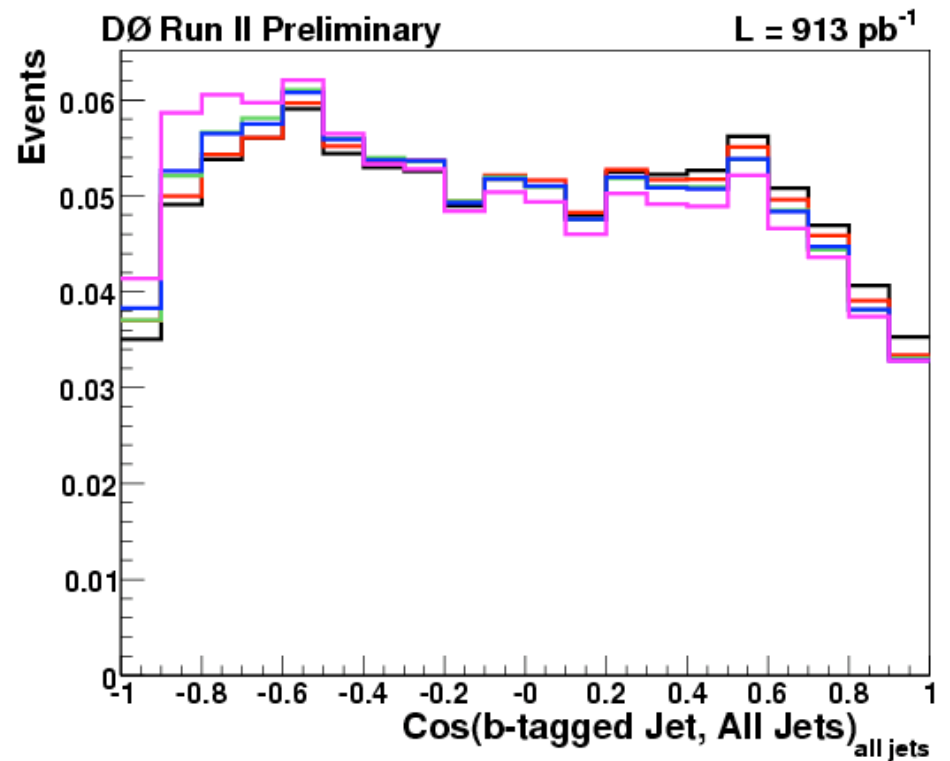
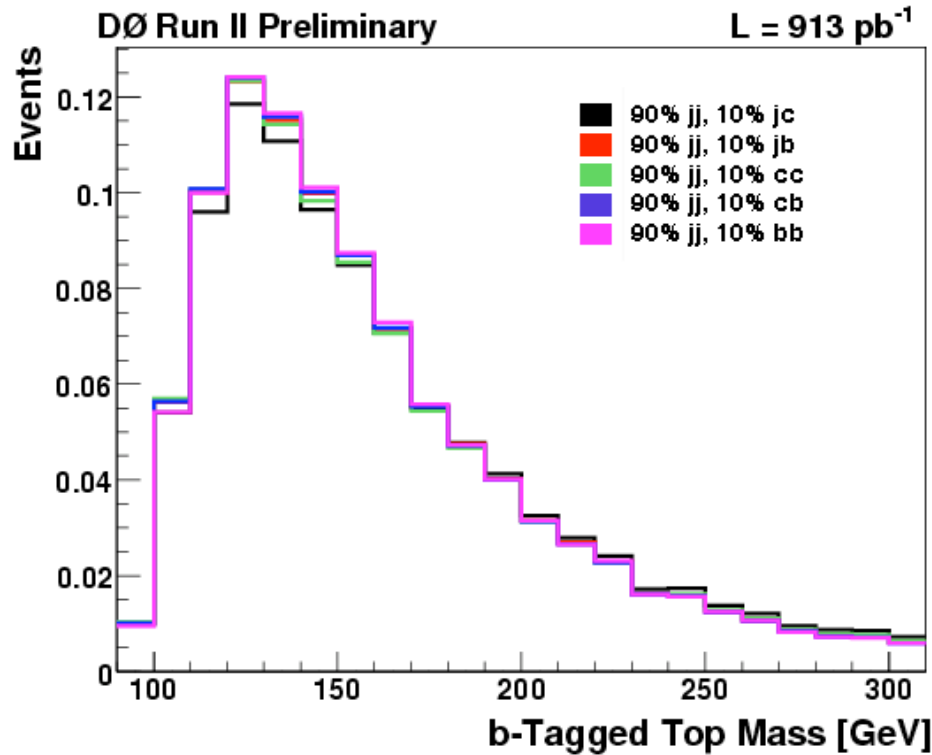
(W+e+2jets, EqOneTag, normed)



The Data, QCD, ttbar, W+jets skims we used are located in:
</rooms/sauna/singletop/p17MC/CC/TopCafe00-08-07/PreTag/filelists/>

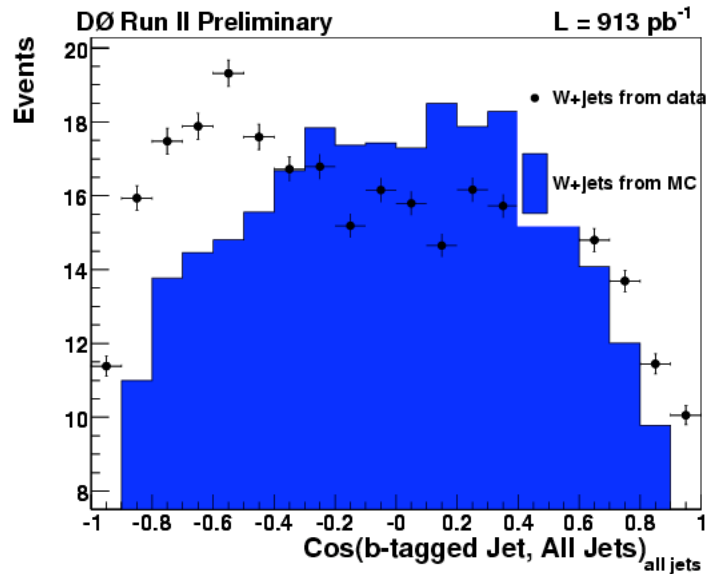
For a given HF content, model is insensitive to HF ratios

(W+e+2jets, EqOneTag, normed)

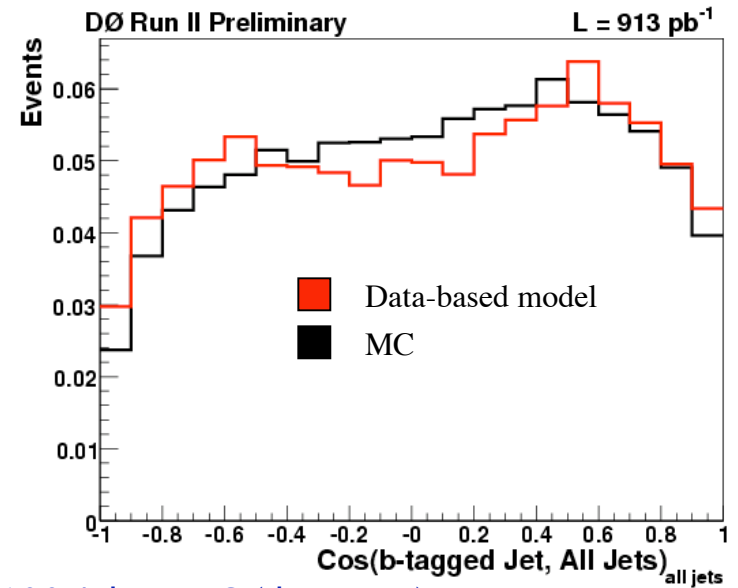


The deviation of our model from MC corresponds to the HF content of our model

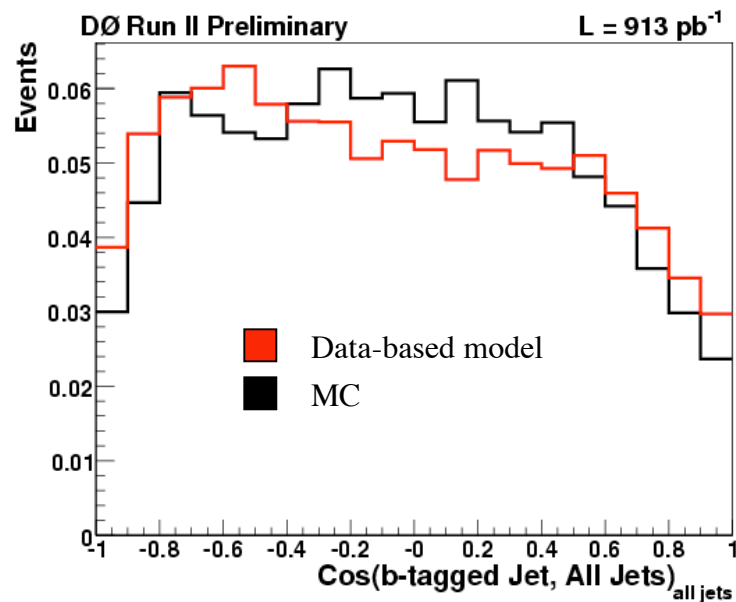
Data-based model vs MC



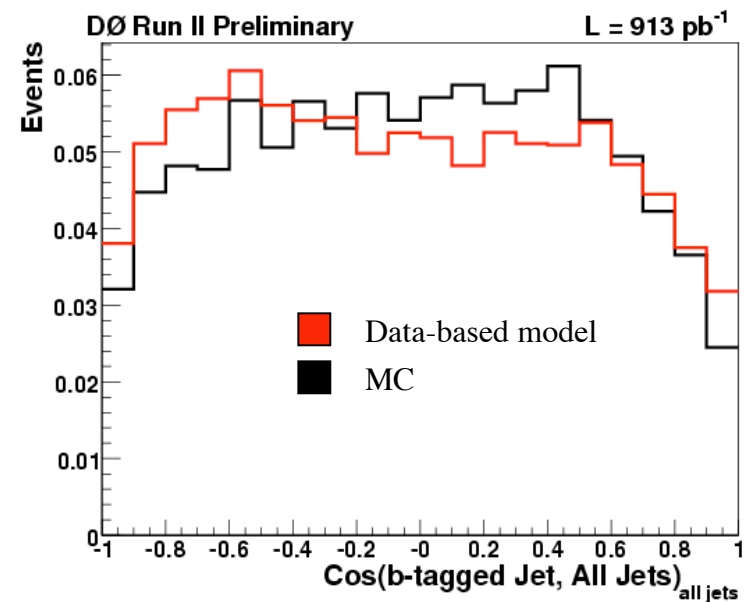
100% jj vs MC (jj events)



100% jc vs MC (jc events)

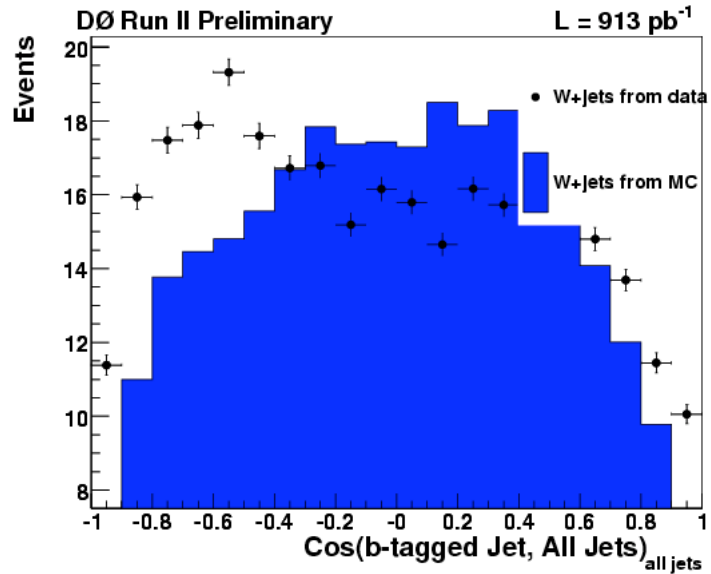


100% jb vs MC (jb events)

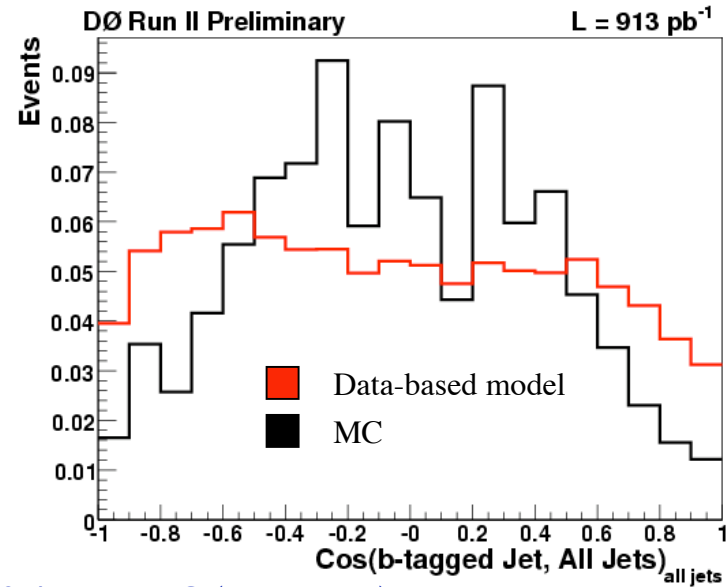


... continued

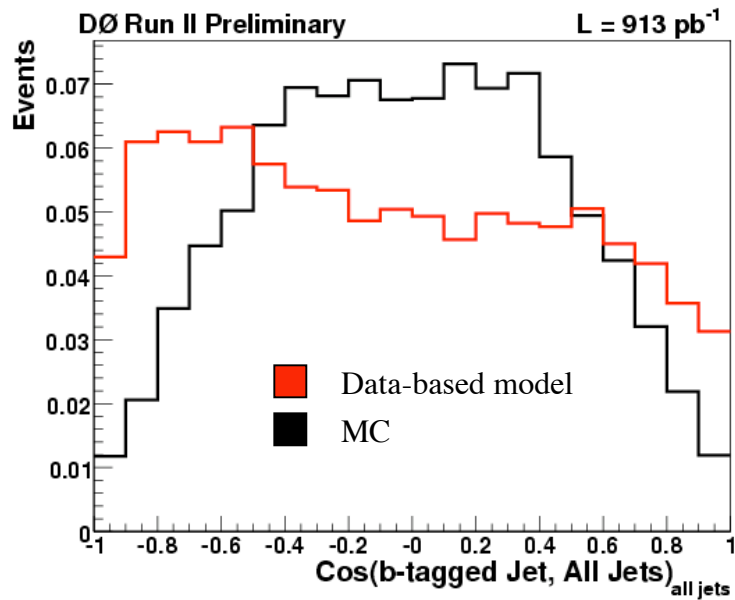
Data-based model vs MC



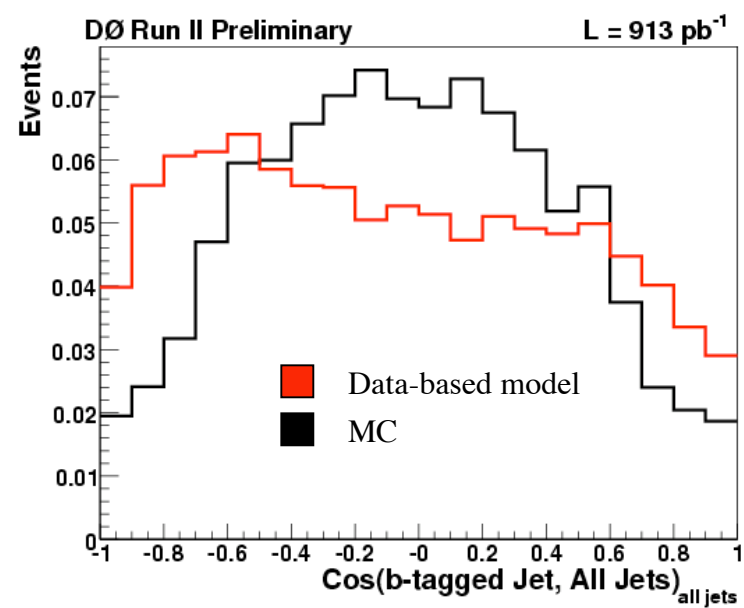
100% cb vs MC (cb events)



100% bb vs MC (bb events)



100% cc vs MC (cc events)



It appears likely that:

- Given that our data sample consists of $\sim 80\%$ light jet events, our model is being polluted by misapplying HF TRFs to light jet events for $\sim 80\%$ of the data sample on which we base our model

This is a problem because:

- Real bb events have a certain characteristic distribution for the angle between jets
- Our model does not recreate this distribution:
 - “bb” events in our model may actually be jj or cb (or any permutation of flavors), which have their own characteristic distributions, distinct from bb
 - Our weights only depend directly on jet pT and eta (the TRF’s are functions of these variables alone), not the angle between jets
 - Our model therefore produces an angle distribution that’s some combination of the characteristic distributions for jj, cb, bb, etc.

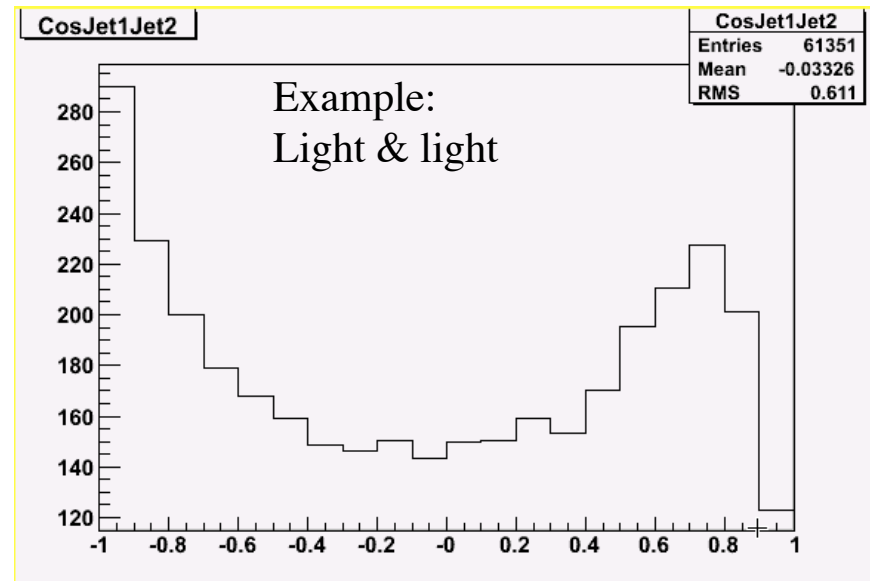
Flavor dependence on the angle between jets

-- Could be used to selectively apply jet flavors to data --

So far have found little discriminating power:

Mean Cosine of angle between:

- light & light: -.033, StdDev: .611
- light & c: -.021, StdDev: .590
- light & b: -.039, StdDev: .597
- c & b: .126, StdDev: .608
- c & c: .150, StdDev: .584
- b & b: .140, StdDev: .594



Similar studies involving jet Pt, Eta, Phi are no more promising.

W+light jets from data

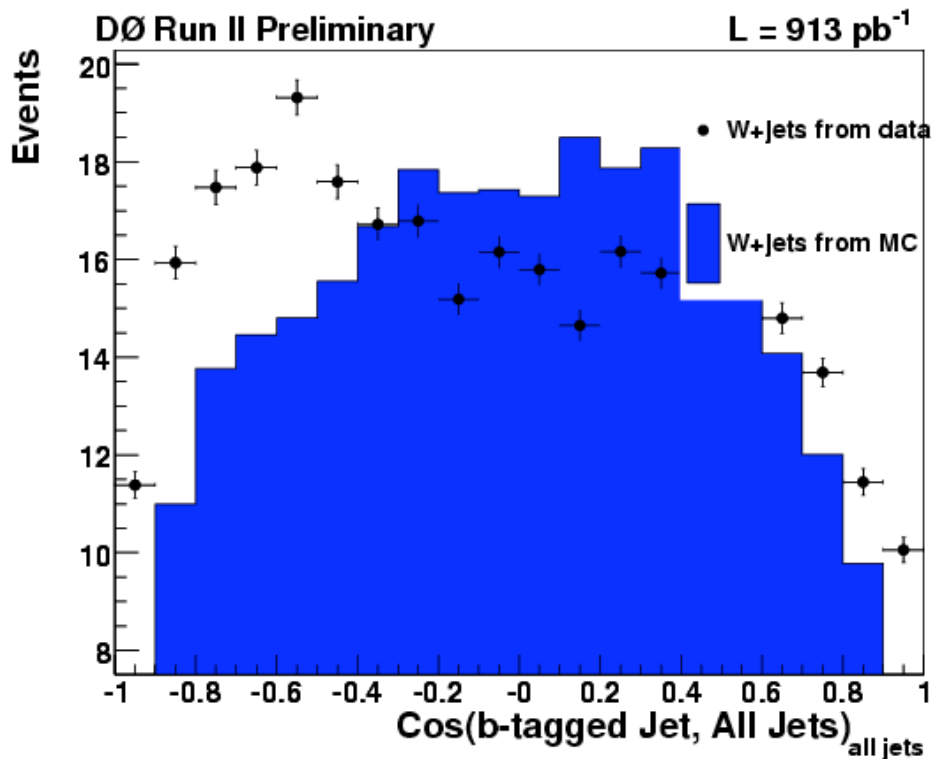
- Our method fails to adequately model W+HF due to data being predominately W+light
- It may still be useful to use data to model just W+light, letting MC model W+HF
- CDF and D0 have done this before¹

Method

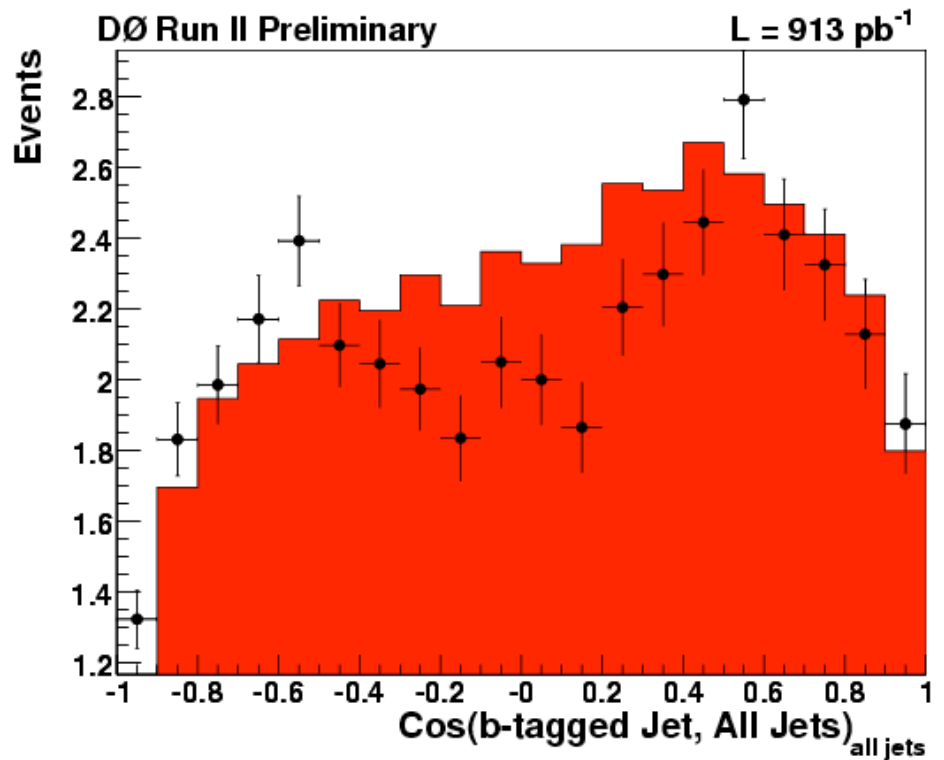
- The method is simpler than before: just send data to the ordinary b-tag permuter, except applying only light TRFs to all events.
- Then subtract (in addition to the QCD, etc.) the HF in data that gets mis-tagged as light, by applying our method to the HF portion of W+jets MC, and subtracting after applying the matrix method.

W+jets model, preliminary W+light jets model

W+jets model vs MC



W+light jets model vs MC (light jets)



Conclusion

- A W +jets background model from data is unlikely to be viable unless a way is found to better discriminate between light and HF events.
- A W +light jets background model from data is worth investigating further.

(additional plots can be temporarily found at:

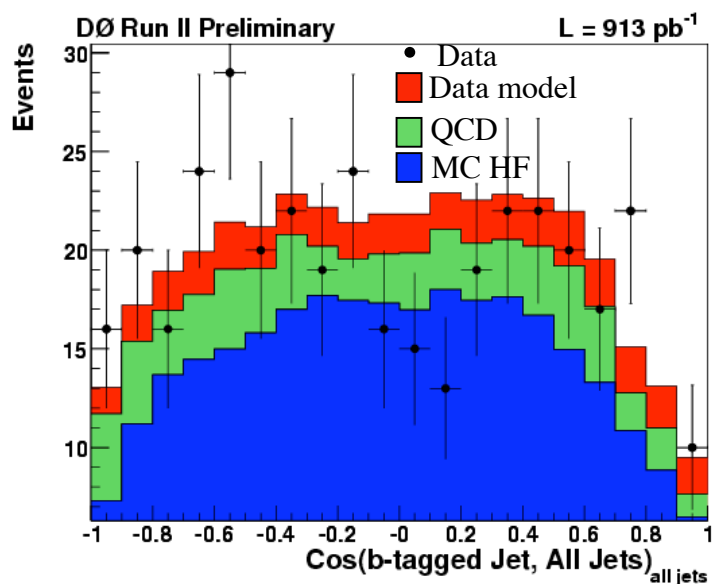
http://www-clued0.fnal.gov/~orin/Graphs/2j_EqOneTag/)

End Notes

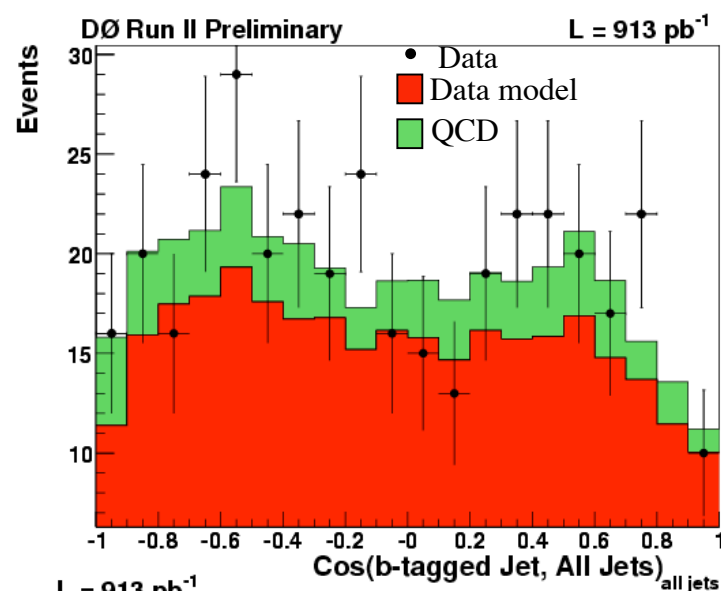
- 1
 - CDF notes 8110 and 8185, and Phys. Rev. D, Vol 50, No. 5, p2966, 1994
 - D0 note 4967

How do the various models compare to data?

W+light model vs Data



W+jets model vs Data



W+jets MC vs Data

