Study of Jet Substructure in pp Collisions at 7 TeV in CMS

CMS PAS JME-10-013

Salvatore Rappoccio
for the CMS Collaboration

Jet Substructure

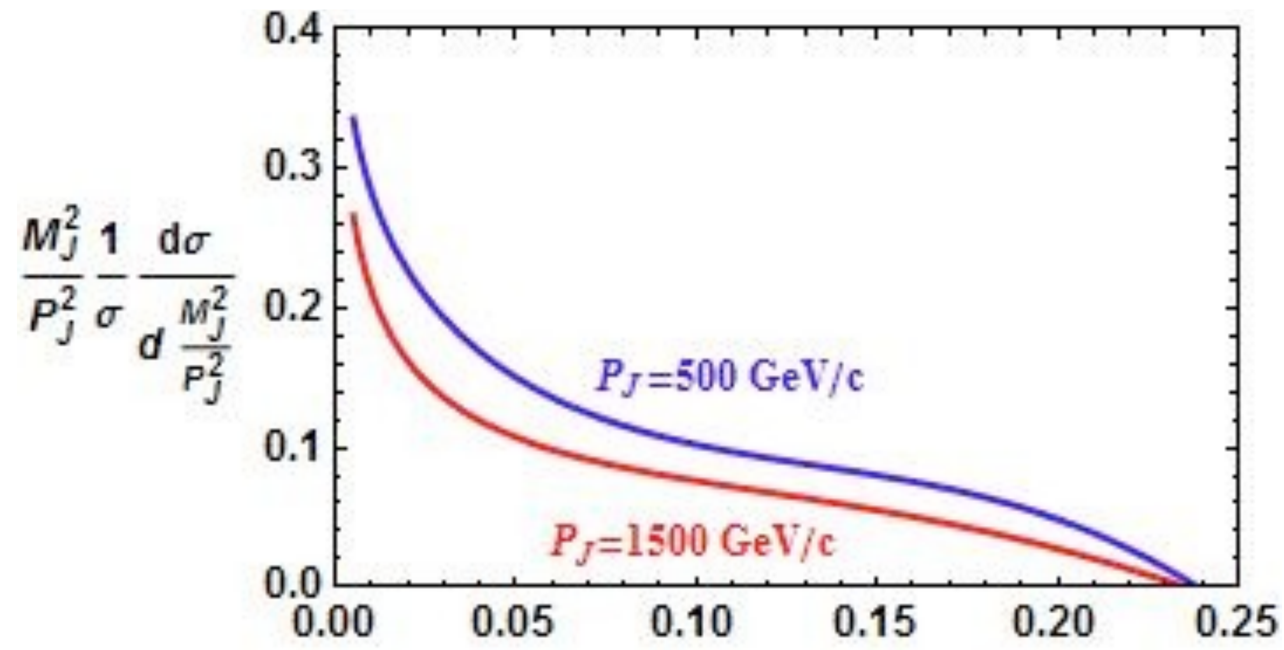
- General concept to recover information from boosted hadronic final states
- General strategy: Use the principle that boosted hadronic objects have a mass scale and different kinematics than QCD
  - See e.g. Ellis et al arXiv: 0712.2447v1

Finite-size effects from cutoff

Scales ~linearly with momentum

Log. divergence at low mass

$$\langle M_J^2 \rangle_{NLO} \simeq \overline{C}\left(\frac{p_J}{\sqrt{s}}\right)\alpha_s\left(\frac{p_J}{2}\right)p_J^2 R^2,$$



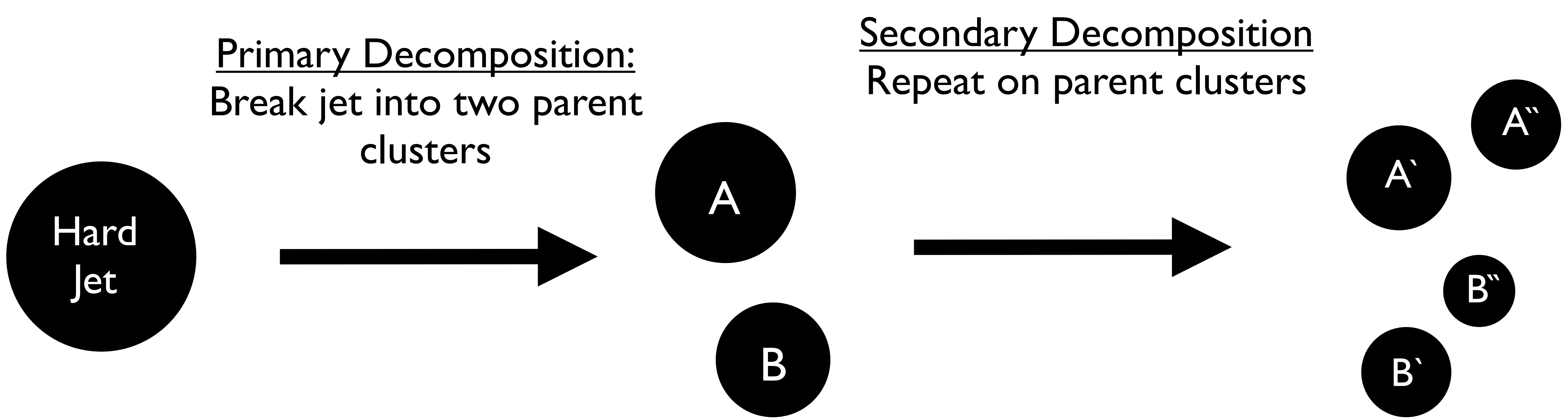
Graphic taken from Steve Ellis

Jet Substructure

- Examined two of the algorithms just mentioned
  - JHU "top tagger"
    - Targeted tagger for identifying boosted top quarks (3-body decay)
    - Phys.Rev.Lett.101:142001,2008, arXiv:0806.0848
  - University of Washington "jet pruning"
    - General-purpose tool to "clean up" jets
    - Phys. Rev. D80 (2009) 051501, arXiv:0903.5081
    - Can be used for any decay topology
      - We look at boosted W's in our PAS (2-body decay)
- Performance of these algorithms in dijet data is examined

Top Tagging Details

- Based on Kaplan et al. (arXiv:0806.0848)
- Cluster particle flow candidates using Cambridge Aachen
- Reverse the clustering sequence in order to find substructure
- Subjets must satisfy two requirements
  - Momentum fraction criterion: $pTsubjet > 0.05 \times pThard\ jet$ ← Removes soft clusters
  - Adjacency criterion: $\Delta R(A, B) > 0.4 - 0.0004 \times pT$ ← Removes wide angle clusters
- Iterative process - throw out objects that fail momentum fraction cut and try to decluster again

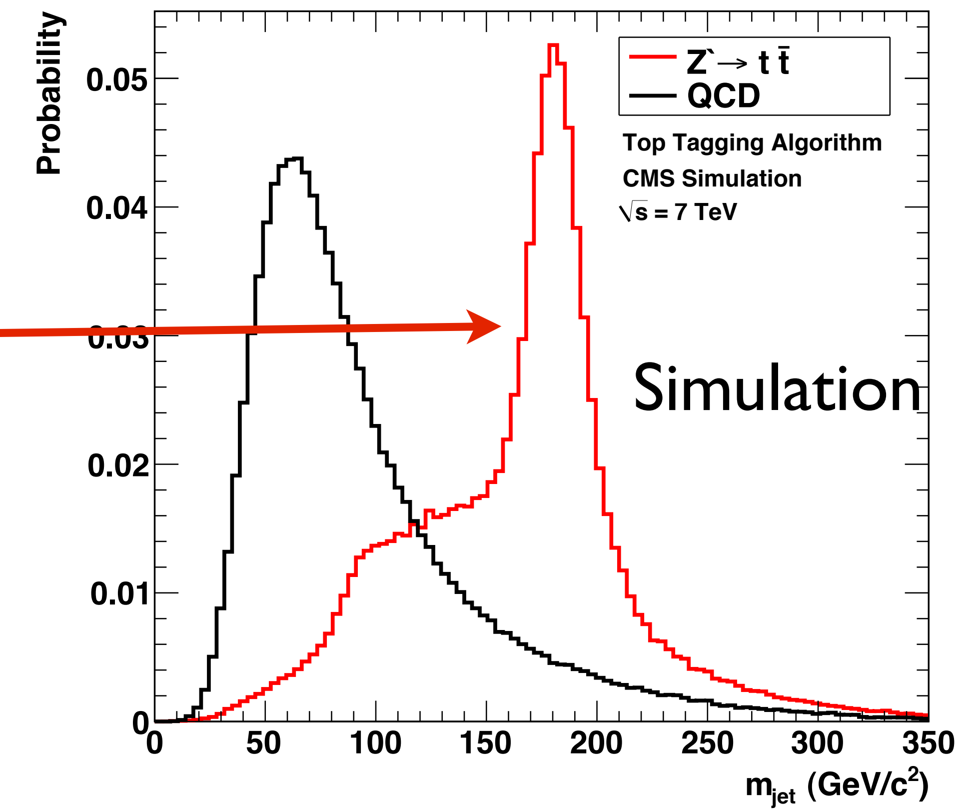


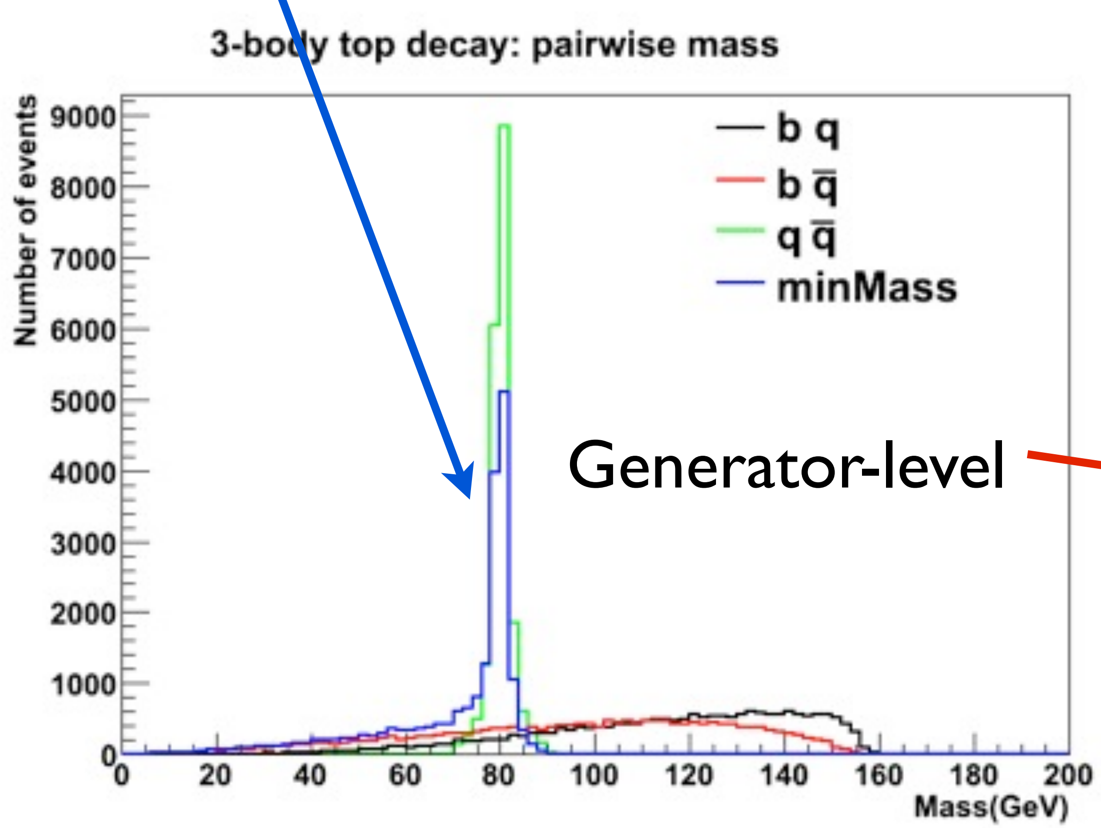
Require >= 3 subjets

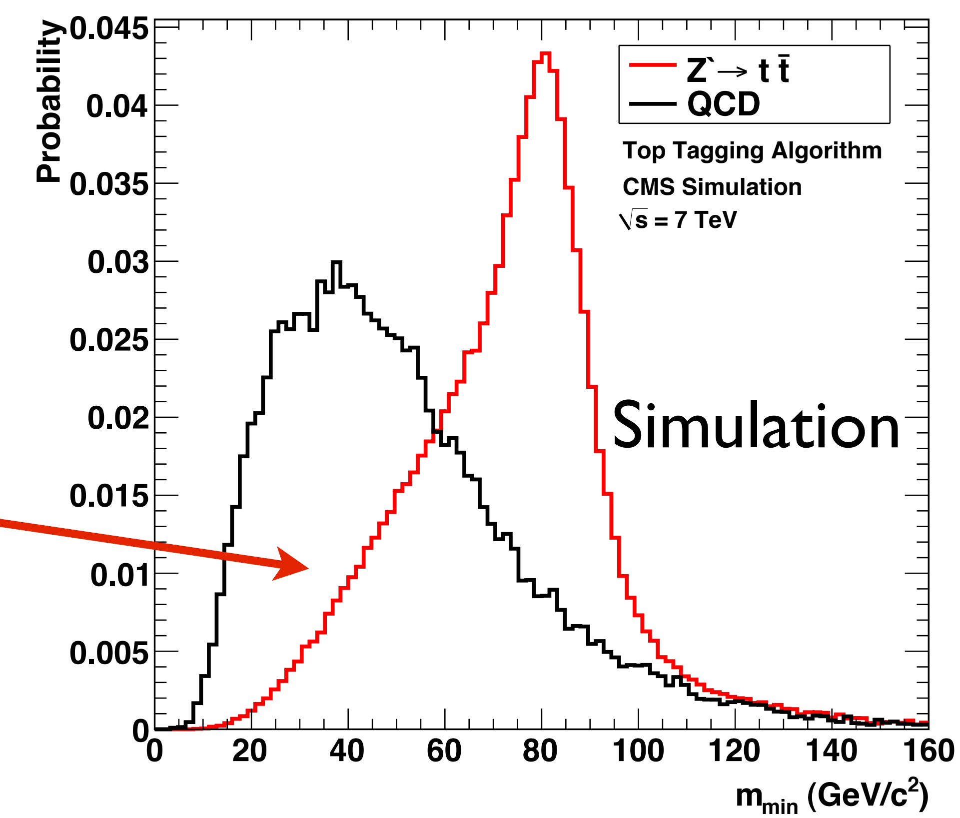
Top Tagging Details

- Discriminating variables:
  - Number of subjets: 3 or 4
  - Top Mass: Approximated by jet mass
    - Mass in 100-250 GeV/c^2
  - W Mass: Approximated by min pairwise mass
    - Min mass > 50 GeV/c^2

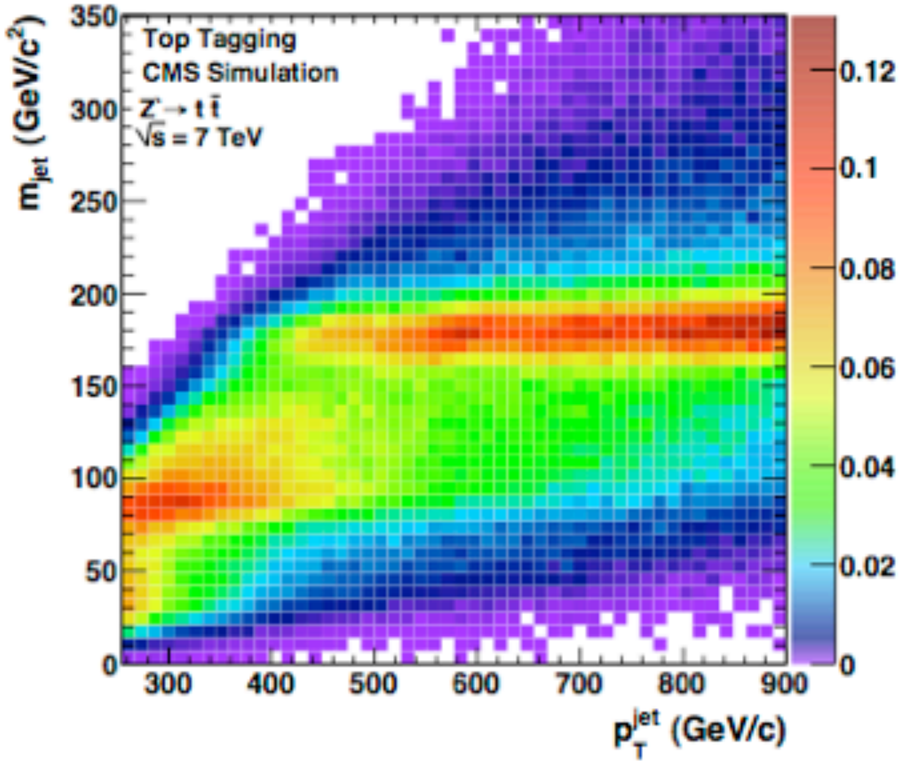
$$m_{\min} = \min[m_{12}, m_{13}, m_{23}]$$



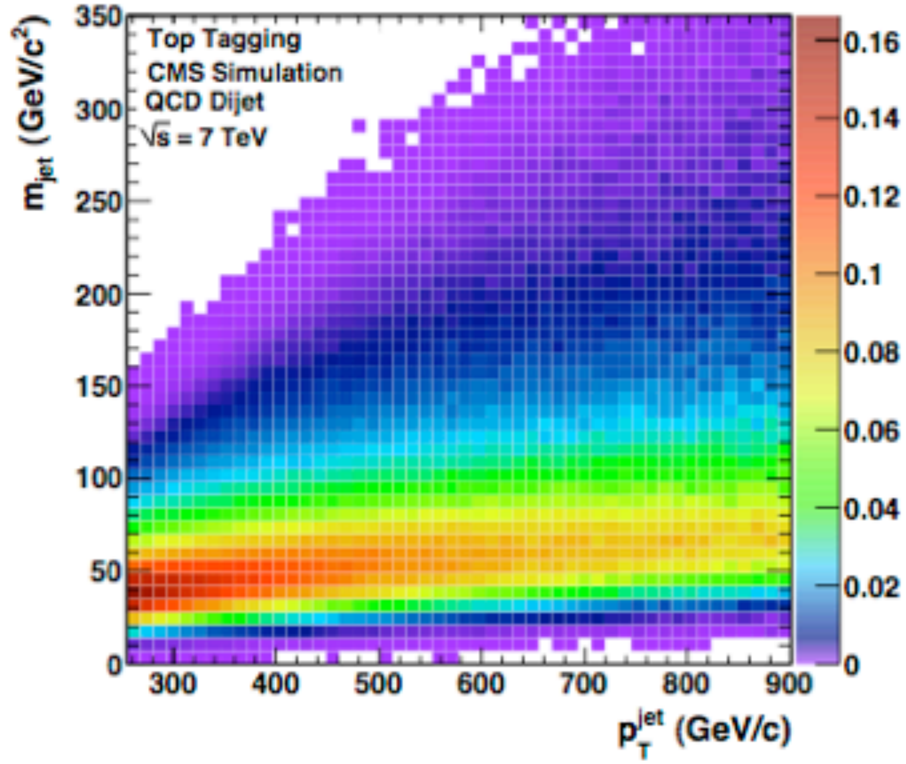




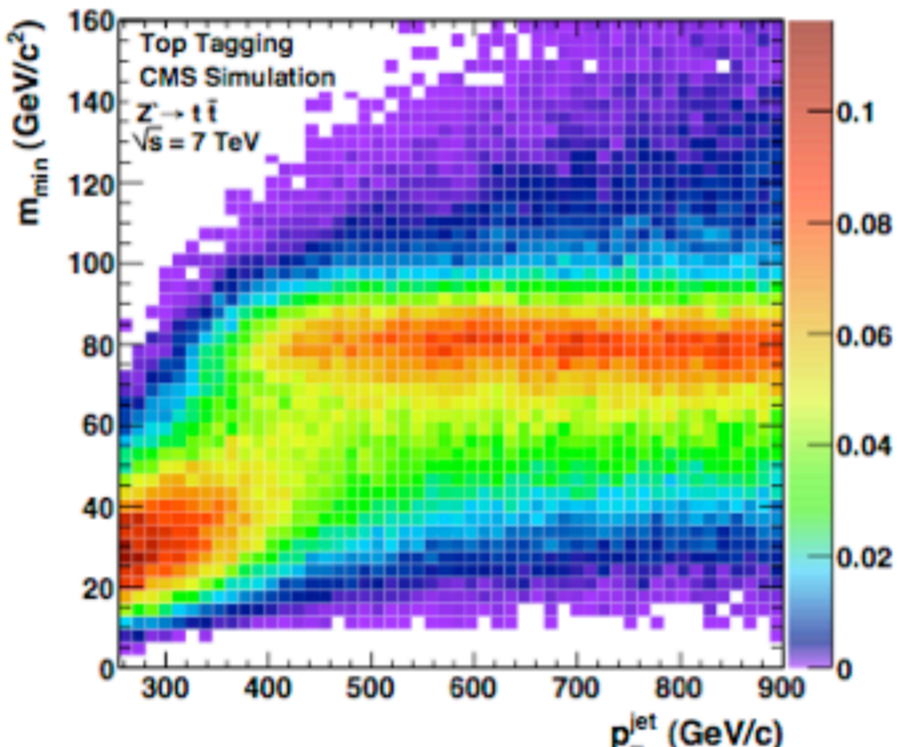
Top Tagging Details



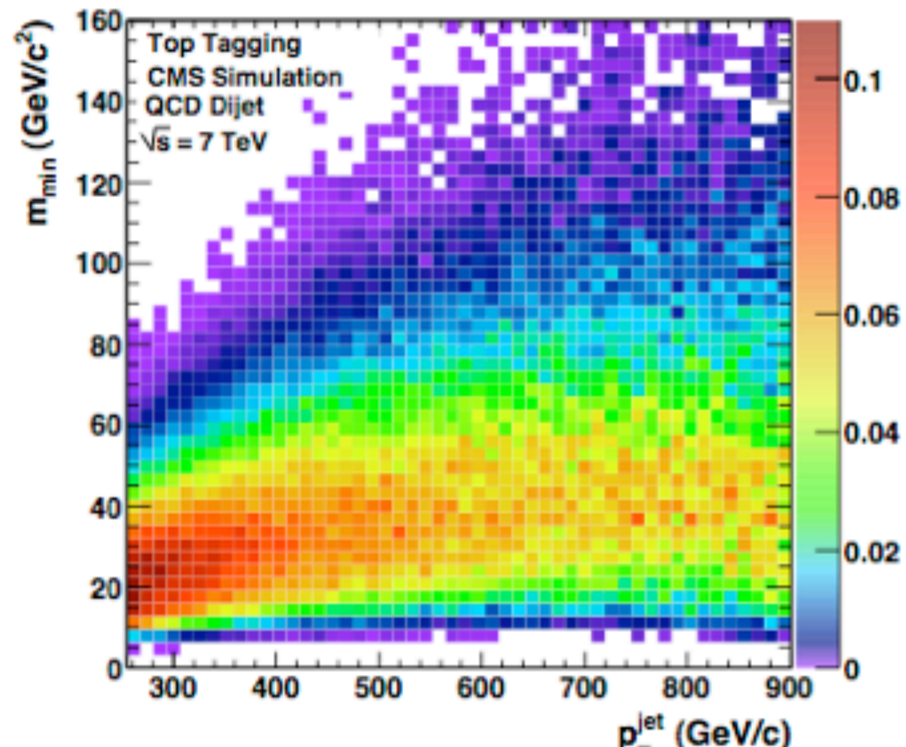
(a)



(b)



(c)



(d)

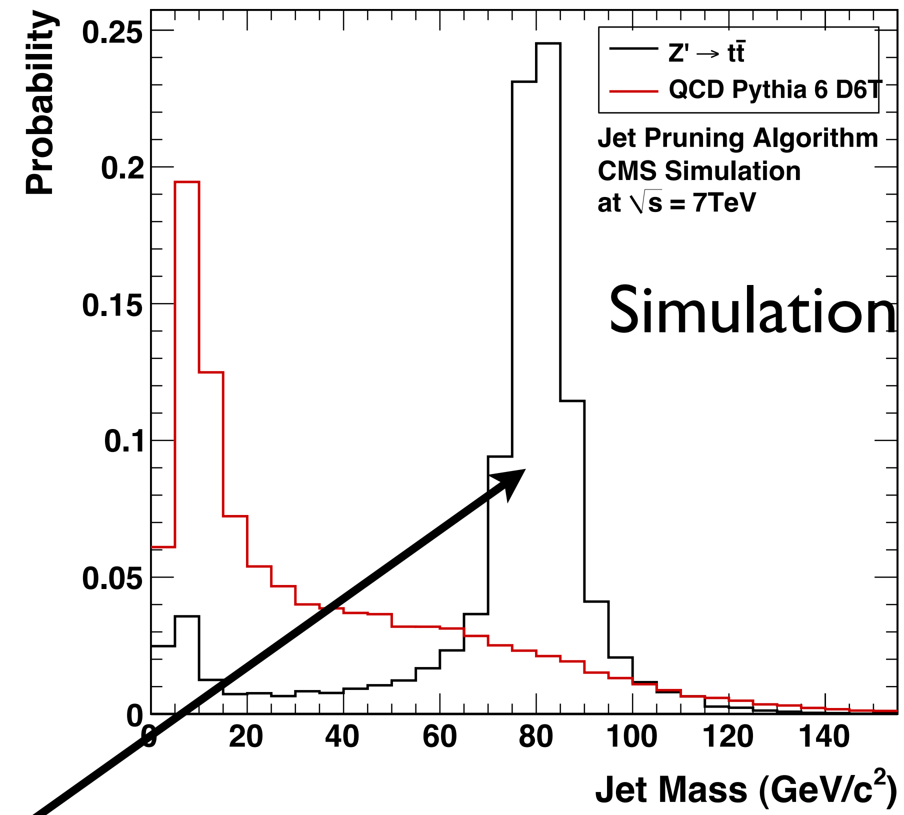
Jet Pruning Details

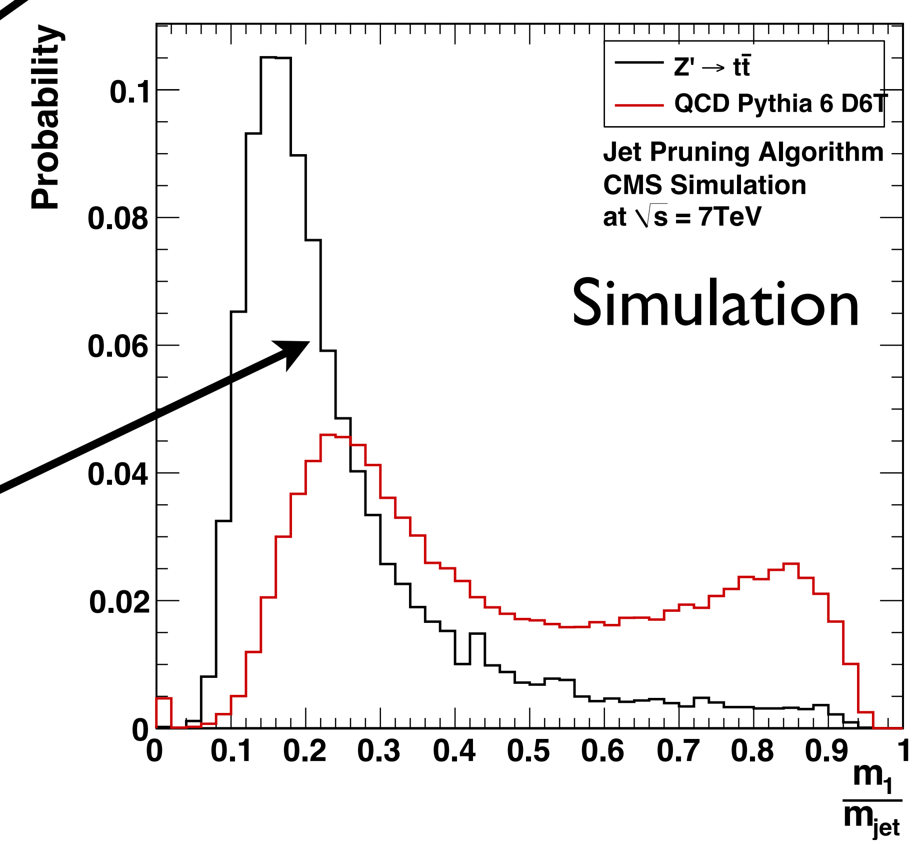
- Ellis et al. (arXiv:0903.5081)
- Improves mass resolution by removing soft, large angle particles from the jet
- Recluster each jet, requiring that each recombination satisfy the following:

$$\frac{\min(p_{T1}, p_{T2})}{p_{Tp}} > 0.1$$

$$\Delta R_{12} < 0.5 \times \frac{m_{\rm jet}}{p_T}$$

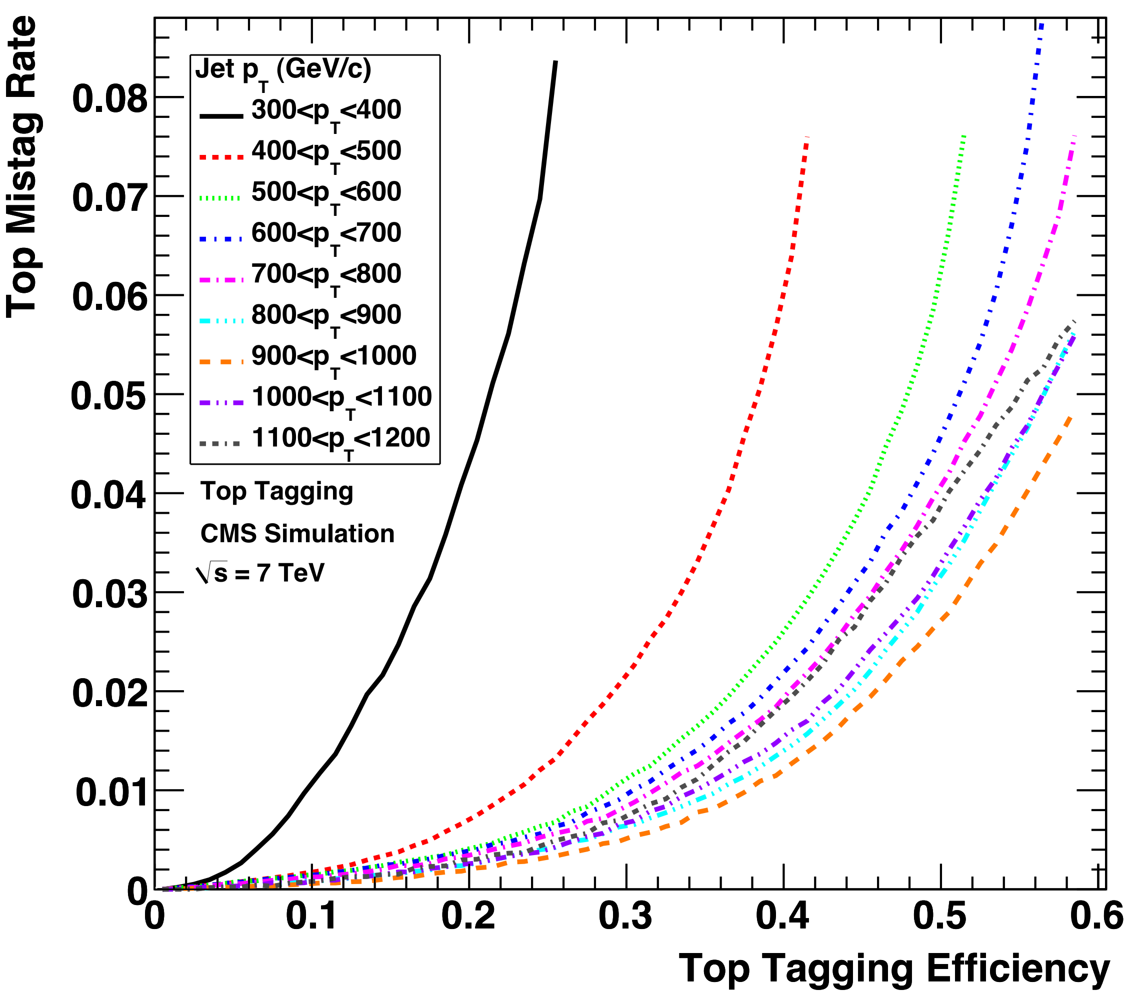
- For W tagging, require:
  - Jet mass in 60-100 GeV/c2
  - Mass drop (mu) < 0.4 $\quad \mu = \frac{m_{j1}}{m_j}$

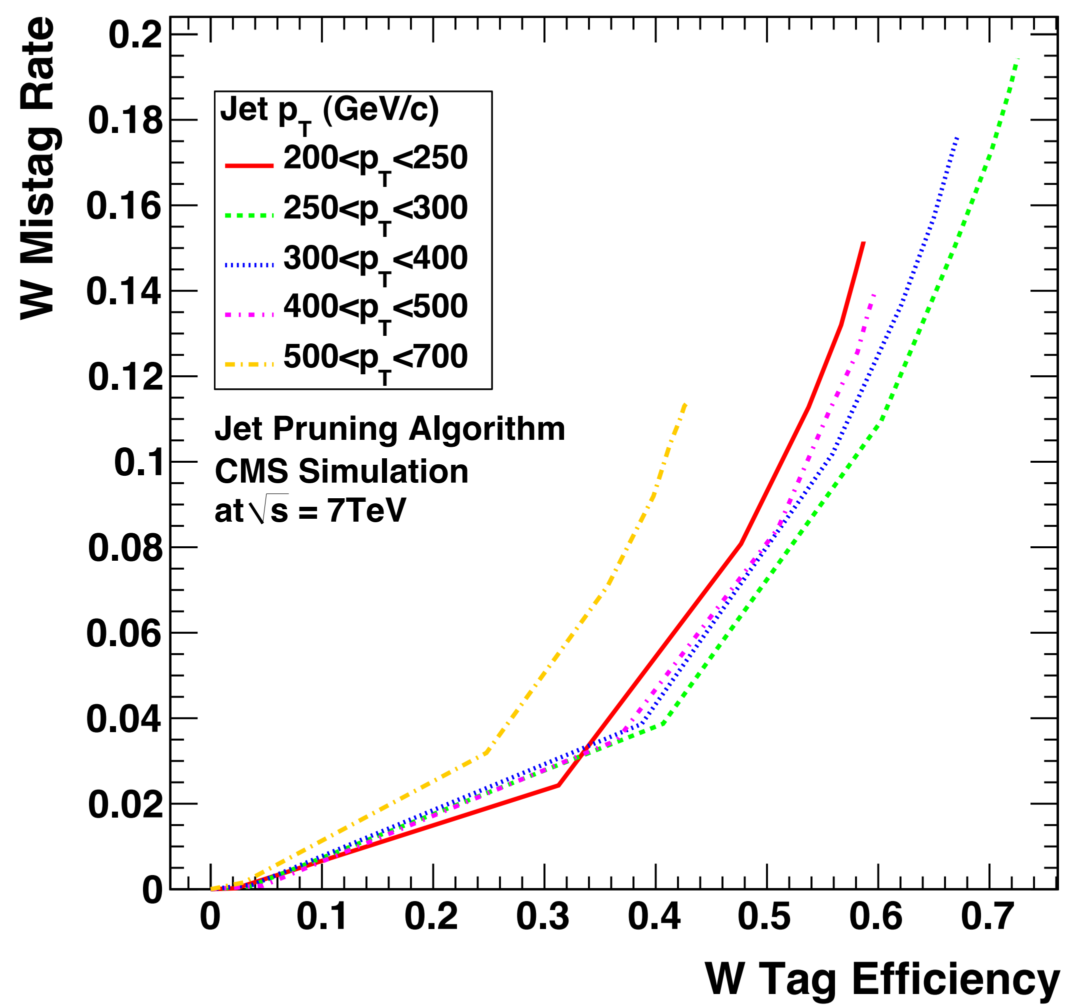




Optimization

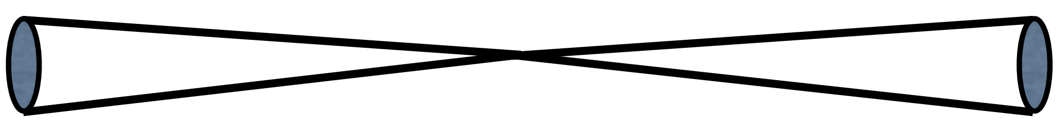
- Extensive optimization effort, documented in AN-10-080
- Used efficiency versus fake rate in MC as figure of merit
- Values shown in previous slides are chosen based on this optimization





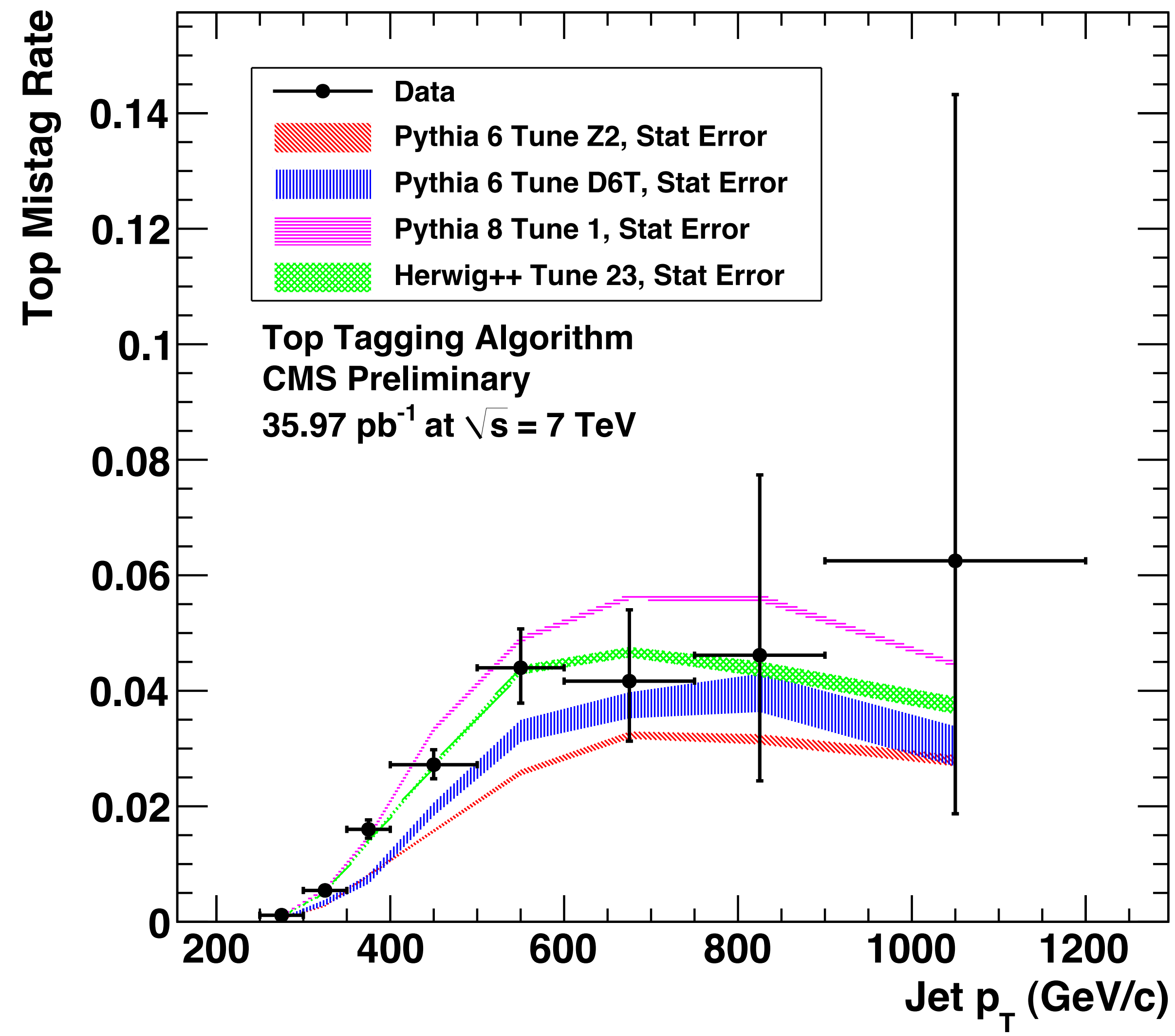
Top Tagging Mistag Rate

- Anti-tag and probe method
  - Randomly select one jet, check if its tagged
  - If the random jet is vetoed, the opposite jet is the probe jet



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

Probe jet

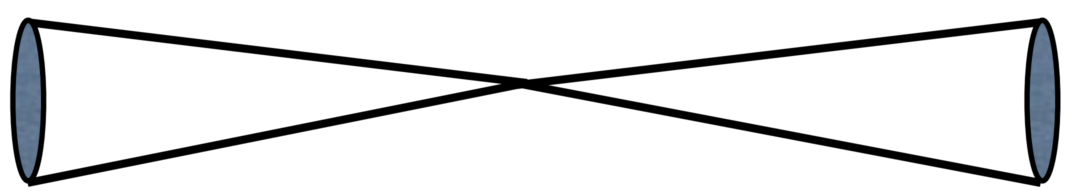


$$\text{Mistag Rate} = \frac{\text{Number of probe jets that are tagged}}{\text{Number of probe jets}}$$

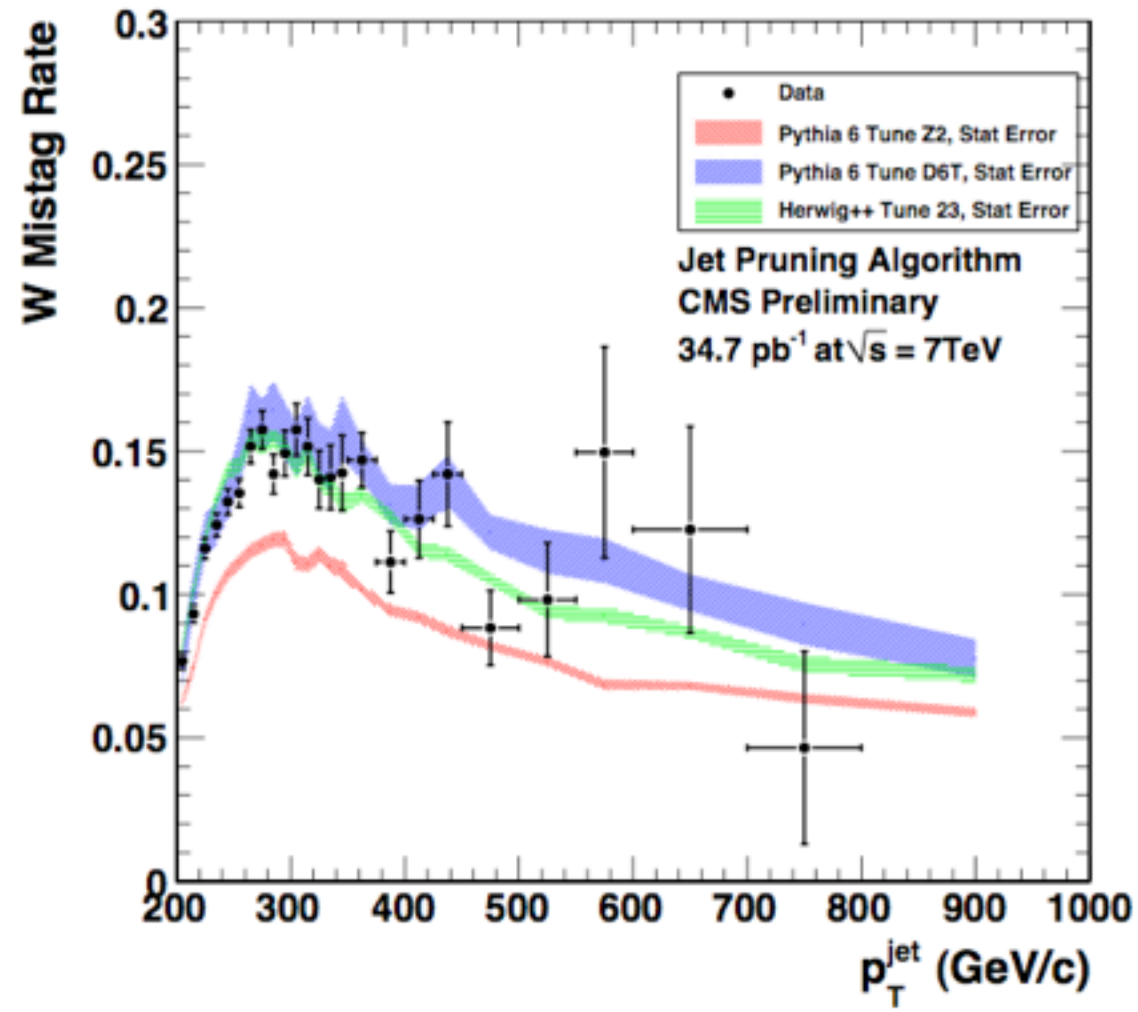
Dependence on shower model and tune, overall good agreement

W Mistag Rate from Data

- Random method
  - Randomly select one jet
  - That is the probe



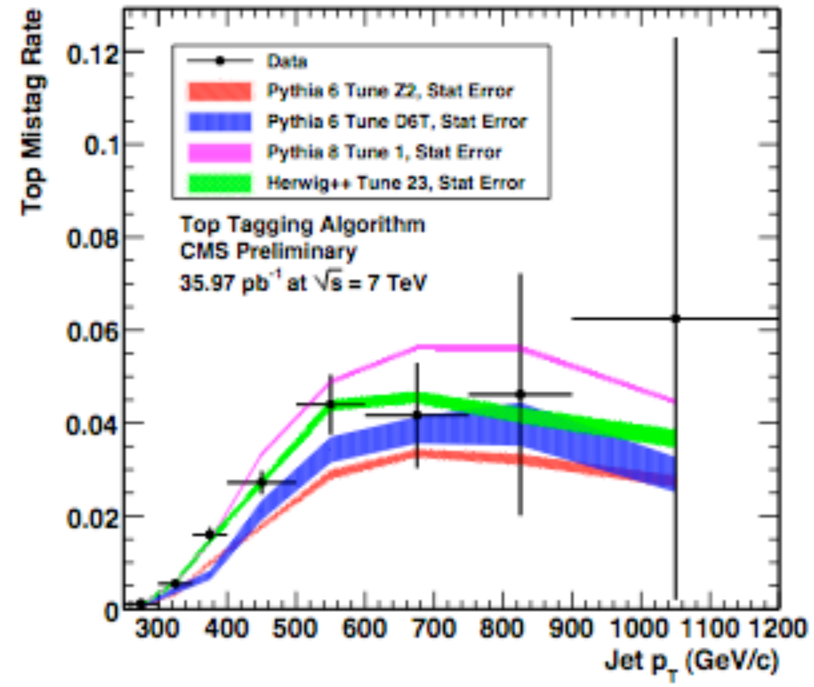
$$\text{Mistag Rate} = \frac{\text{Number of probe jets that are tagged}}{\text{Number of probe jets}}$$



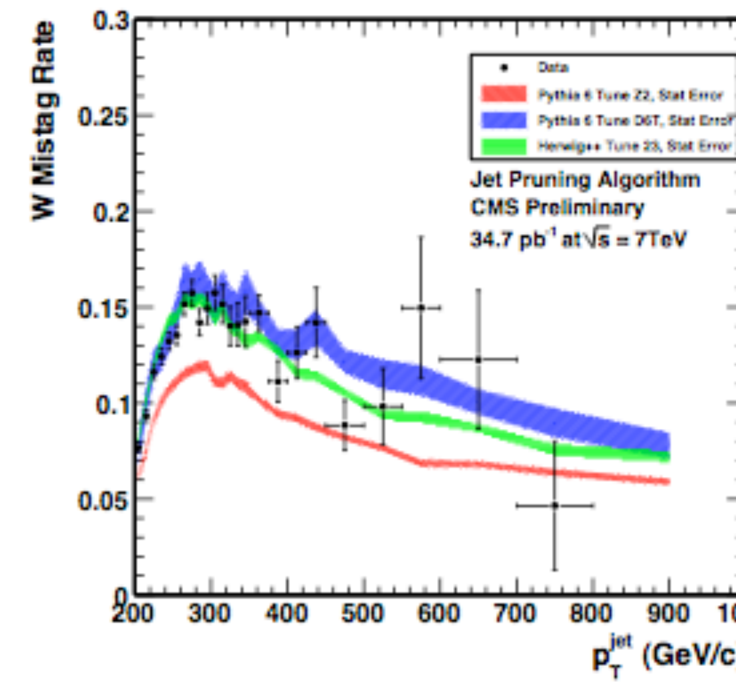
Dependence on shower model and tune, overall good agreement

Efficiency (MC) and Mistag (Data)

Mistag rates from dijet data

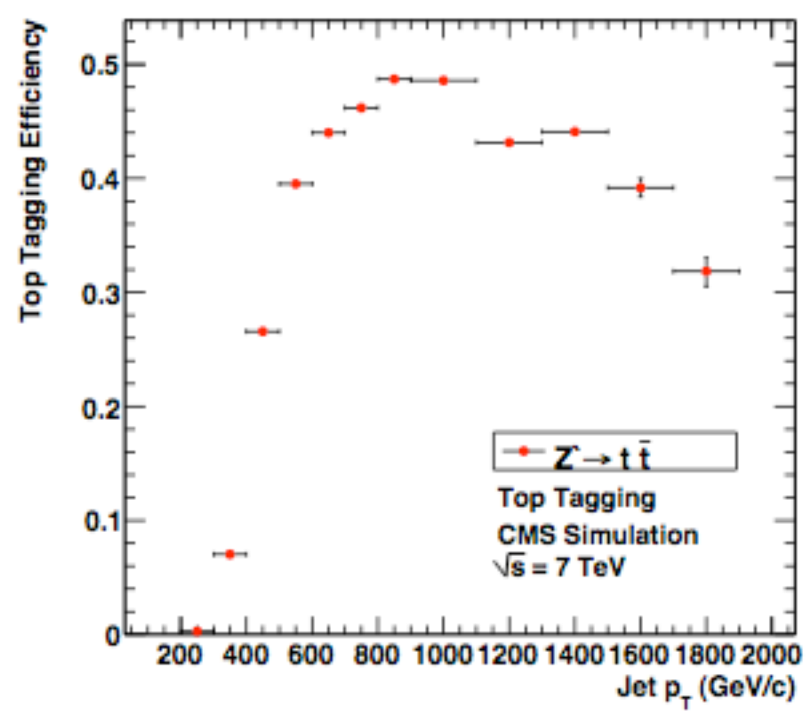


(a)

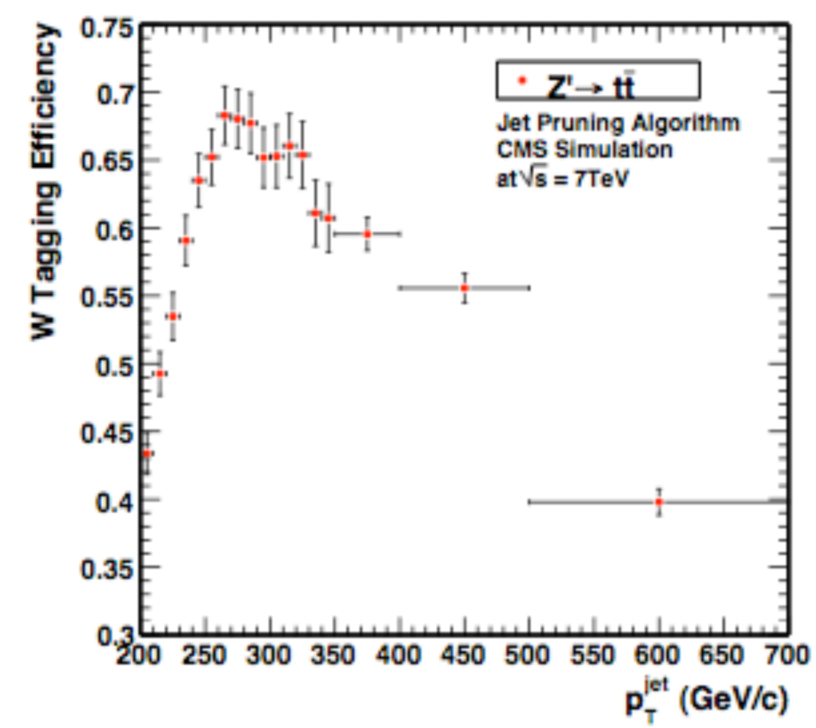


(b)

Efficiency from Z' ->tt in MC



(c)

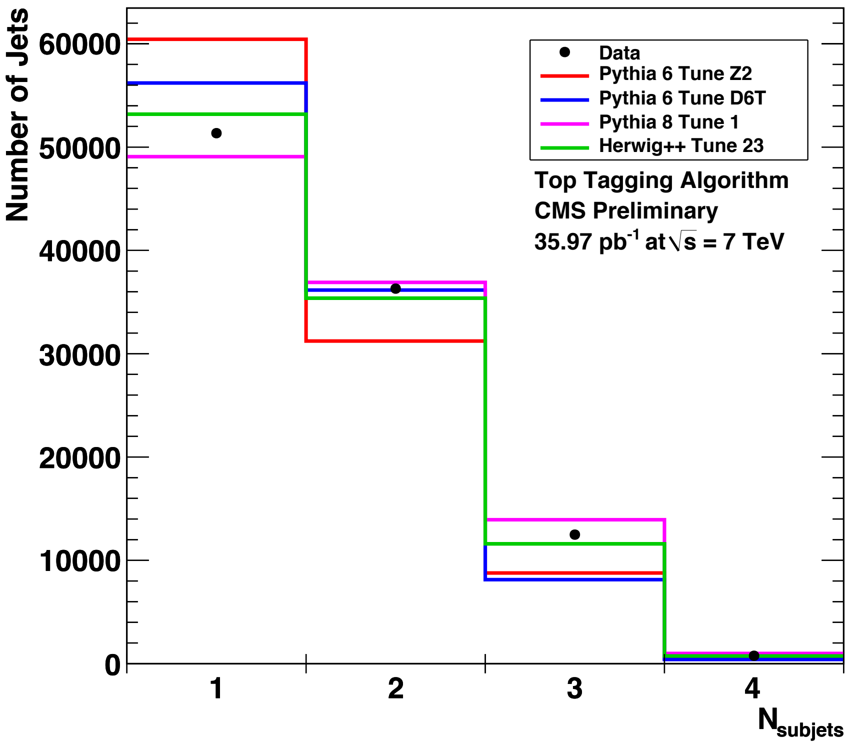


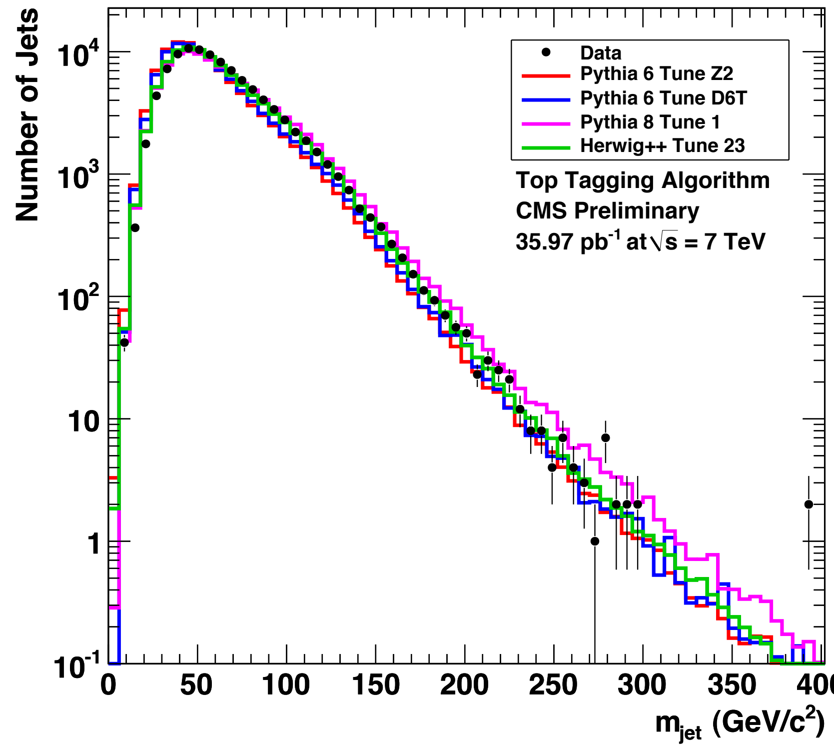
(d)

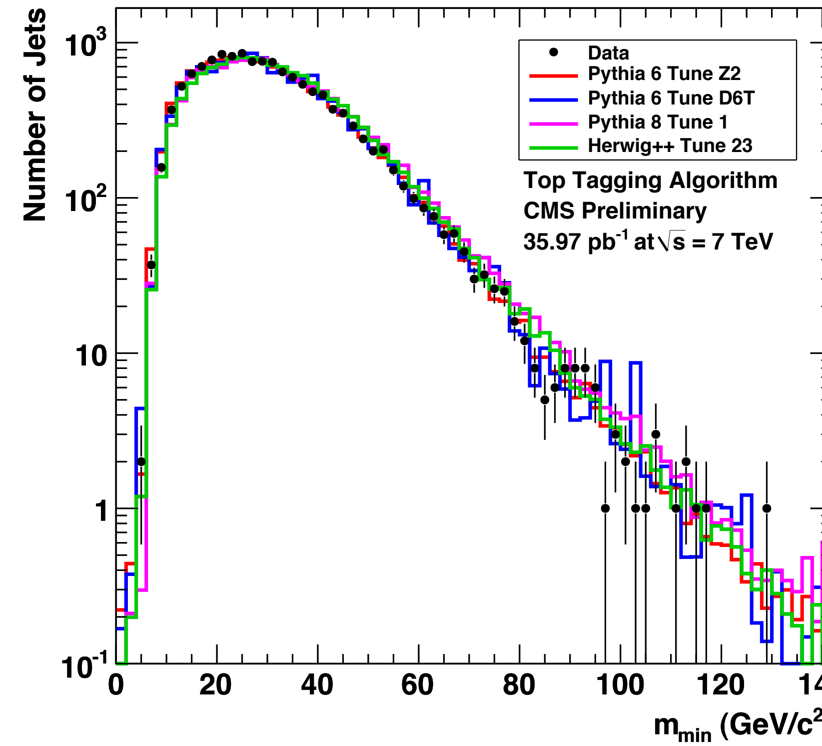
Commissioning Plots

- Compare data to simulation of various MC's and tunes
  - PYTHIA6 Z2, D6T
  - PYTHIA8 1
  - HERWIG++ 23
- Show bare distributions and ratios
- Show "N-1" distributions
  - Apply all but one selection
  - Exception: N_subjets, only require jet mass
  - Min mass undefined for N_subjets < 3

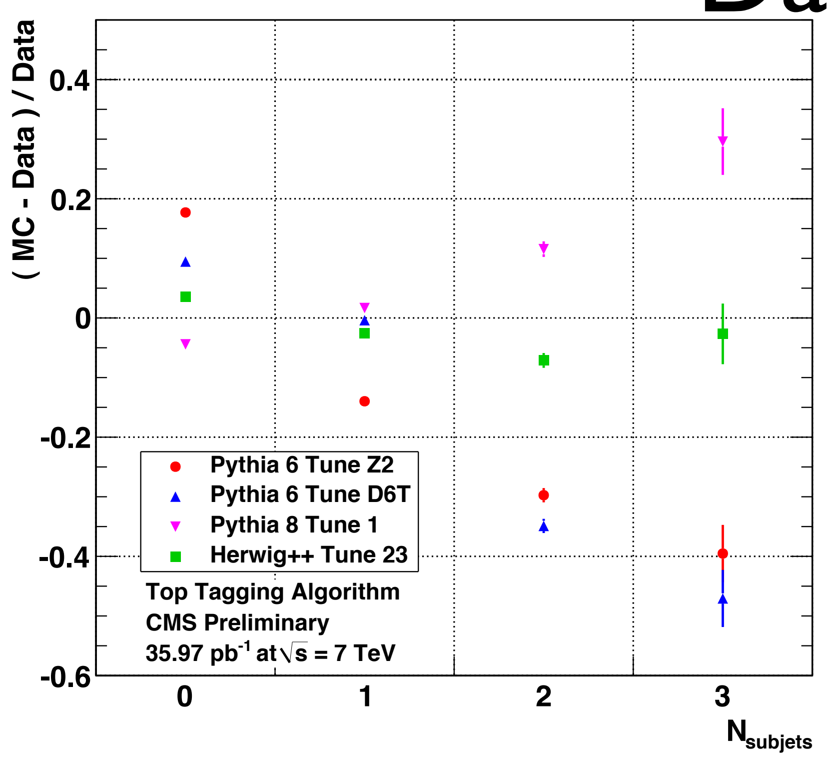
Top Tagging Inclusive Comparison

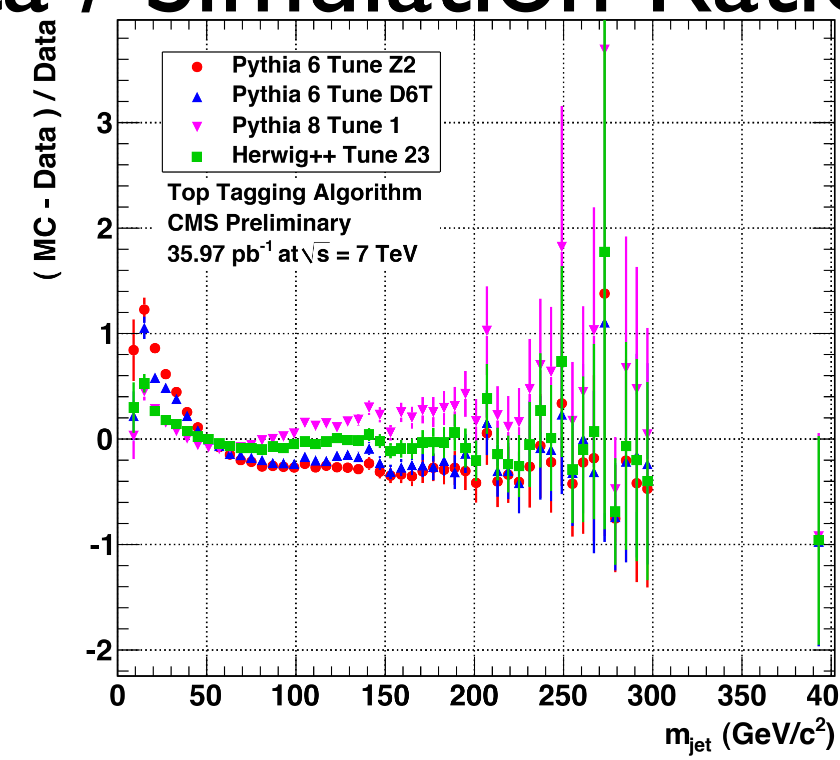


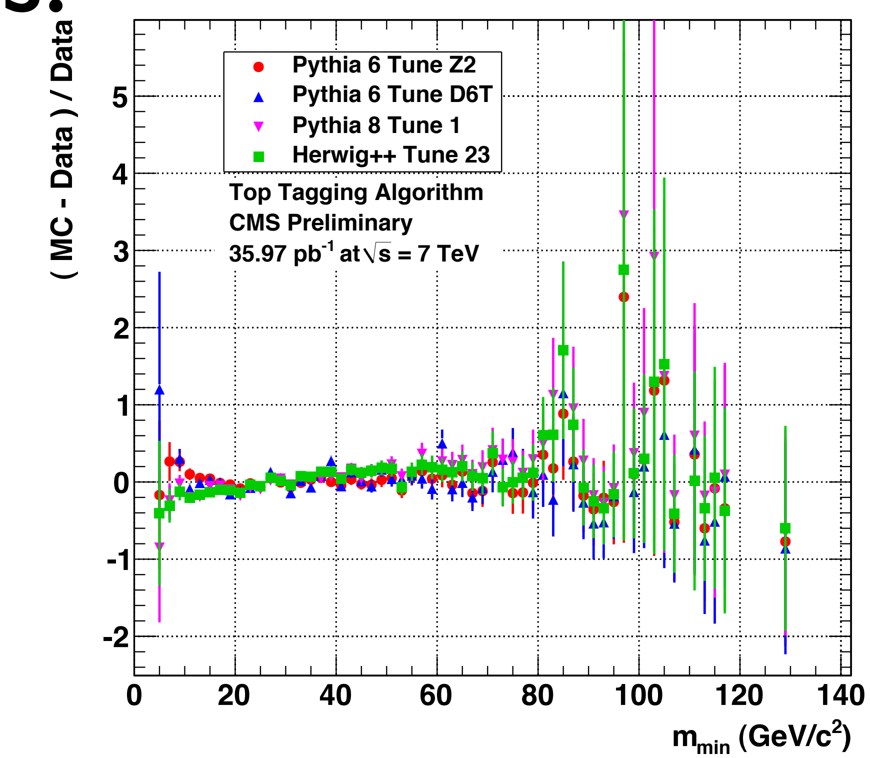




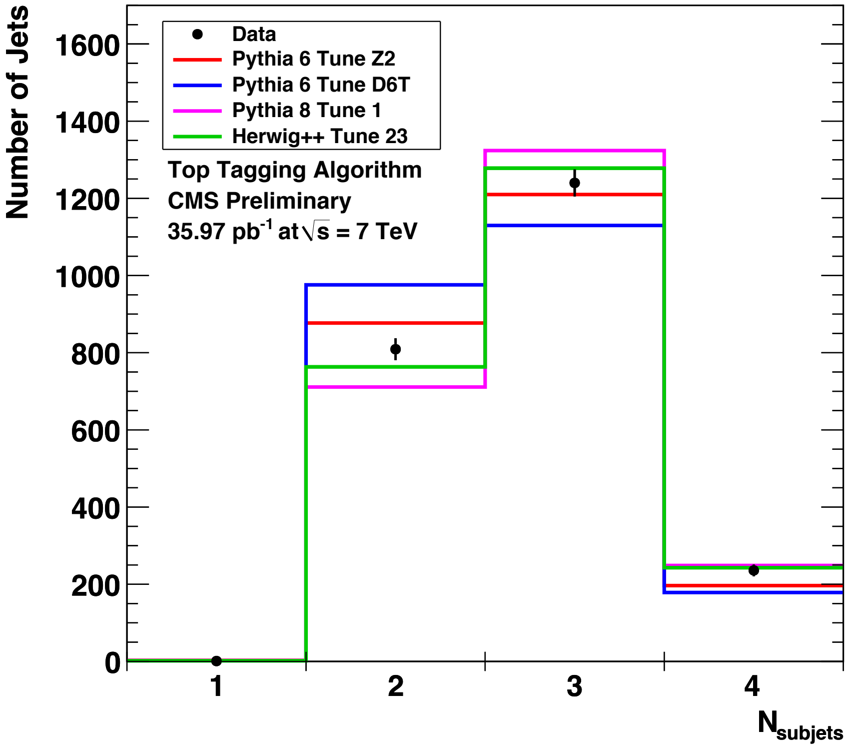
Data / Simulation Ratios:

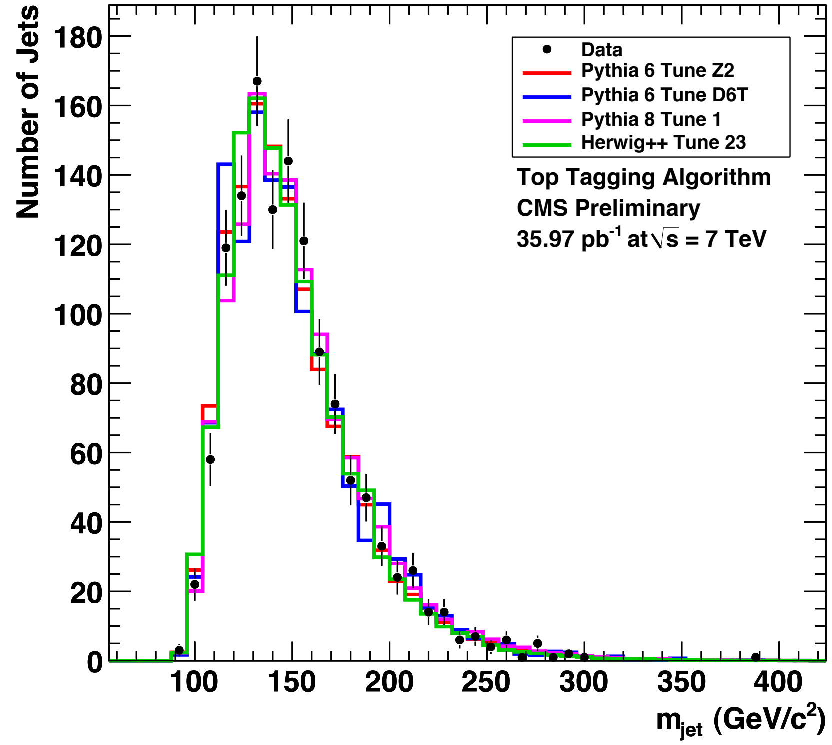


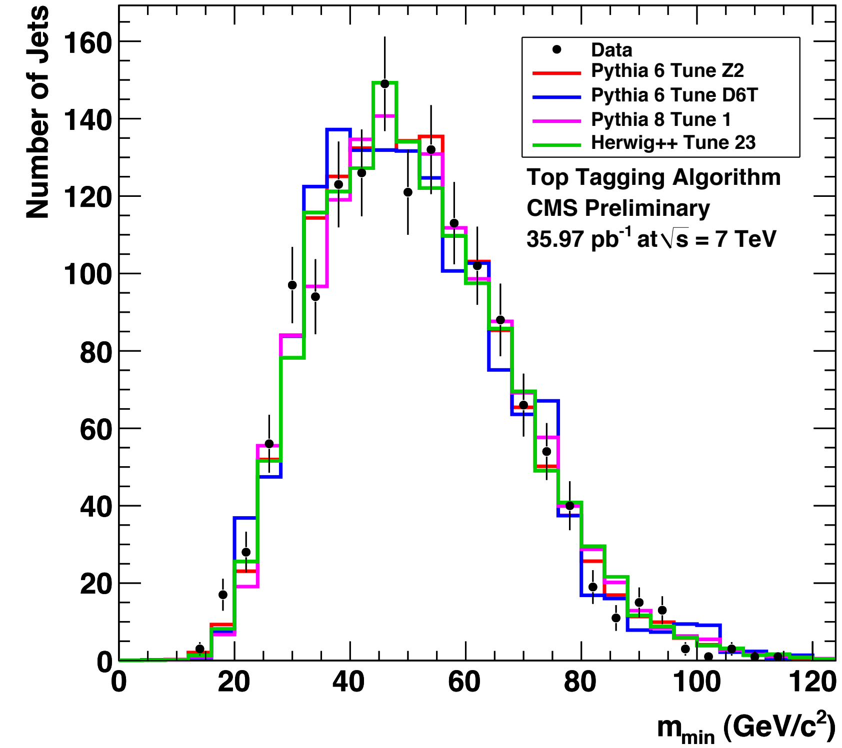




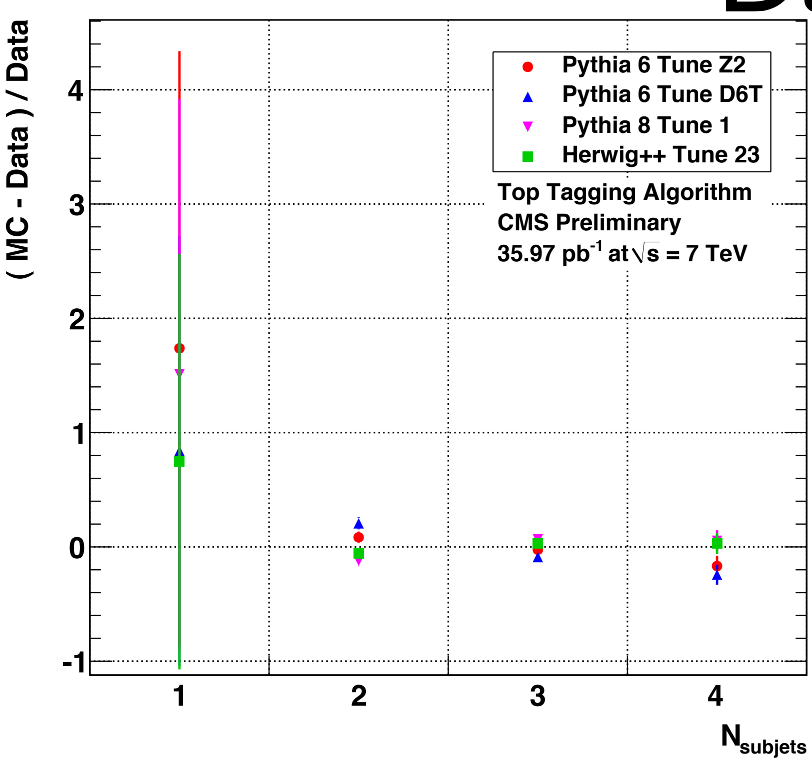
Top Tagging "N-1" Comparison

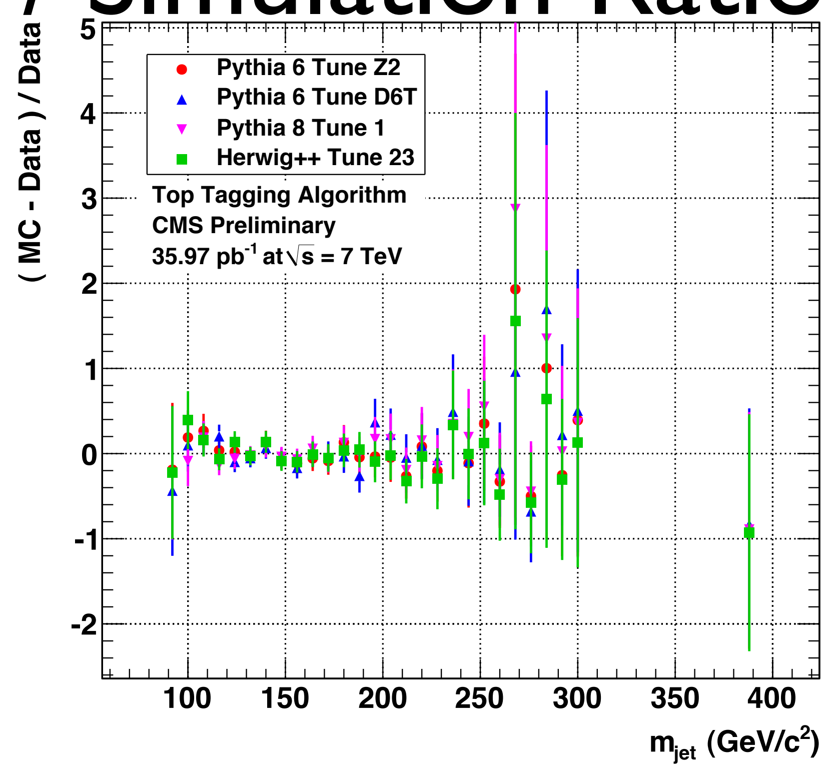


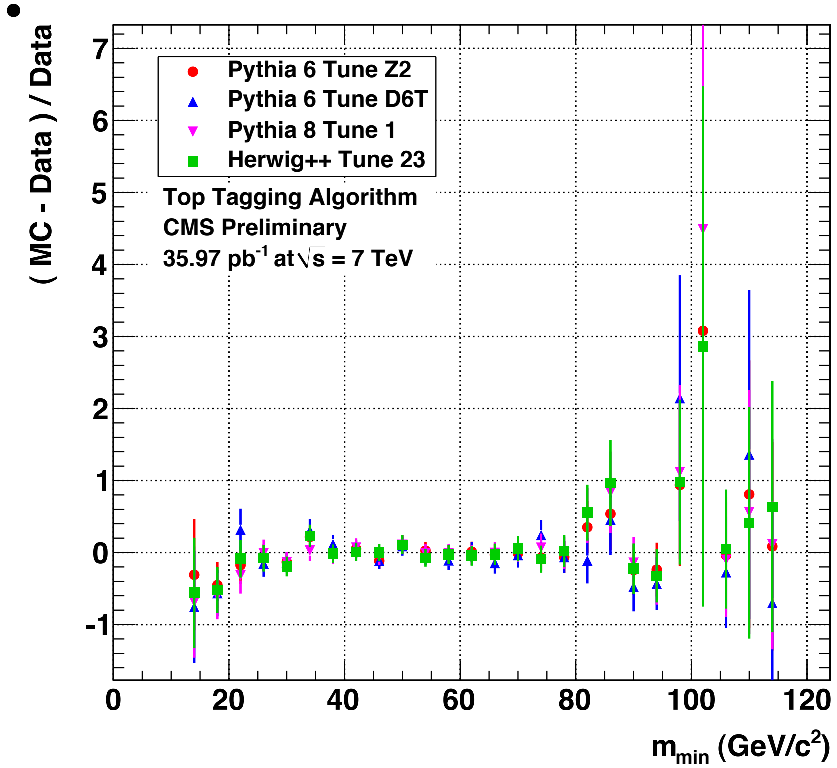




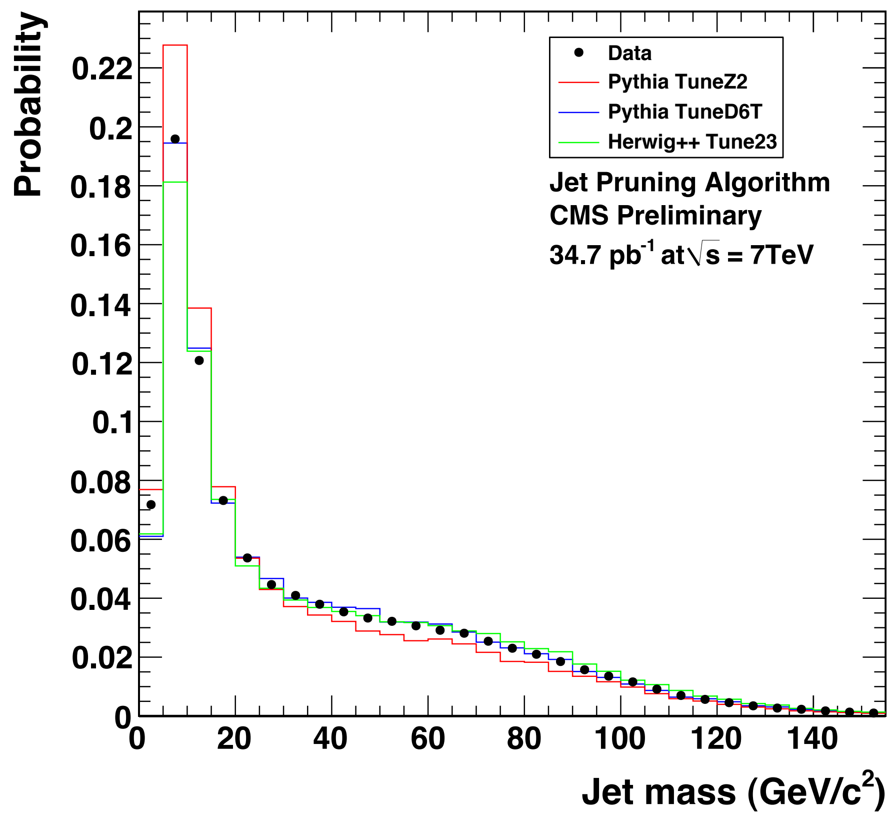
Data / Simulation Ratios:

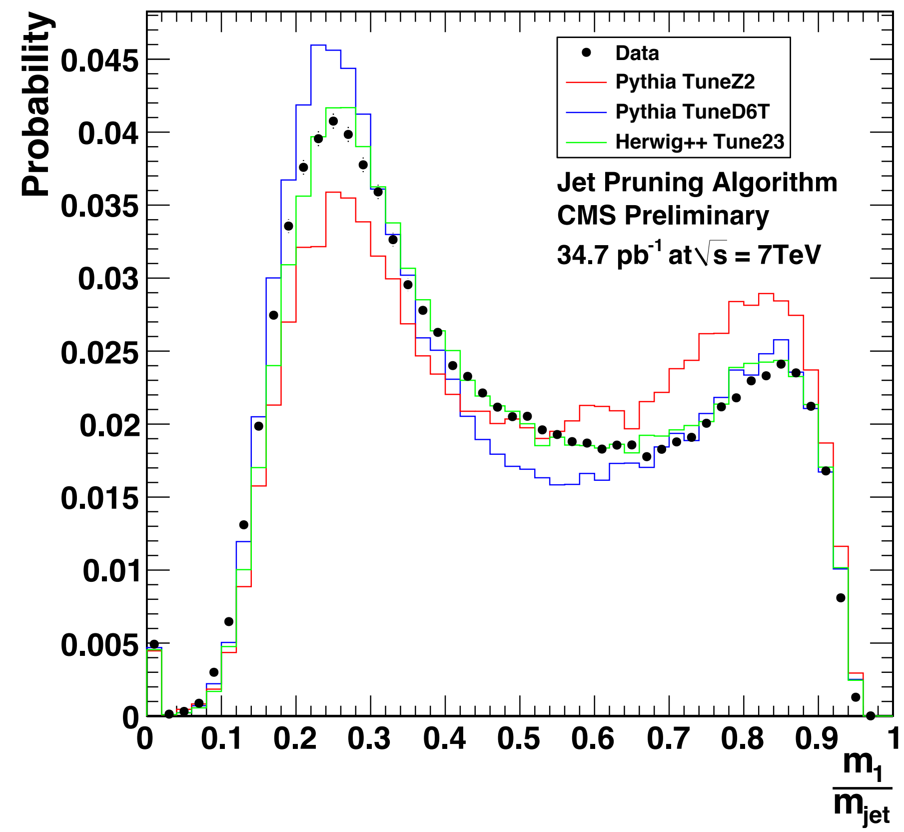


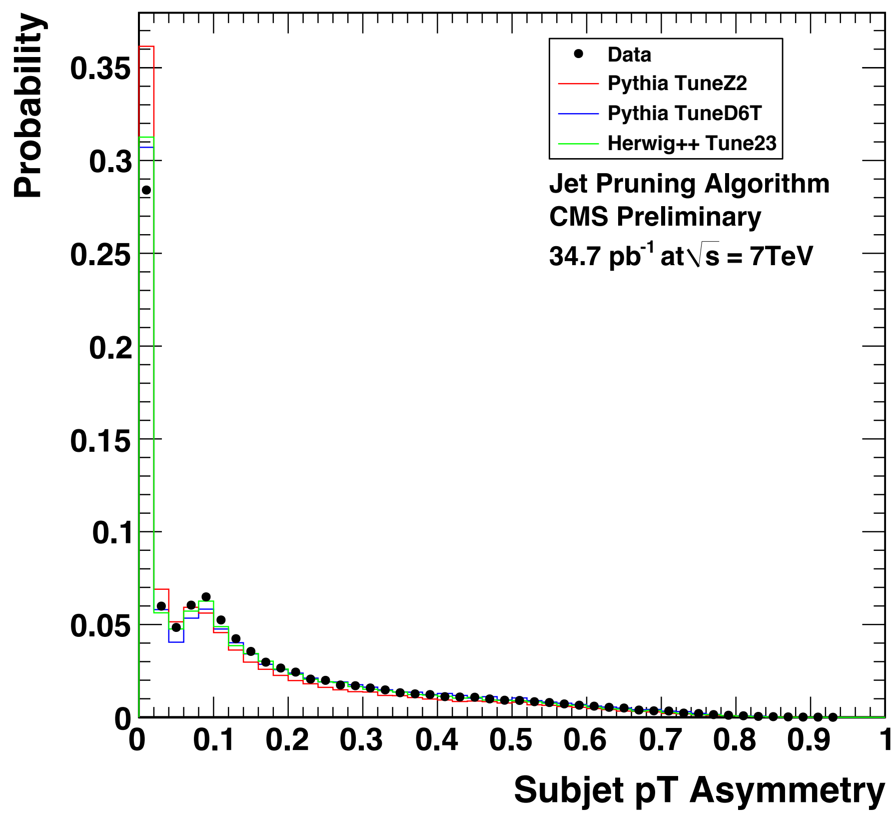


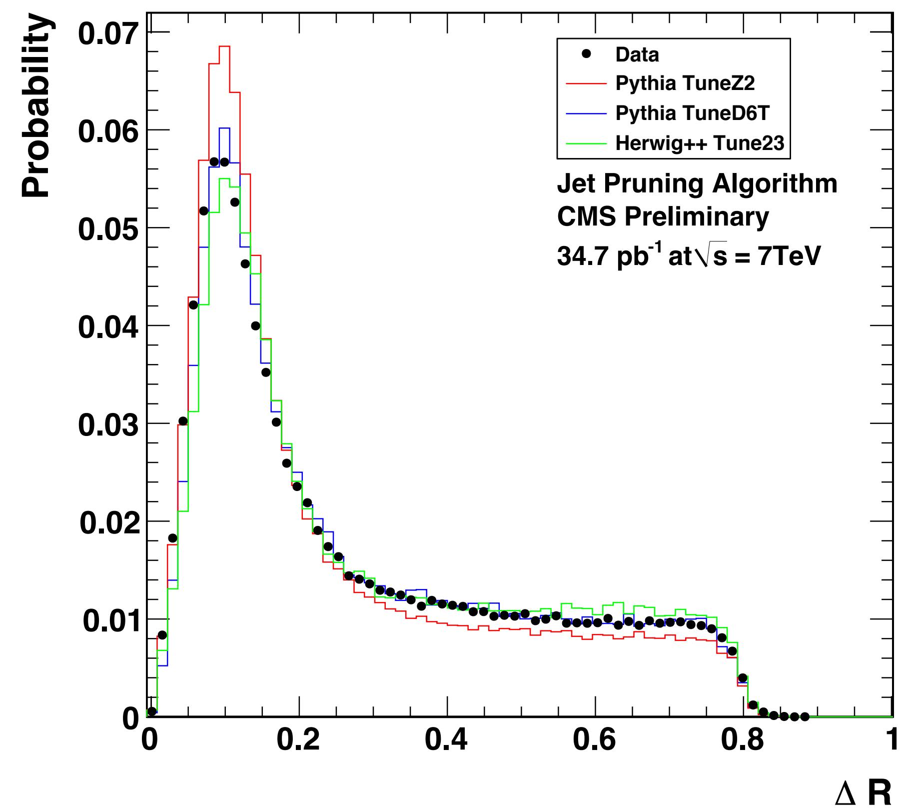


Jet Pruning Inclusive Comparison

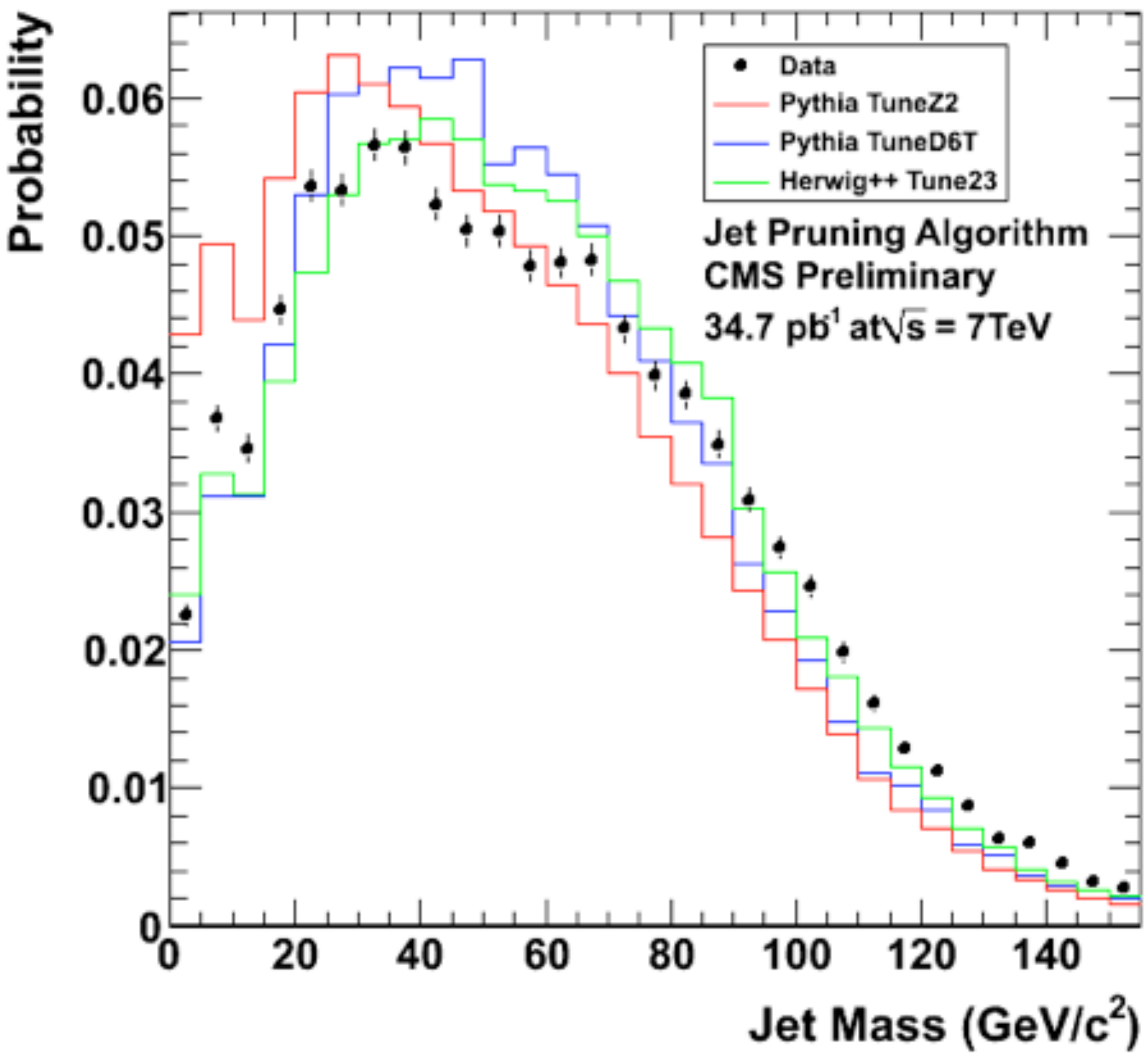


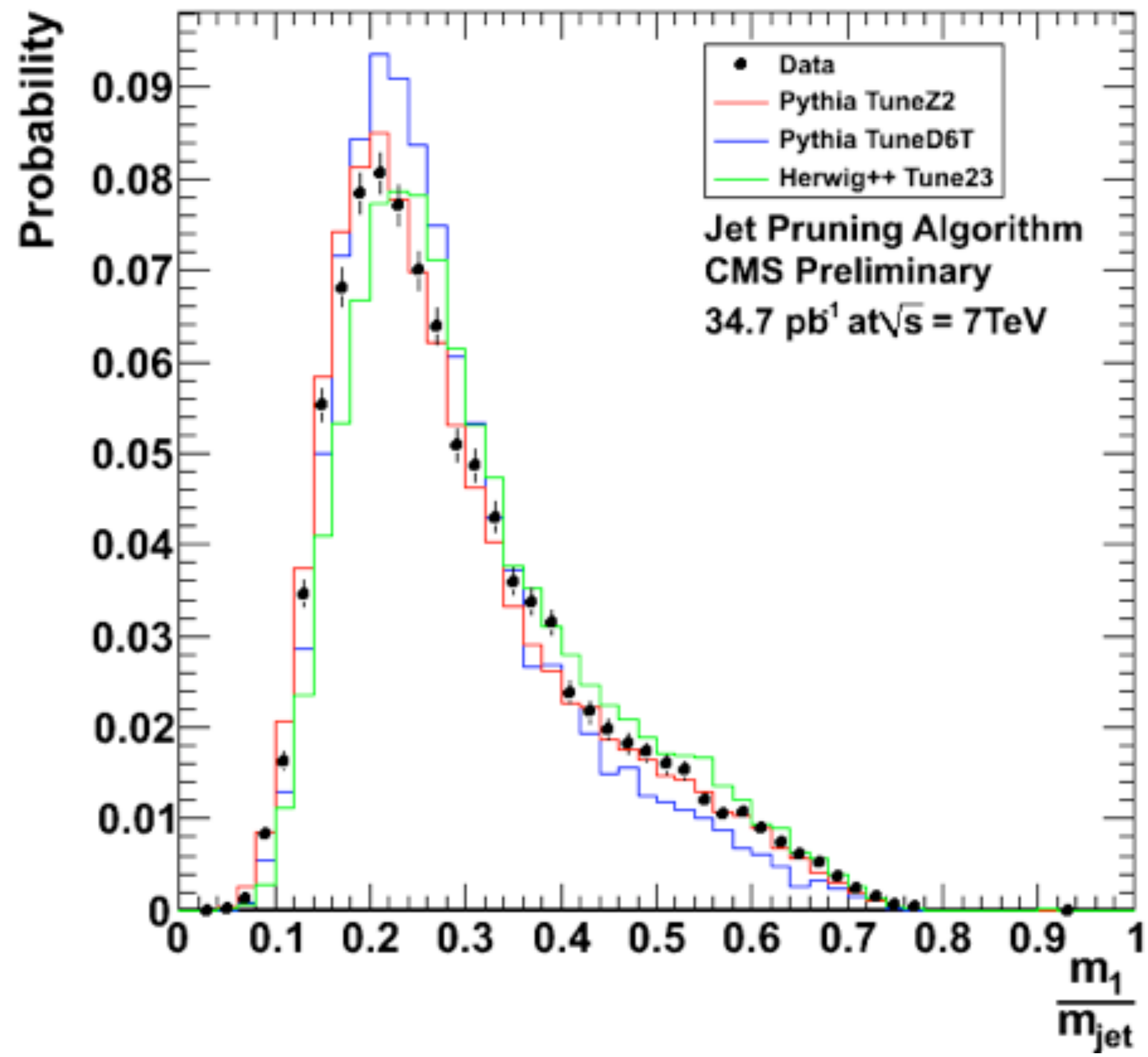




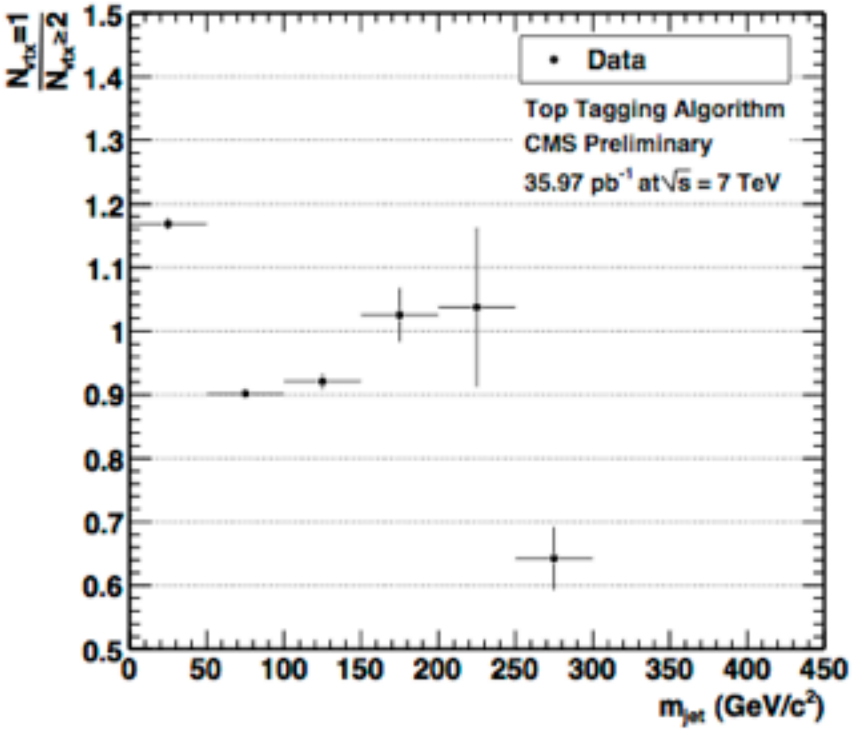


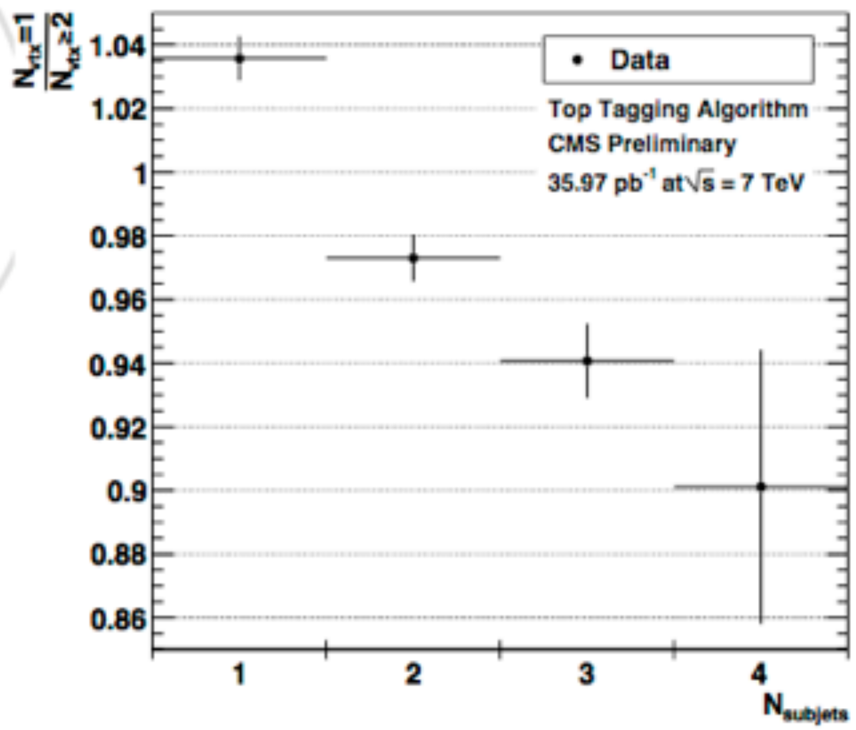
Jet Pruning "N-1" Comparison

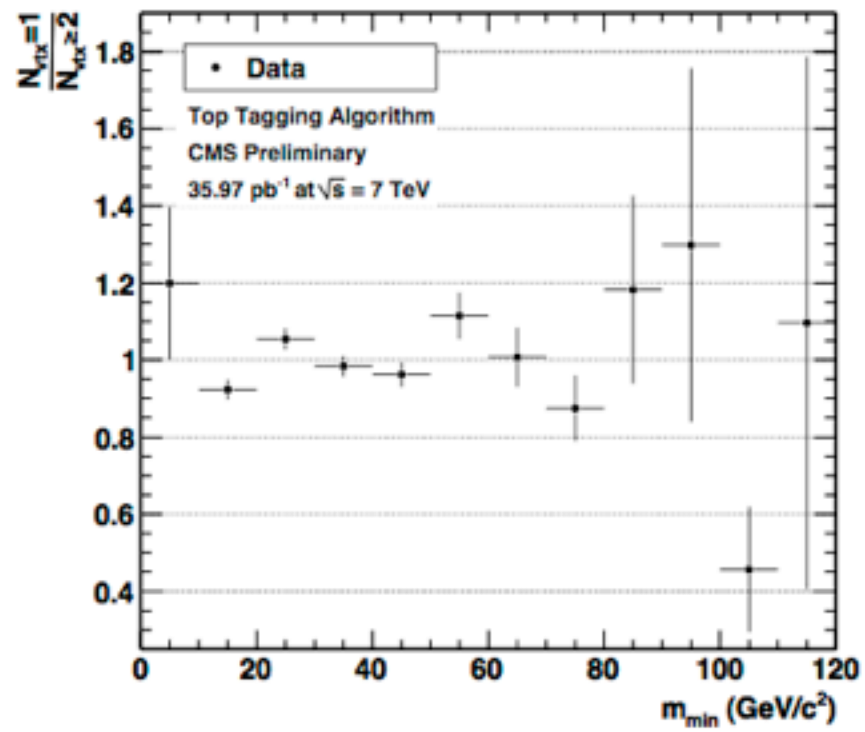


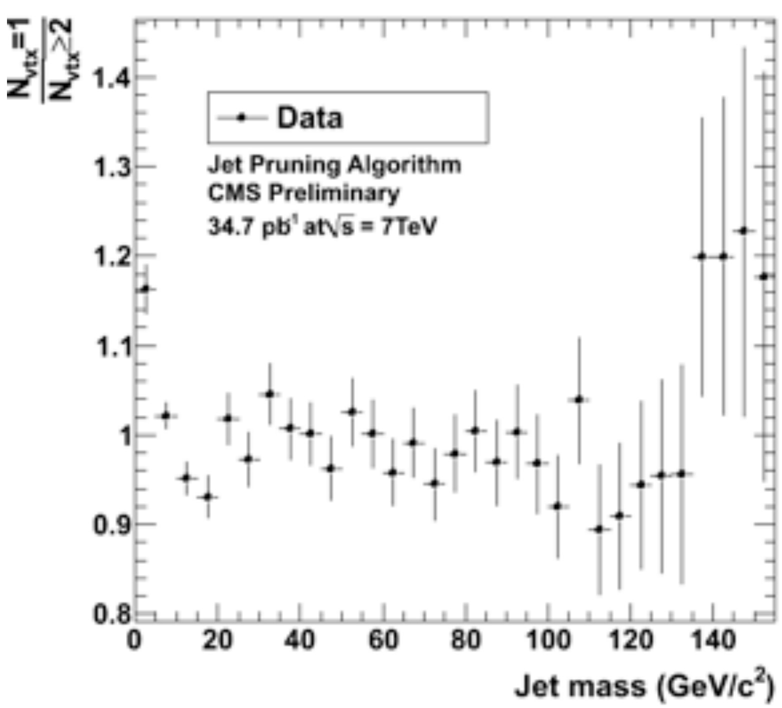


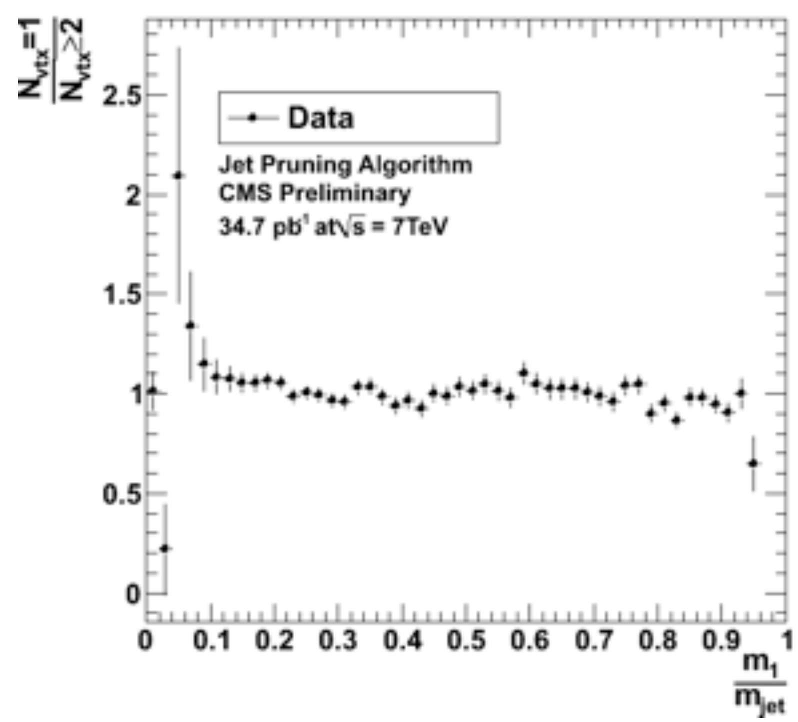
Pileup Dependence











- Some small pileup dependence seen

Conclusions

- Jet substructure tools being actively studied at CMS
- Ready for physics analysis
- Much data coming in from the LHC, these tools will be exercised in new physics searches