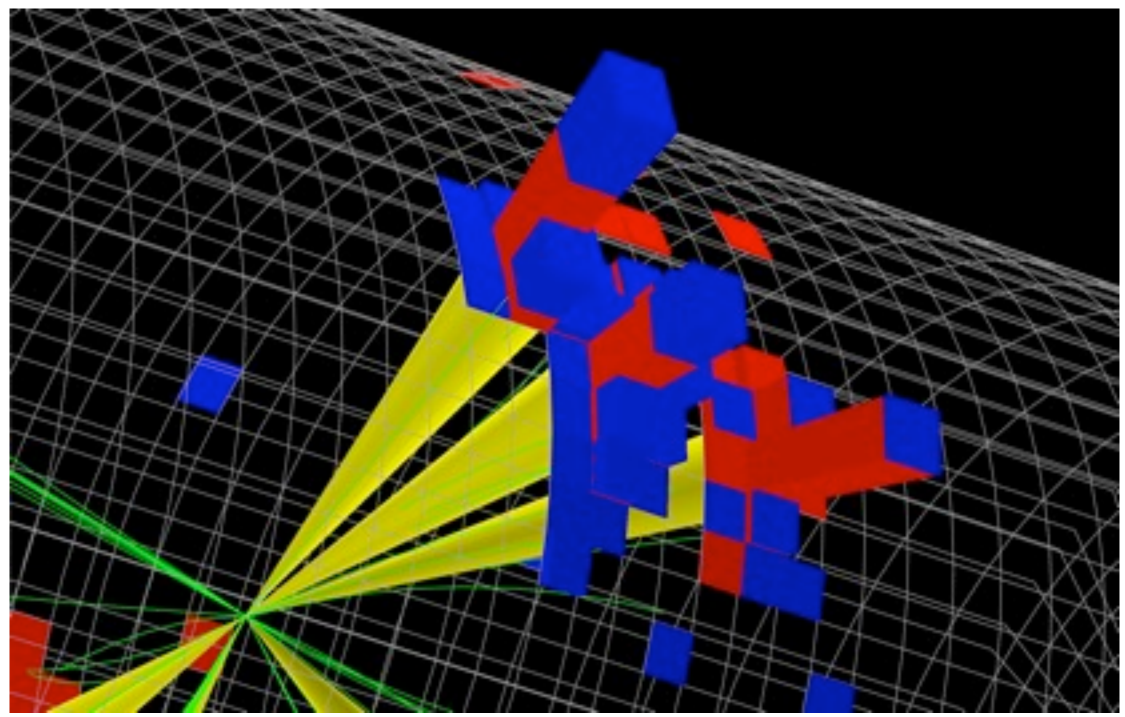
Jet Substructure in pp Collisions at 7 TeV in CMS



James Dolen (UC Davis)

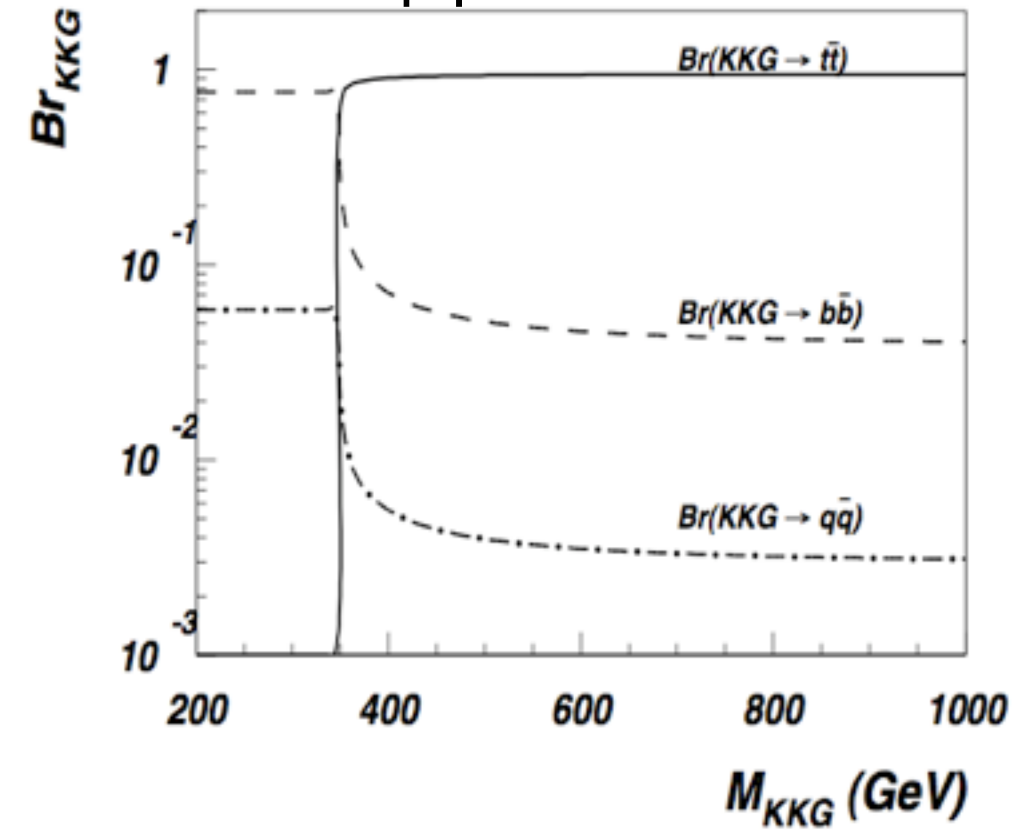
on behalf of the CMS TTBSM and JetMET Groups

Boston Jet Physics Workshop 13-Jan-2011

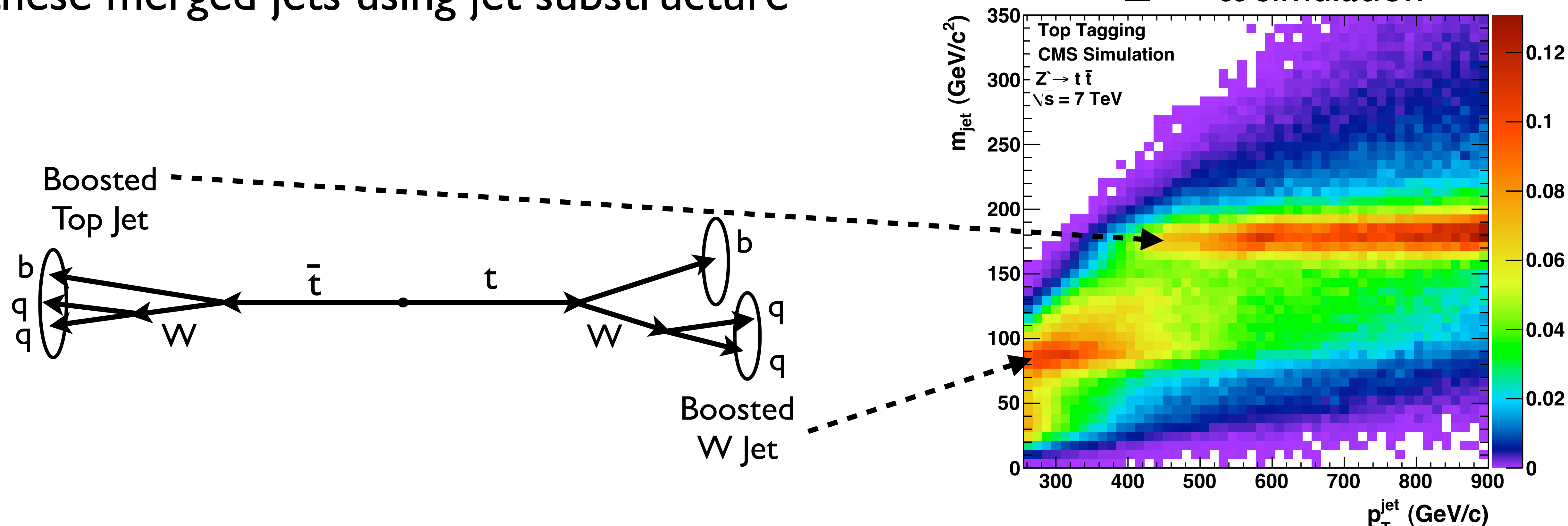
Introduction

- New physics scenarios often predict heavy resonances who's decay products decay hadronically
  - Top, W, Z etc.
- These decay products are highly boosted, and thus may merge within one jet
- CMS is working on techniques for tagging these merged jets using jet substructure

hep-ph/0612015v1



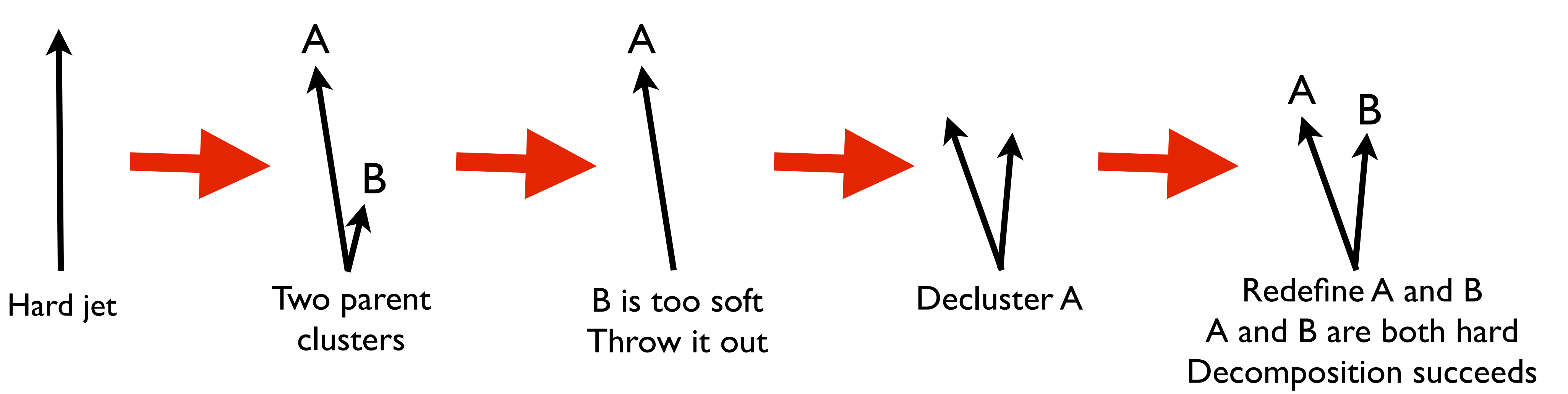
$Z` \to t\bar{t}$ simulation



Top Tagging Algorithm

- Based on the Hopkins Algorithm (Kaplan, Rehermann, Schwartz, Tweedie) (arXiv:0806.0848)
- Cluster particle flow candidates using Cambridge Aachen R=0.8
- Reverse the clustering sequence in order to find substructure
- Subjets must satisfy two requirements
  1. Momentum fraction criterion: $p_T^{subjet} > 0.05 \times p_T^{hard\ jet}$
  2. Adjacency criterion: $\Delta R(A, B) > 0.4 - 0.0004 \times p_T$
- Iterative process - throw out objects that fail momentum fraction cut and try to decluster again

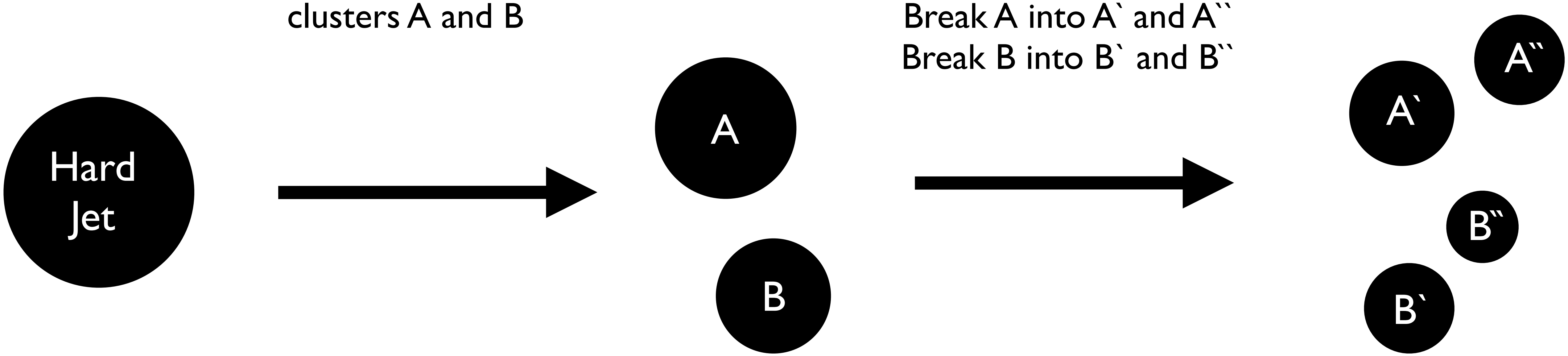
Example Decomposition



Top Tagging Algorithm

<u>Primary Decomposition:</u>
Break jet into two parent clusters A and B

<u>Secondary Decomposition</u>
(Applied individually to A and B)
Break A into A` and A``
Break B into B` and B``



Four Possible outcomes:

1. Primary decomposition fails - the jet has 1 subjet (itself)

2. Primary succeeds, both secondary decompositions fail - the jet has 2 subjets (A and B)

3. Primary succeeds, one secondary decomposition fails, and one succeeds - the jet has 3 subjets ({A, B`,B``} or {A`,A``,B})

4. Primary and both secondary decompositions succeed - the jet has 4 subjets (A`,A``,B`,B``)

Top Tagging Variables

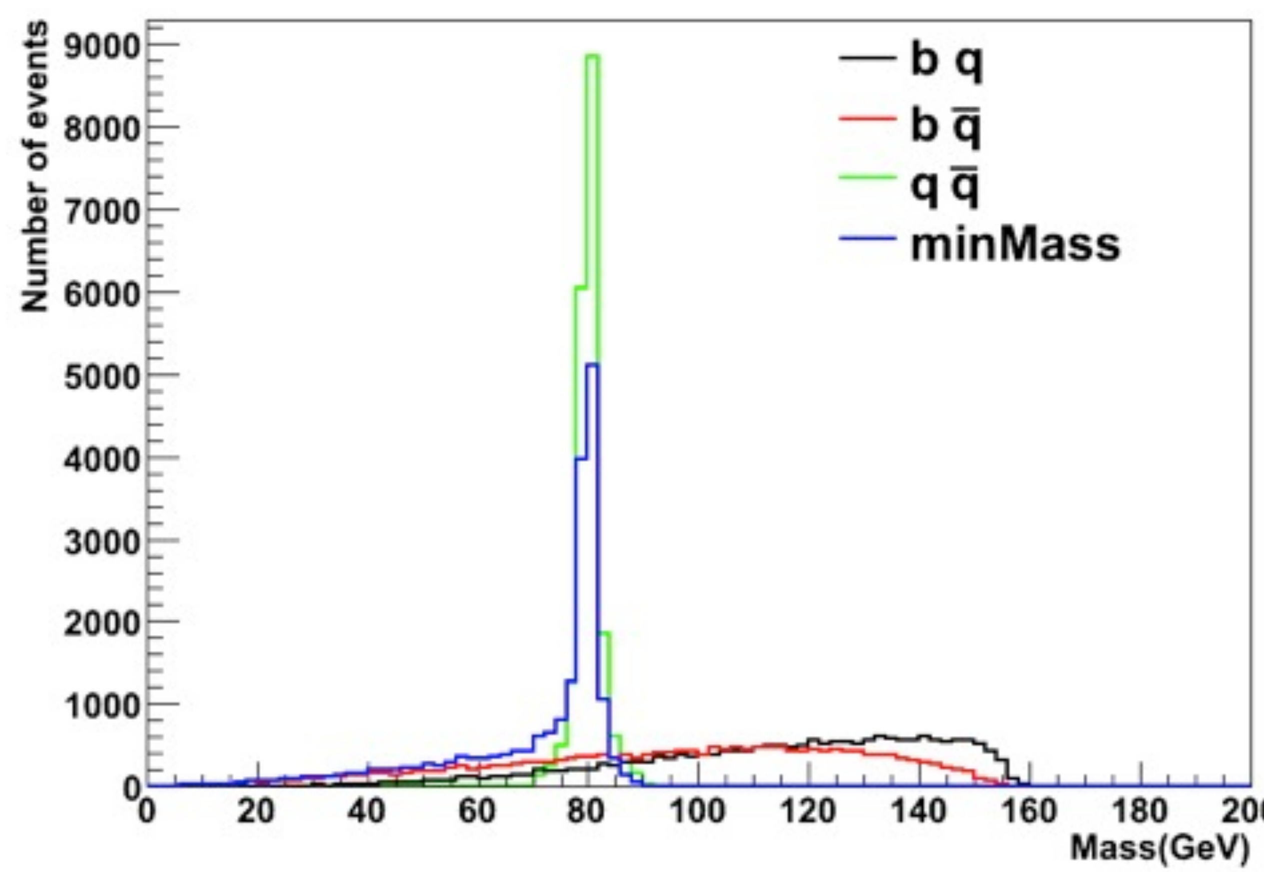
Three Variables are used to tag tops:

1. Jet mass - top jets have mass within a top mass window
2. Number of subjets - top jets have 3 or 4 subjets
3. Minimum Pairwise Mass - pairwise mass of two of the subjets should reconstruct the W mass

Define MinMass = minimum pairwise mass of subjets

$$m_{\min} = \min[m_{12}, m_{13}, m_{23}]$$

Generated top decay products: pairwise mass



Top Tagging Variables

Three Variables are used to tag tops:

1. Jet mass - top jets have mass within a top mass window
2. Number of subjets - top jets have 3 or 4 subjets
3. Minimum Pairwise Mass - pairwise mass of two of the subjets should reconstruct the W mass

Define Loose Top Tag Cuts

$140 < m_{jet} < 250$ GeV/c^2

$N_{subjets} \geq 3$

$m_{min} > 50$ GeV/c^2

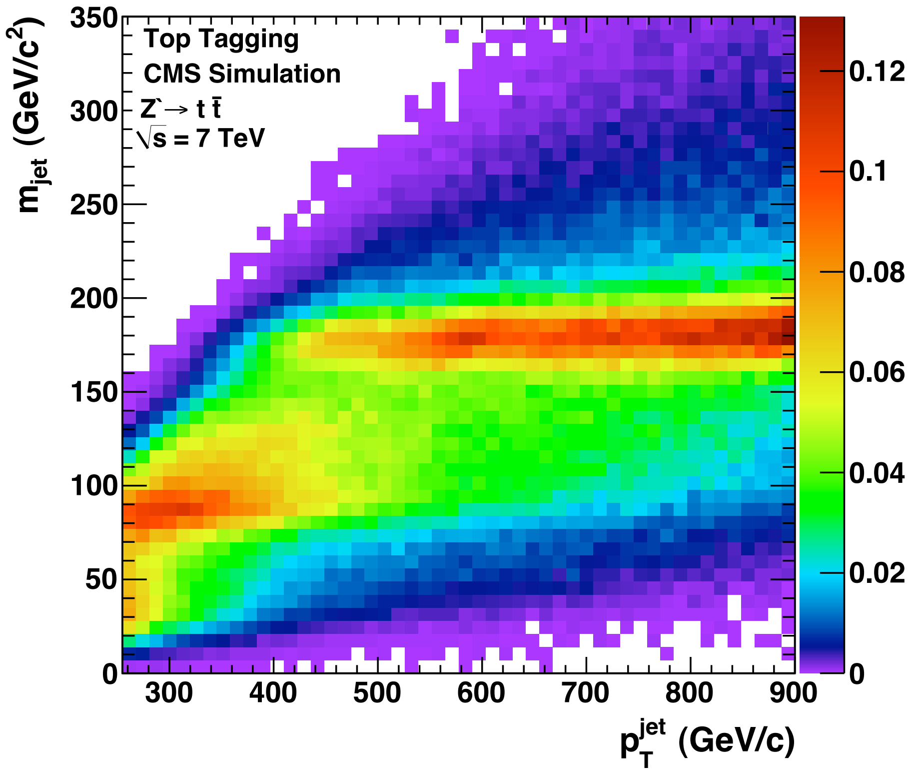
These cuts are purposely loose in order to increase statistics for the study.

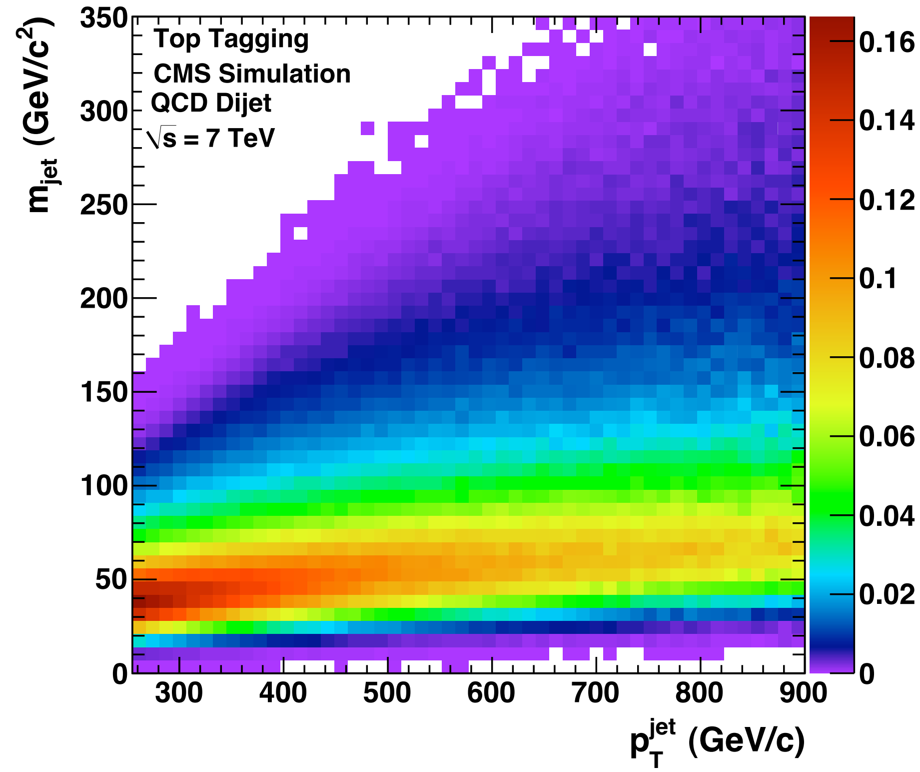
Cuts will be optimized for physics analysis.

Tagging Top Jets

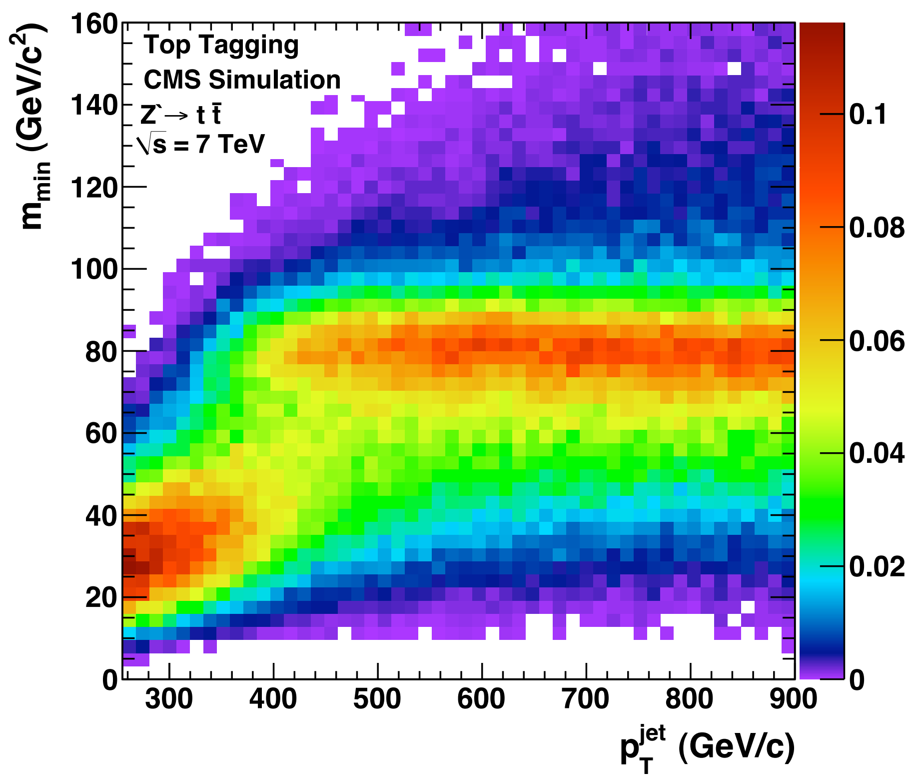
$Z' \rightarrow t\bar{t}$ simulation — QCD simulation

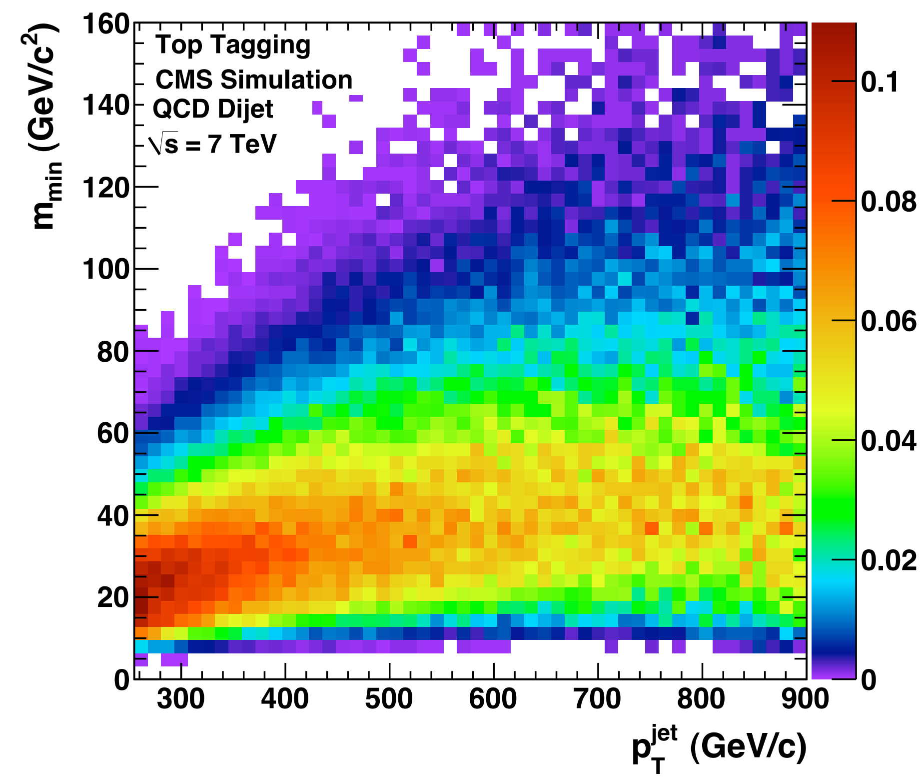
Jet mass vs p_T





Jet MinMass vs p_T



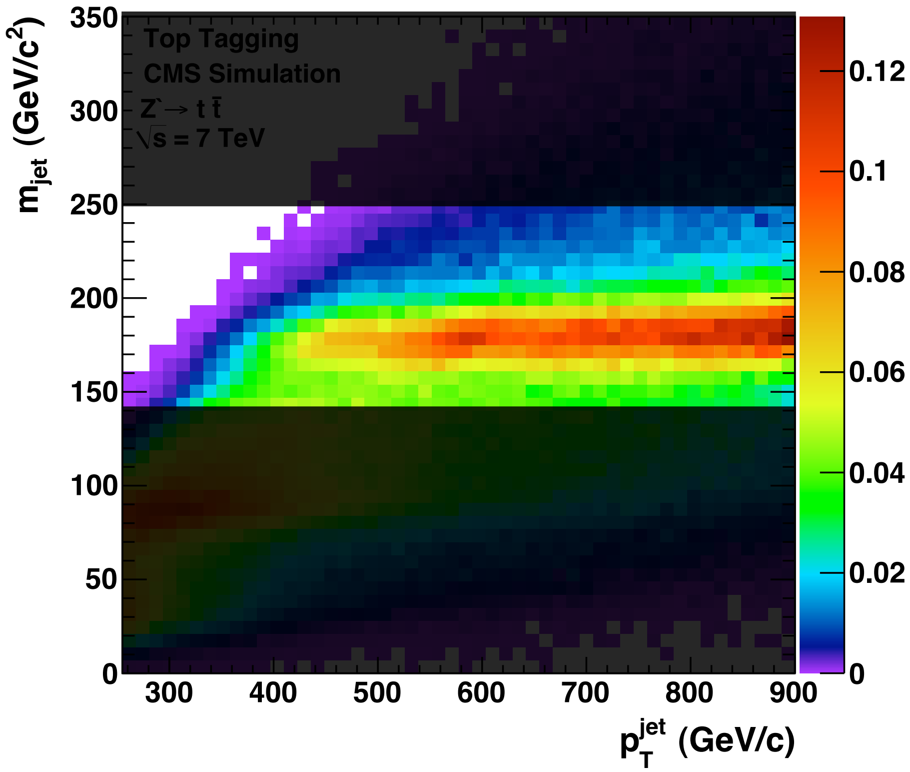


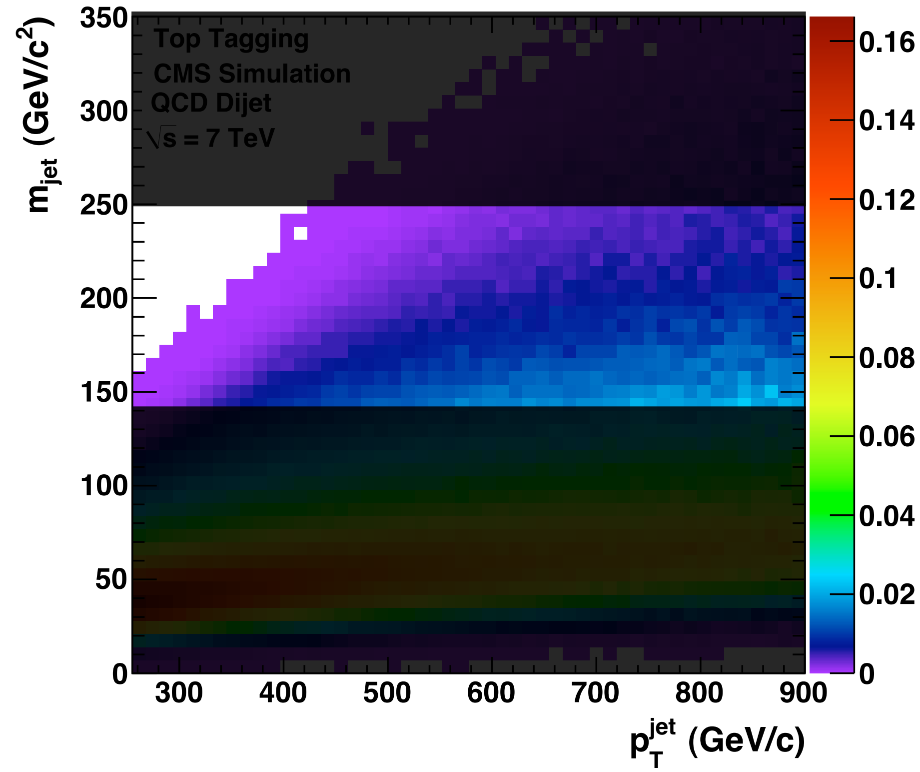
Tagging Top Jets

$Z' \rightarrow t\bar{t}$ simulation

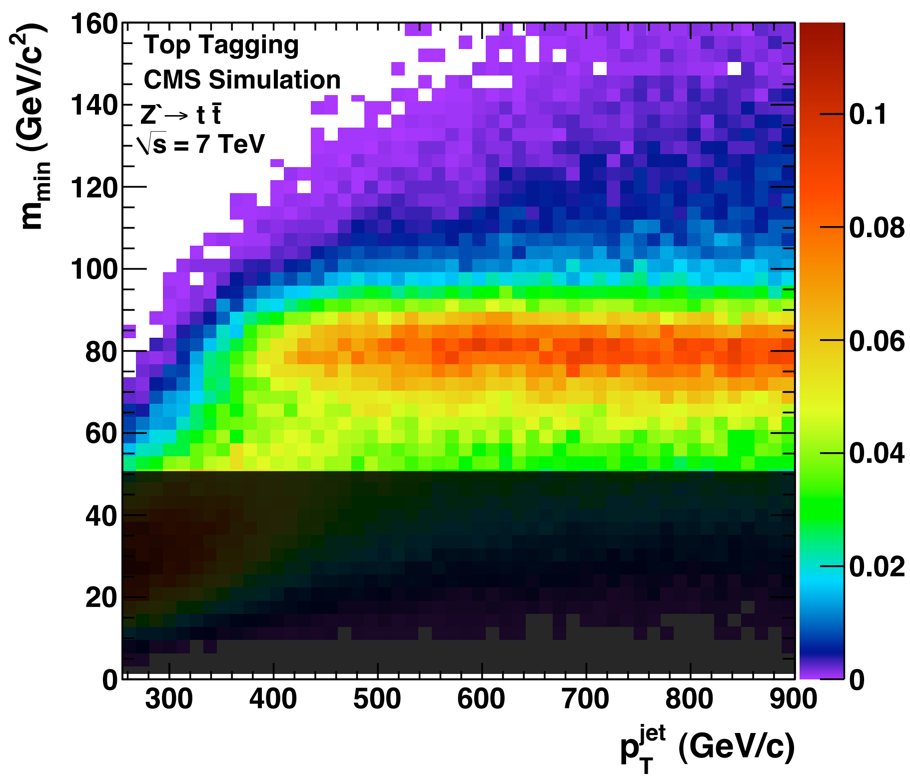
QCD simulation

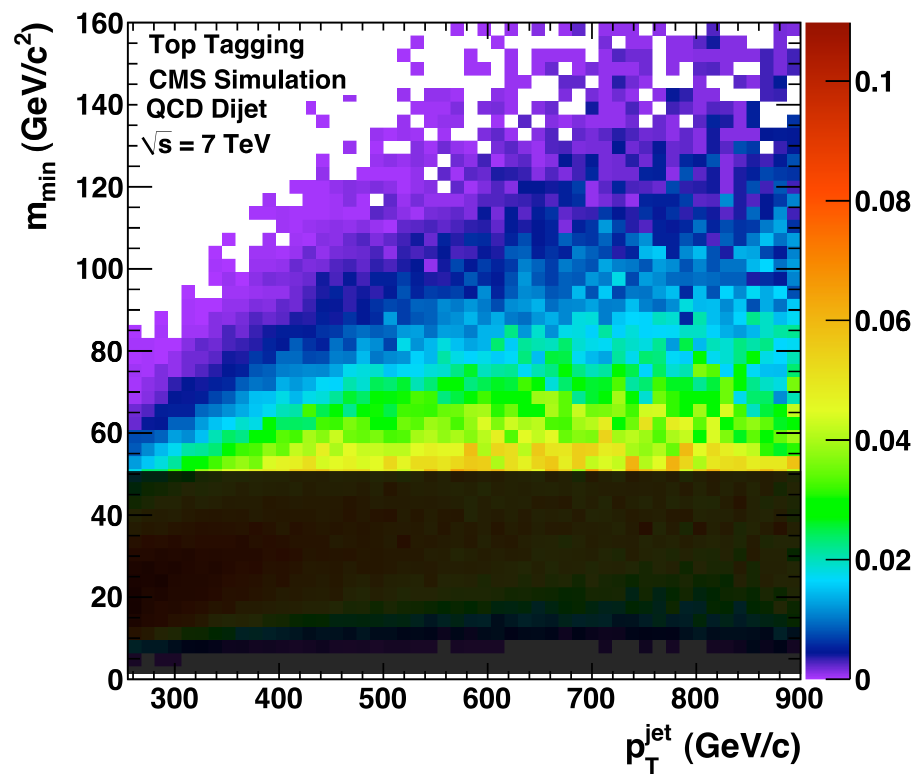
Jet mass vs p_T





Jet MinMass vs p_T



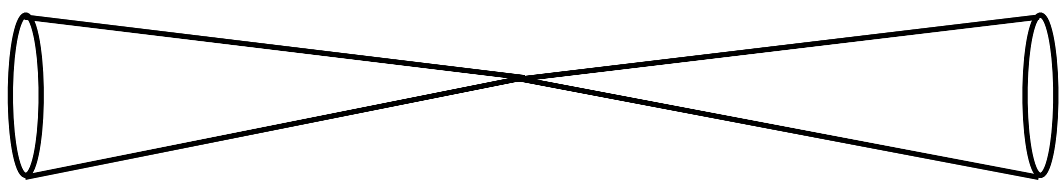


New Results From CMS

- Two days ago CMS approved our first measurements using the top tagging algorithm with data
  - "Study of Jet Substructure in pp Collisions at 7 TeV in CMS"
  - Select dijet events with jet p_T>250 GeV/c
- Main goal: measure mistag rate from data
  - The mistag rate can then be used to estimate QCD background in future search papers
- Secondary goal: commission the algorithm
  - Compare multiple Monte Carlo generators and tunes
    - PYTHIA 6 Tune Z2
    - PYTHIA 6 Tune D6T
    - PYTHIA 8 Tune 1
    - HERWIG++ Tune 23
  - Measure the effect of pileup

Top Tagging Mistag Rate

- Anti-tag and probe method
  - Randomly select one of two leading jets and check if its tagged
  - If the random jet is not tagged, the opposite jet is the probe jet



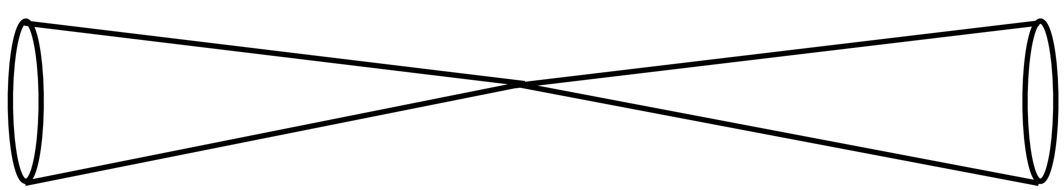
Anti-tagged jet
(mass<140 or mass >250 or minmass<50 or nsubjets <3)

Probe jet

$$\text{Mistag Rate} = \frac{\text{Number of probe jets that are tagged}}{\text{Number of probe jets}}$$

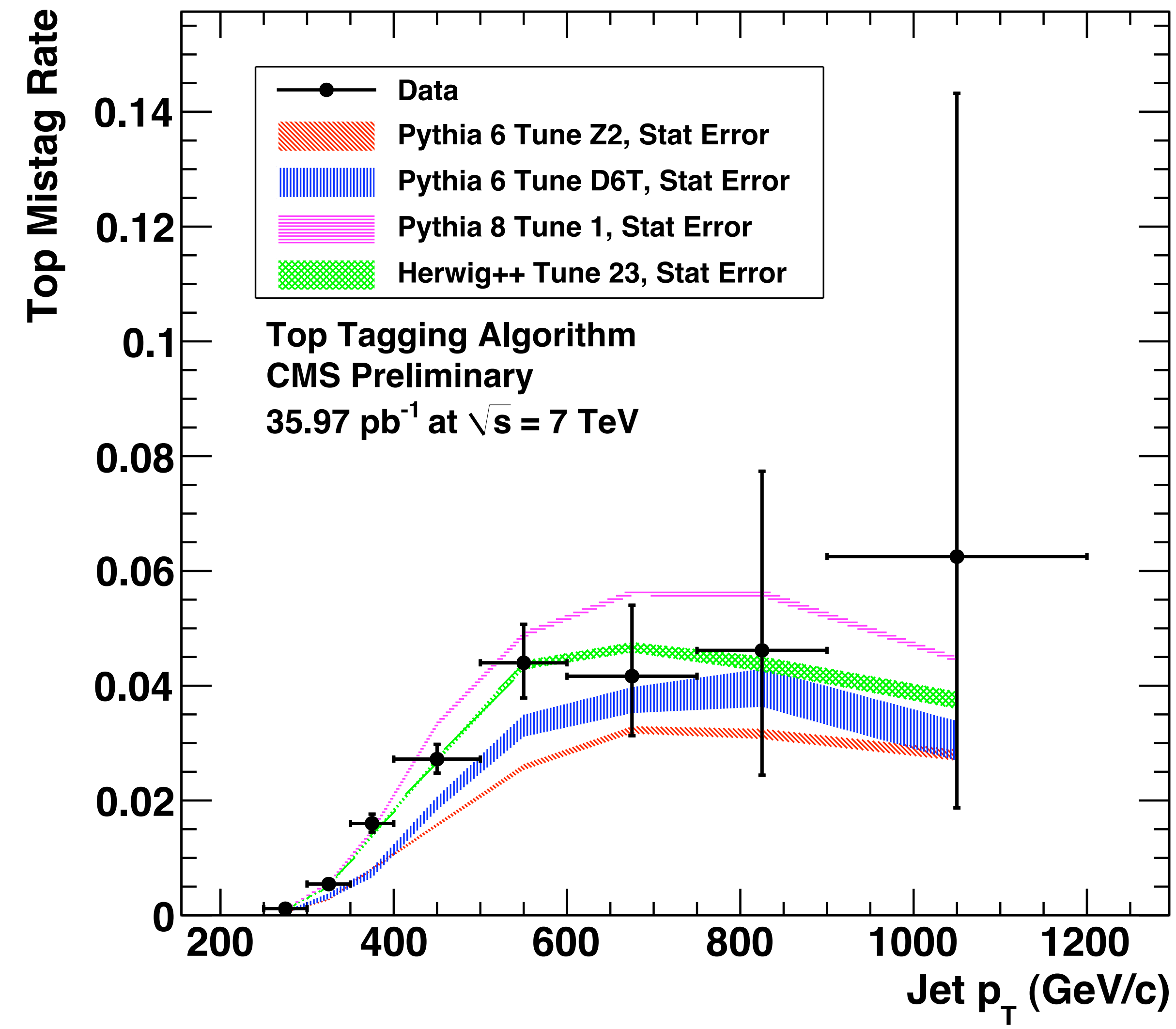
Top Tagging Mistag Rate

- Anti-tag and probe method
  - Randomly select one of two leading jets and check if its tagged
  - If the random jet is not tagged, the opposite jet is the probe jet



Anti-tagged jet
(mass<140 or mass >250 or minmass<50 or nsubjets <3)

Probe jet

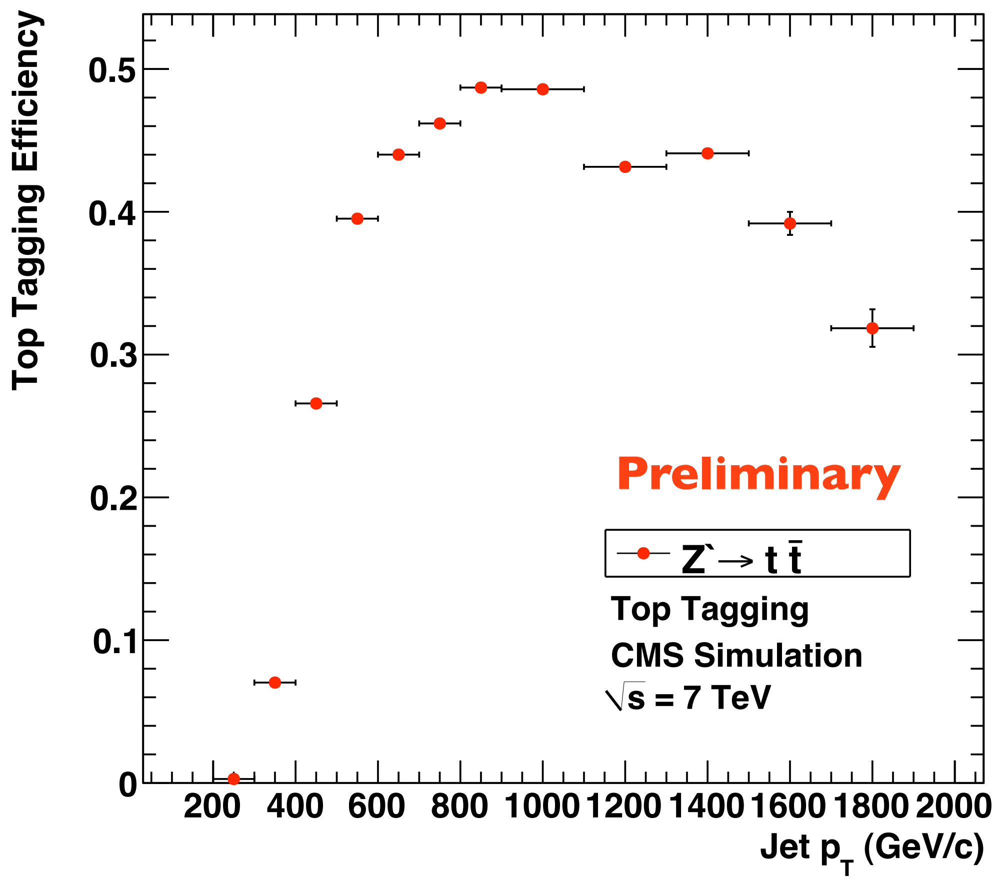


$$\text{Mistag Rate} = \frac{\text{Number of probe jets that are tagged}}{\text{Number of probe jets}}$$

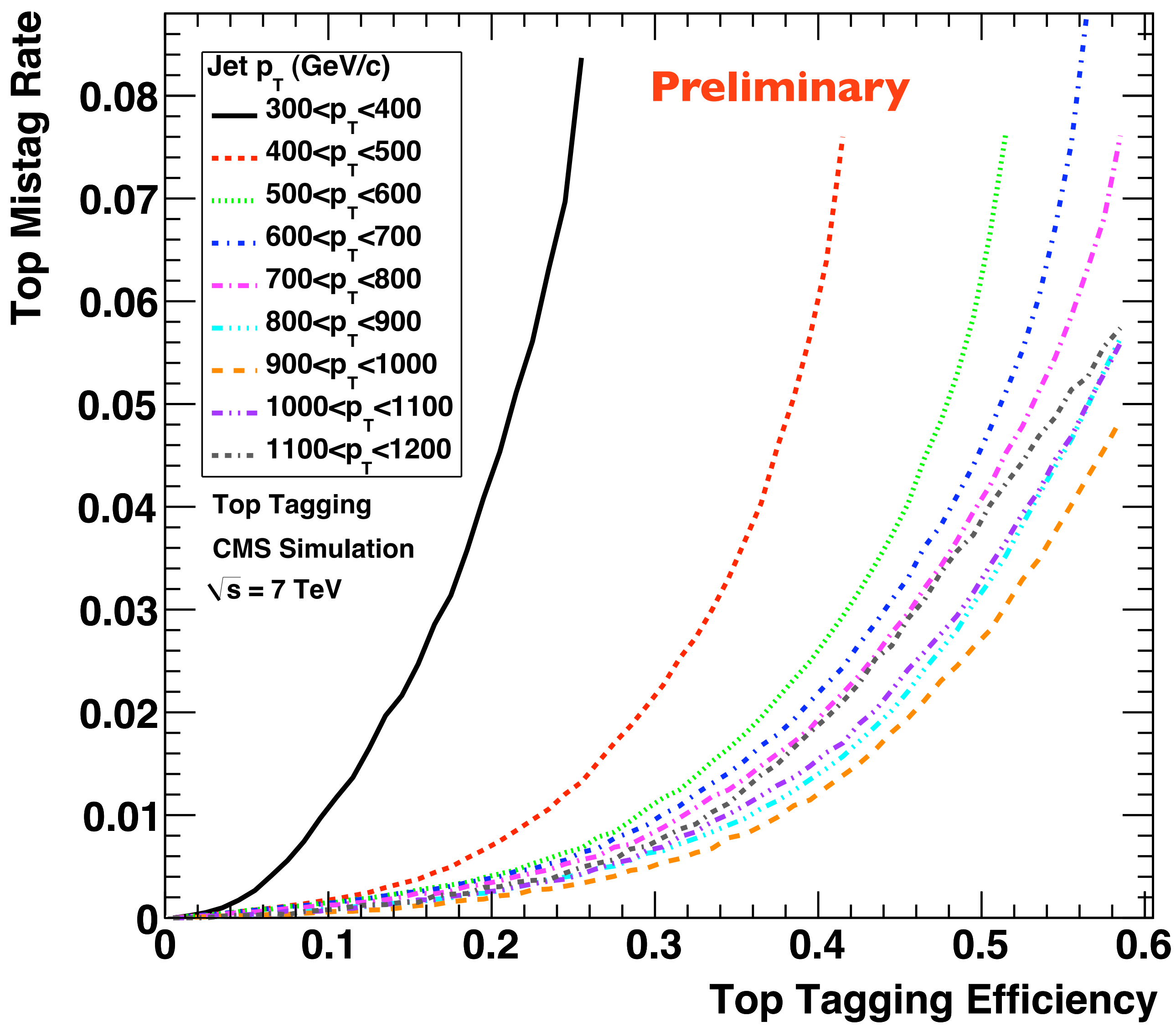
Dependence on shower model and tune, overall good agreement

Top Tagging MC Efficiency

Monte Carlo signal efficiency for our loose top tag working point

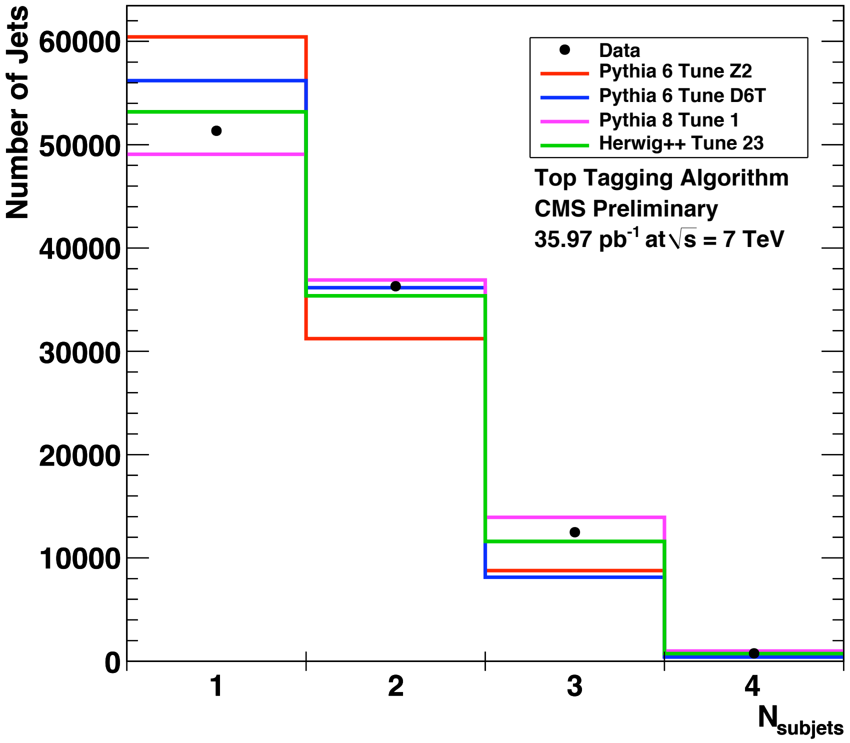


Minimum mistag rate vs efficiency

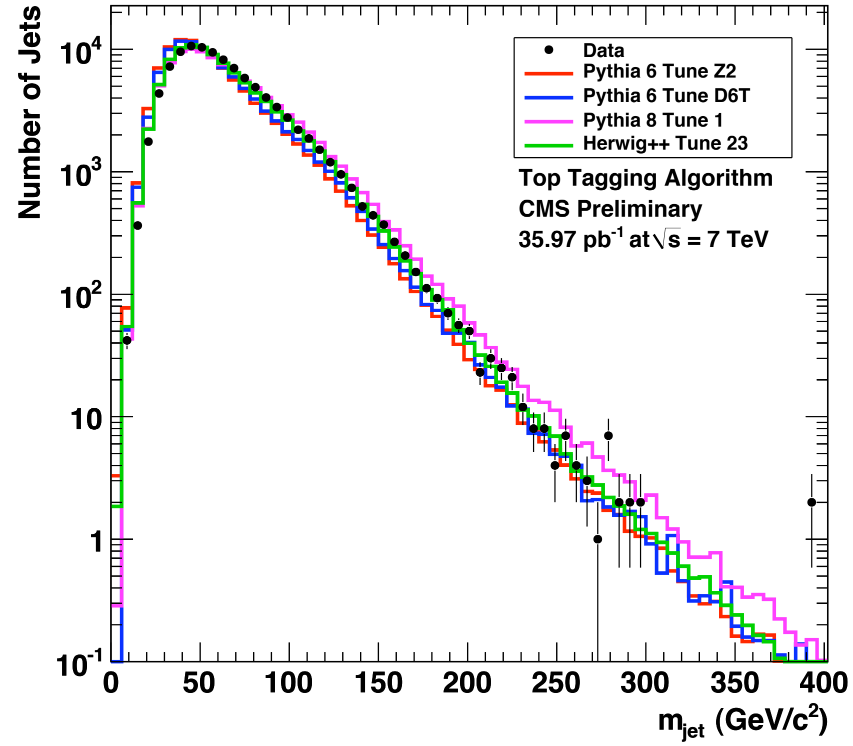


Commissioning

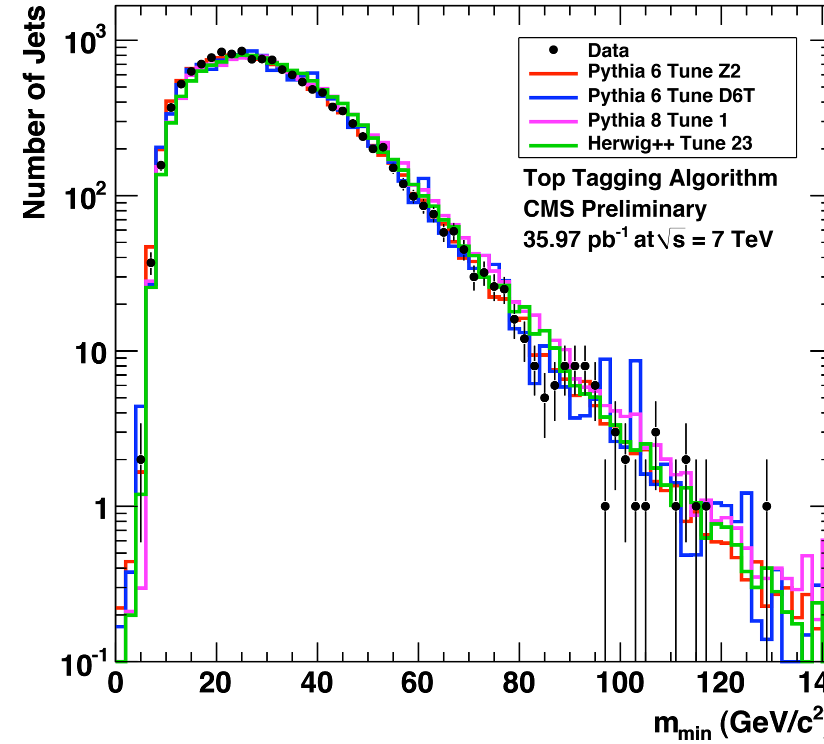
Number of Subjets



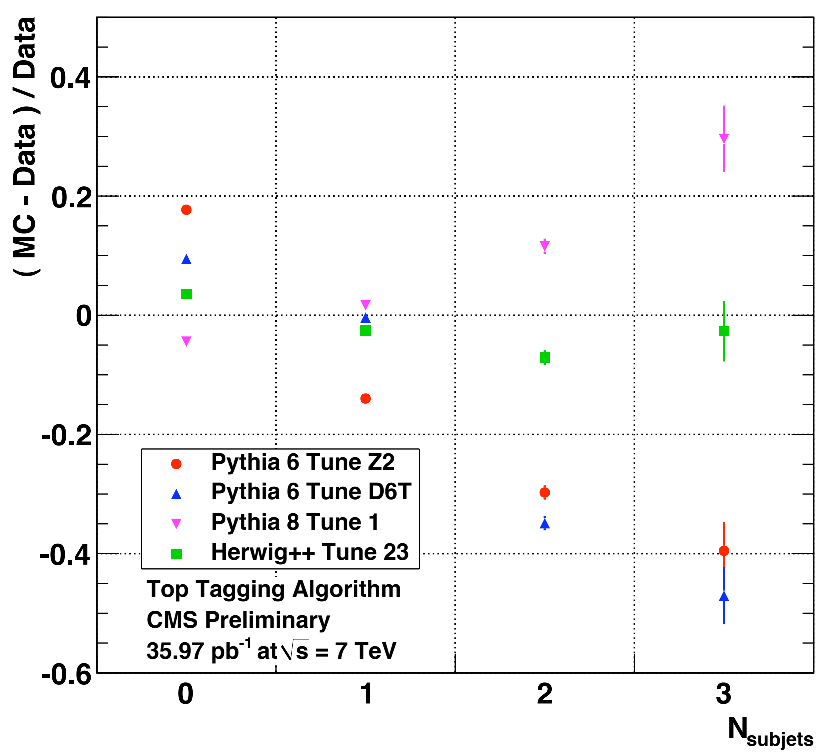
Jet Mass

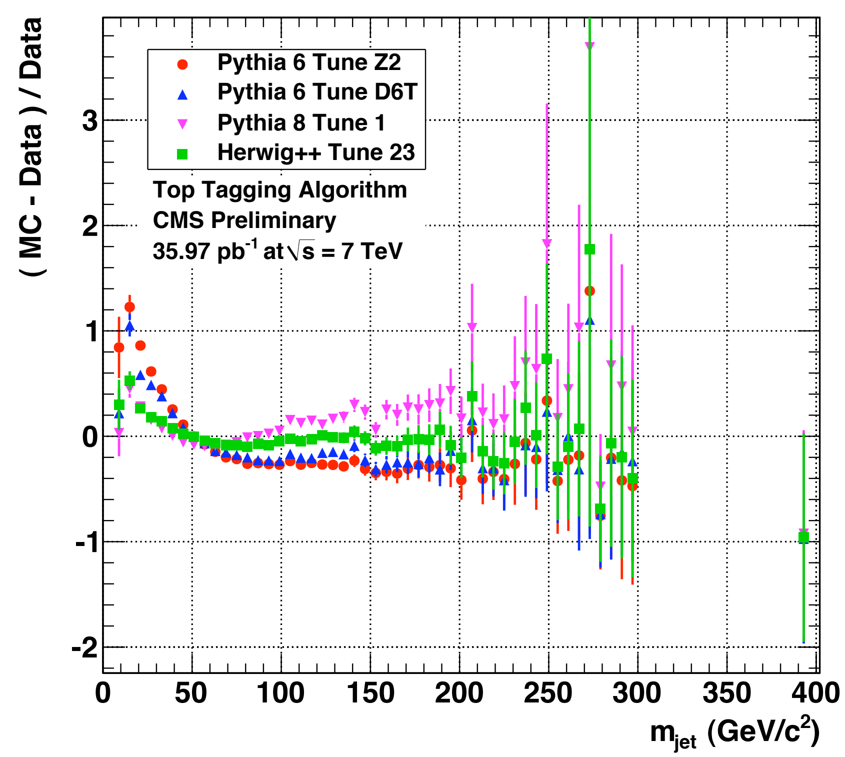


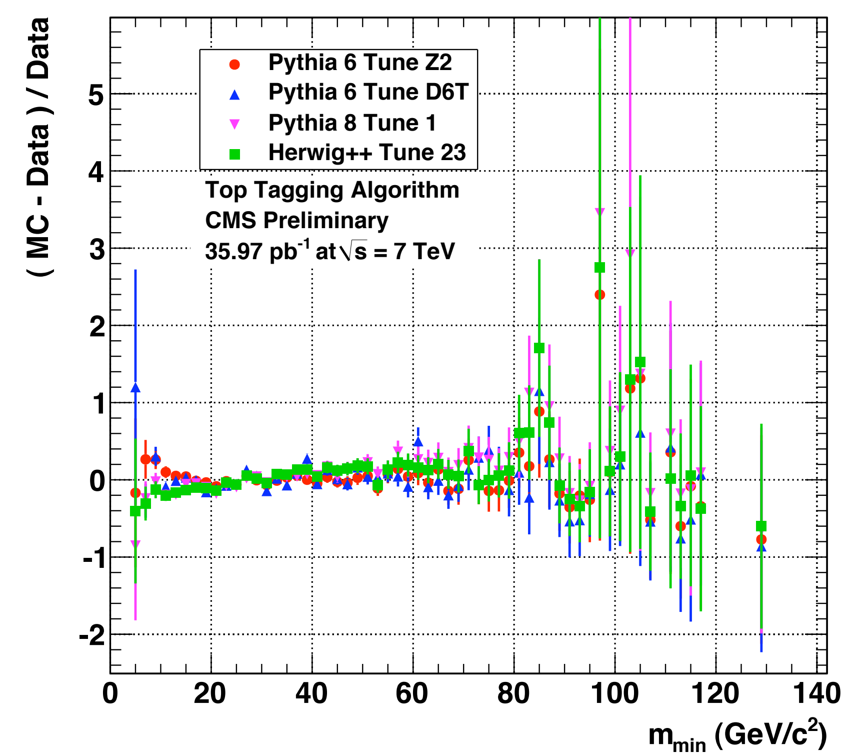
MinMass



Percent difference:

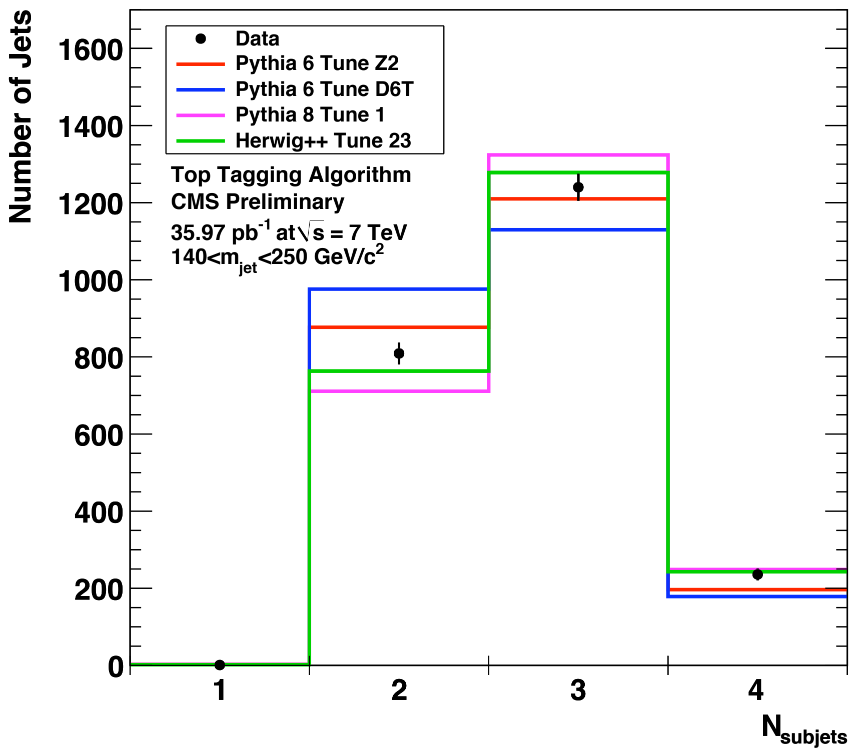




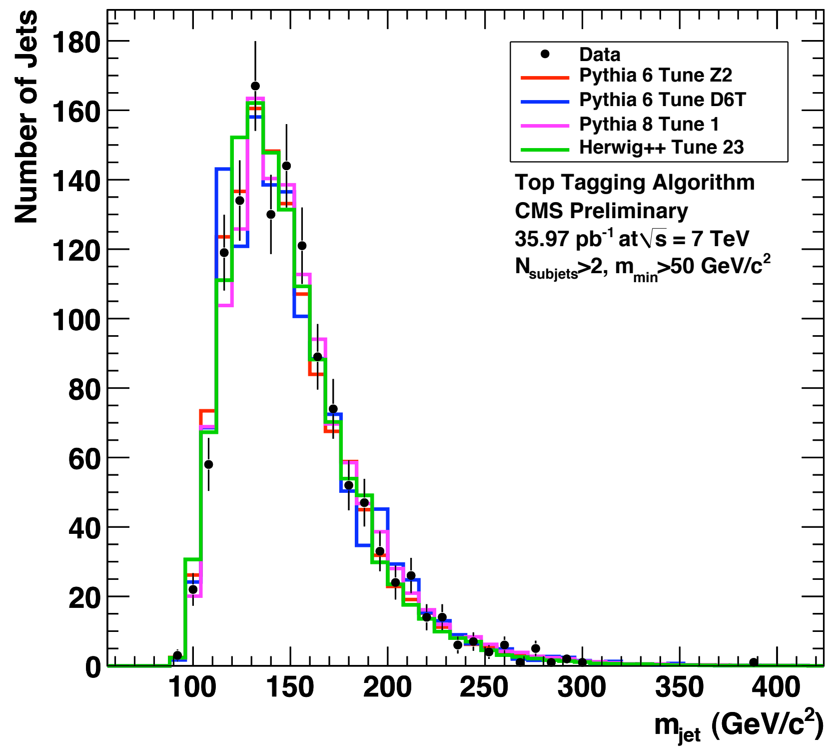


Apply "N-1" cuts

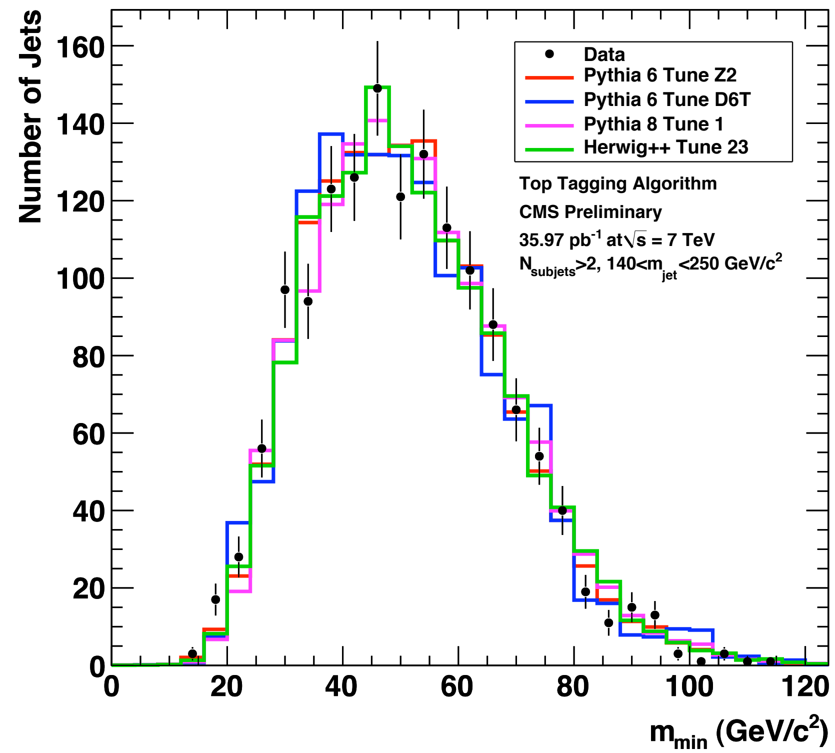
Number of Subjets (Apply jet mass cut)



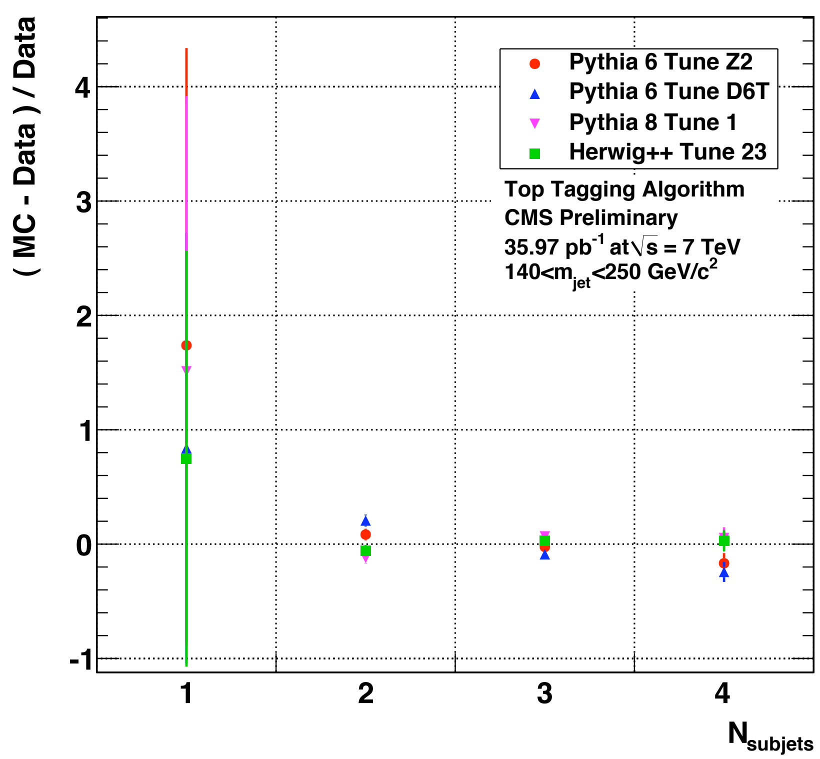
Jet Mass (Apply m_{min} and $N_{subjets}$ cut)

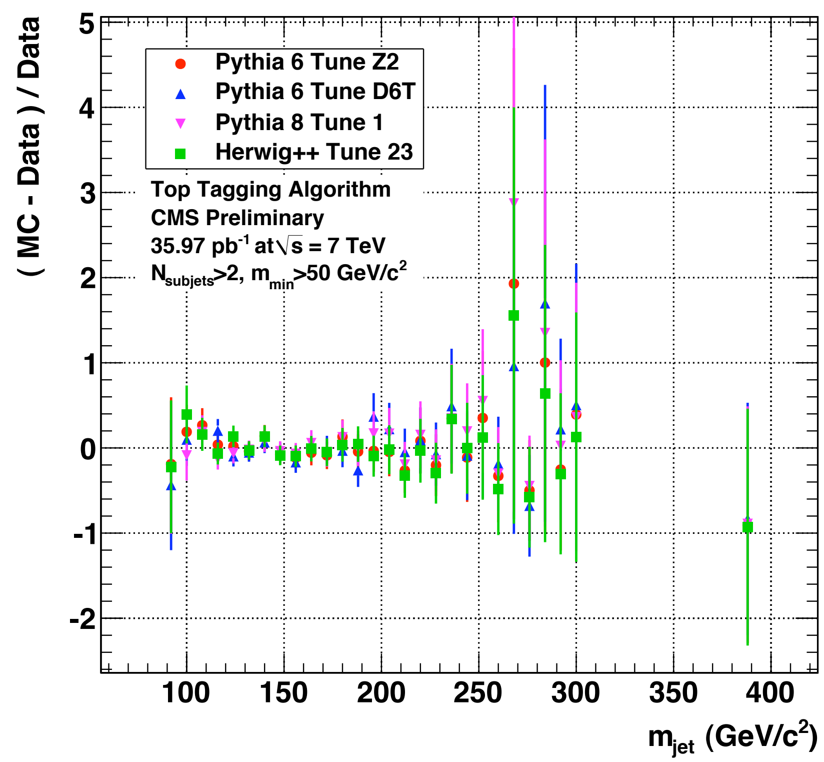


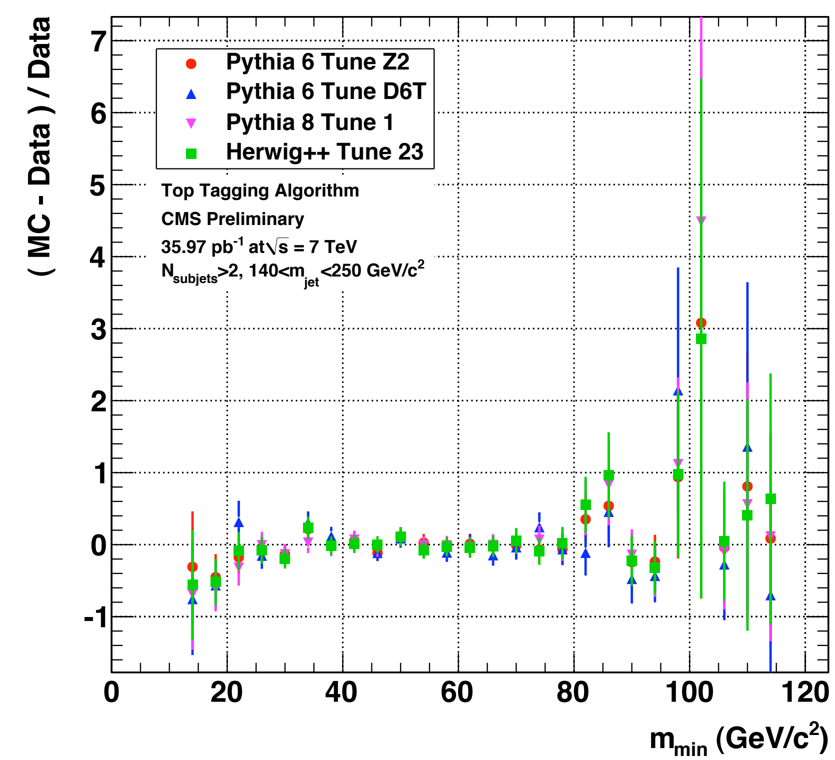
MinMass (Apply m_{jet} and $N_{subjets}$ cut)



Percent difference:



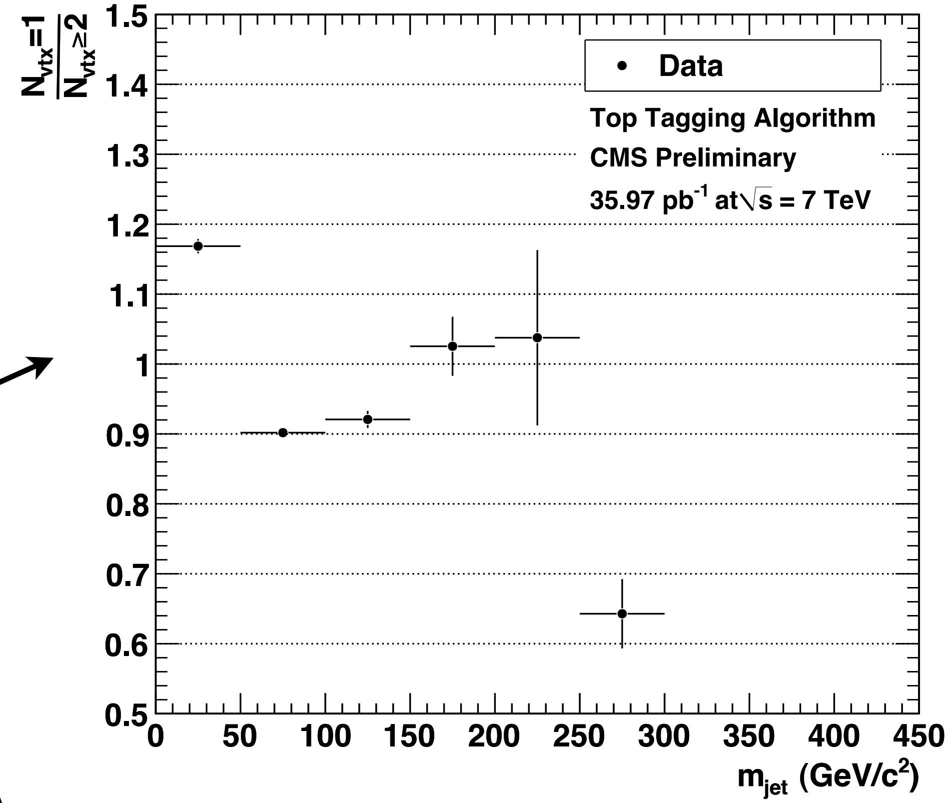


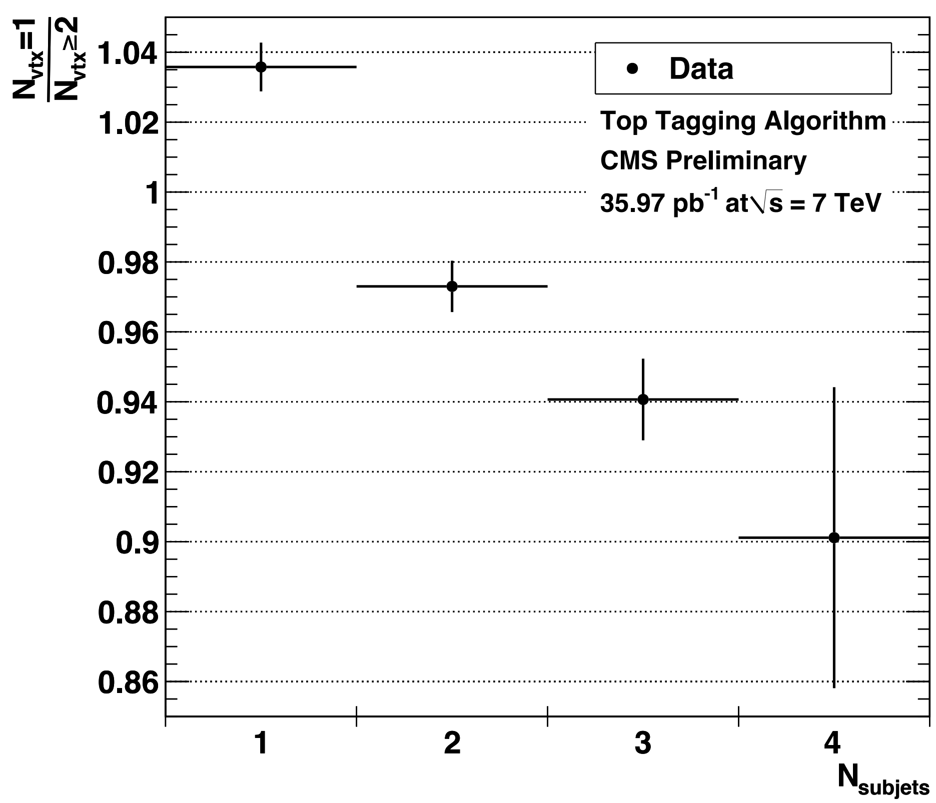


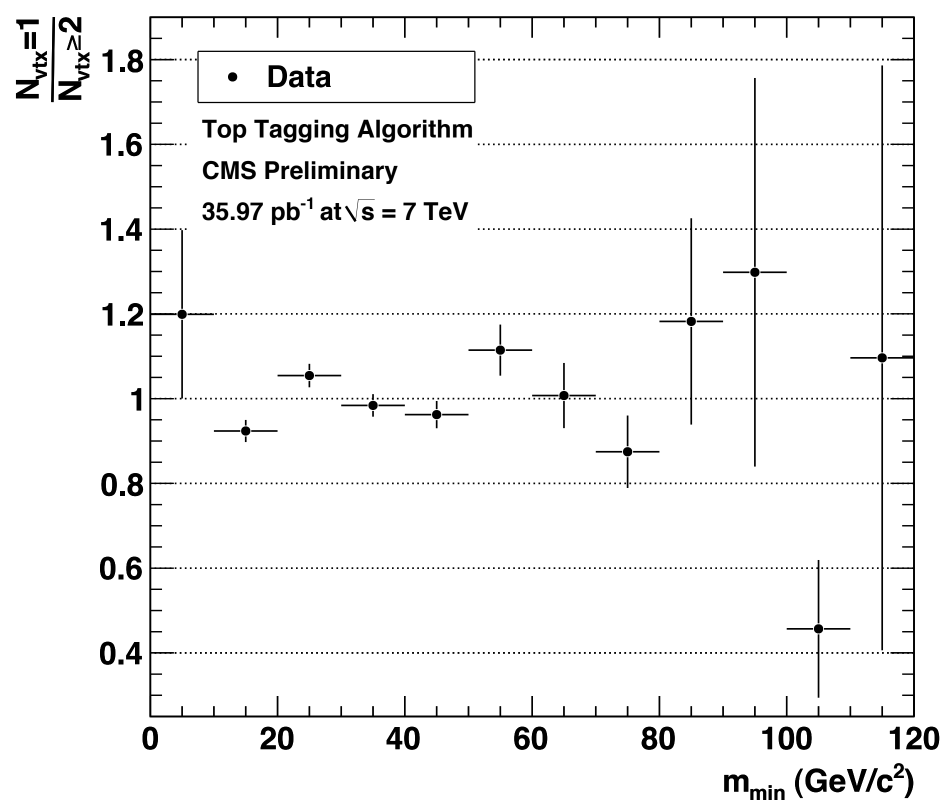
Pileup Dependance

$$\text{Ratio} = \frac{\text{Events with one vertex}}{\text{Events with more than one vertex}}$$

- Measured with data
- Pileup increases the jet mass distribution
- MinMass is less susceptible to pileup
- Events with pileup have more subjets







Conclusion

- Top mistag rate has been measured from data
  - Background estimations for physics measurements can now be completed using data-driven techniques
- First measurements have been made with CMS data and compared to Monte Carlo
  - Reasonable agreement
  - Some dependency on the tune and shower model