Experience with jet studies, and needs/prospects for using jets in new physics searches in CMS

Ricardo Vásquez Sierra, UC Davis

QCD and new physics searches, LPCC Workshop

February 4th, 2011

Outline

- CMS Jets
  - Algorithms, sizes and input types
- CMS searches/analyses with jets
- Towards a common jet in ATLAS and CMS
- Extreme cases: identifying boosted particles decaying hadronically using larger jets
  - Jet substructure
  - W tagging and top tagging
- Conclusions

CMS Jets

Algorithms officially supported

- antik_T — R=0.5, R=0.7
- k_T — R=0.4, R=0.6

Special boosted objects algorithm

- Cambridge-Aachen — R=0.8

Inputs

- MC particles
- Calorimeter towers
- Calorimeter towers plus tracks
- Particle flow candidates

- FastJet is used for all algorithms

CMS Searches with Jets

list not exhaustive

Group	**ANALYSIS**	**Algo**	**"Size"**
EXO	**Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.7
EXO	**Search for Stopped Gluinos in pp collisions at sqrt(s) = 7 TeV**	iterative cone	0.5
EXO	**Search for Pair Production of First-Generation Scalar Leptoquarks in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.5
EXO	**Search for Pair Production of Second-Generation Scalar Leptoquarks in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.5
EXO	**Search for Dijet Resonances in 7 TeV pp Collisions at CMS**	anti-k_T	0.7
EXO	**Search for Microscopic Black Hole Signatures at the LHC**	anti-k_T	0.5
SUS	**Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy**	anti-k_T	0.5

- Some dijet searches use larger cone size to reduce jet energy loss from final state radiation

CMS Analyses with Jets

list not exhaustive

Group	ANALYSIS	Algo	"Size"
QCD	**Dijet Azimuthal Decorrelations in pp Collisions at sqrt(s) = 7 TeV**	anti-k_T	0.5
QCD	**Measurement of Dijet Angular Distributions in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD	**Jet Transverse Structure and Momentum Distribution**	anti-k_T	0.5
QCD	**Hadronic Event Shapes in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD	**Measurement of the 3-jet to 2-jet Cross Section Ratio in pp Collisions at sqrt(s) = 7 TeV**	anti-k_T	0.5
QCD	**Measurement of the Inclusive Jet Cross Section in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD	**Measurement of the Underlying Event Activity at the LHC**	SISCone	0.5

- Other analyses/searches that are not public yet are mostly using anti-k_T 0.5

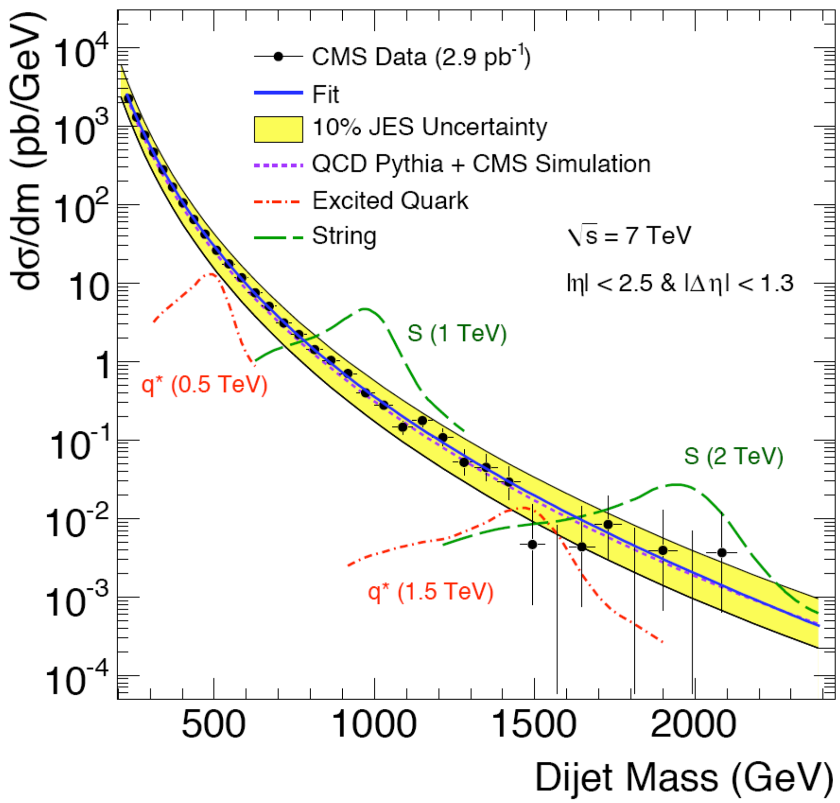
Some results

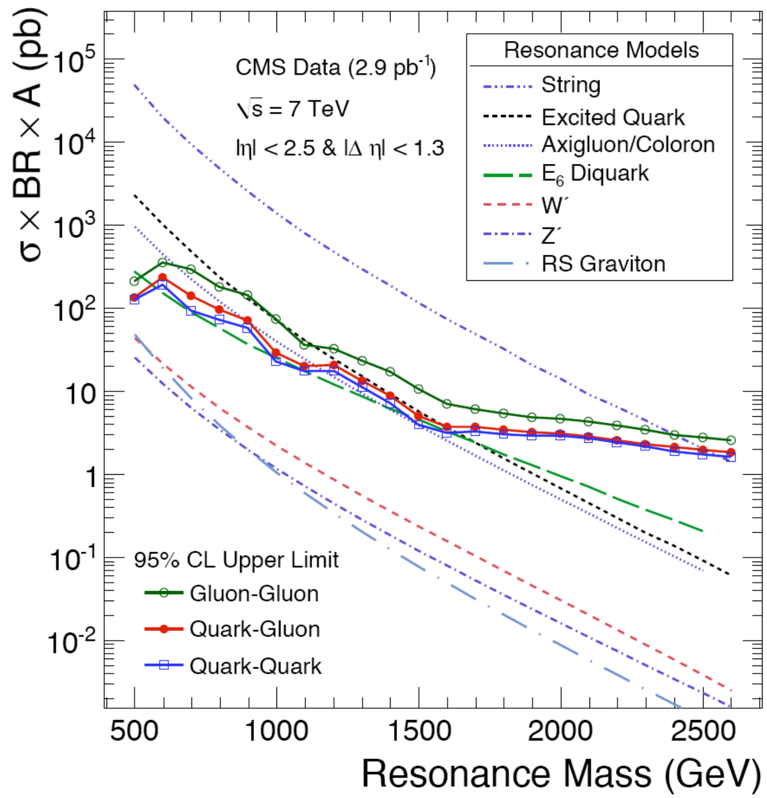
Dijet Bump Searches

Phys. Rev. Lett. **105**, 211801 (2010)

- Parametrize dijet mass spectrum with a smooth 4-parameter fit function:

$$\frac{d\sigma}{dm} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2 + P_3 ln(m/\sqrt{s})}}$$





anti-k_T 0.7

$$m_{q^*} < 1.58\,TeV$$

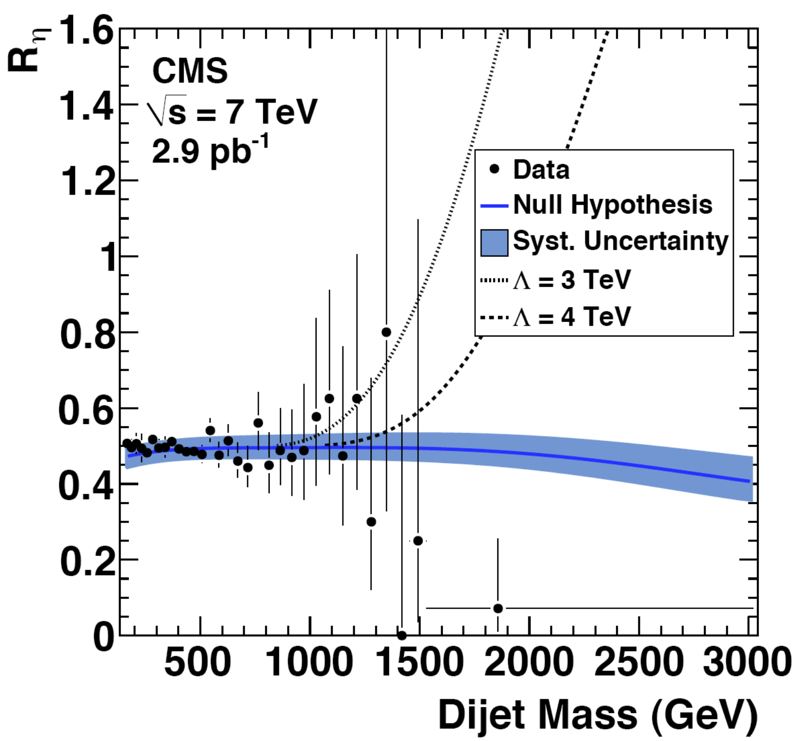
$$m_S < 2.5\,TeV$$ excluded at 95% CL

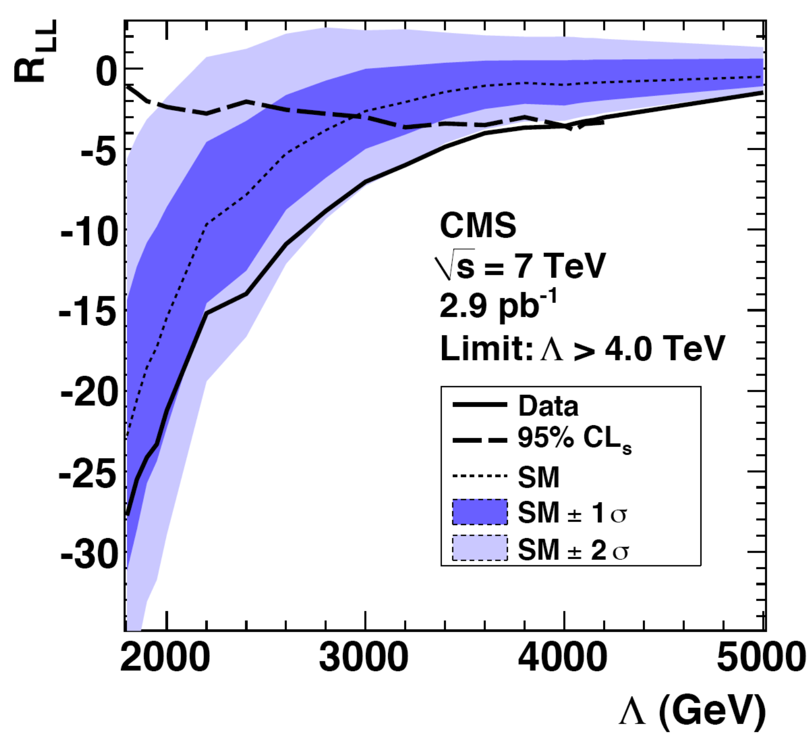
Dijet Centrality Ratio

Phys. Rev. Lett. **105**, 262001 (2010)

$$R_\eta = \frac{N_{events}(\text{leading jets with } |\eta| < 0.7)}{N_{events}(\text{leading jets with } 0.7 < |\eta| < 1.3)}$$

- Measurement is sensitive to quark compositeness since the distribution is expected to be very flat for SM

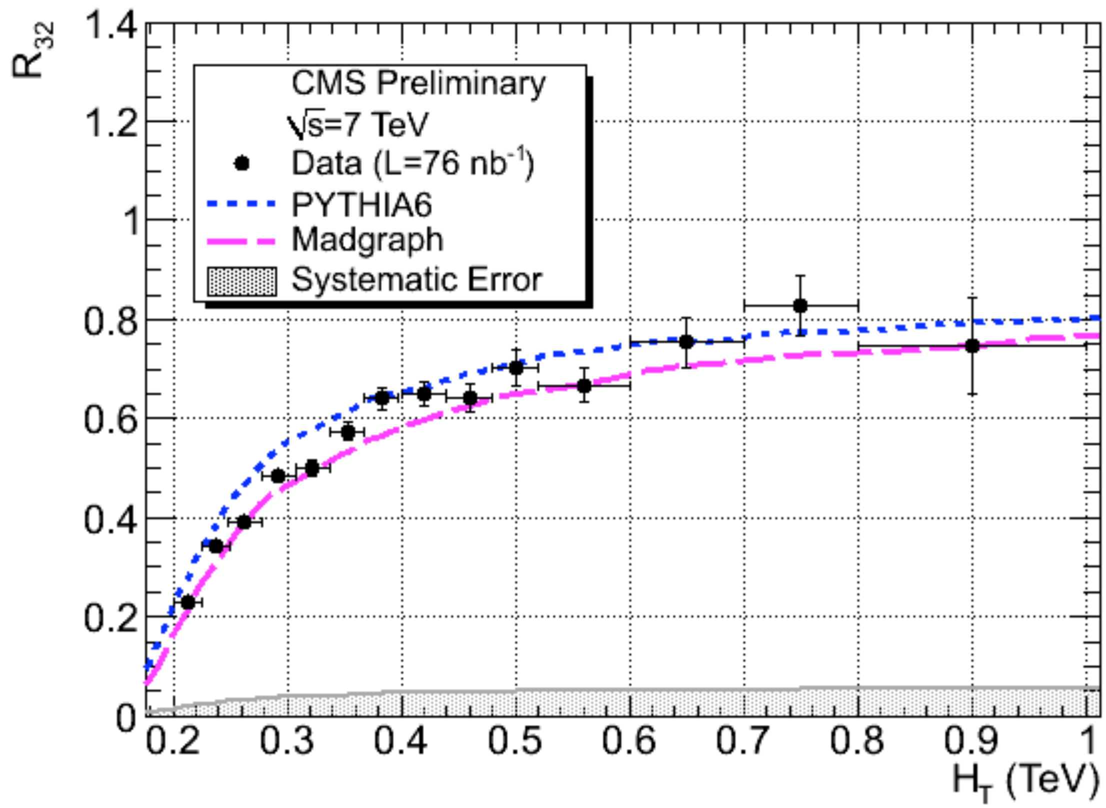




anti-k_T 0.7

3-jet to 2-jet XS ratio

Jet analysis with > 2 objects in the event



- H_T = scalar transverse momentum sum of all jets

$$R_{32} = \frac{d\sigma_3/dH_T}{d\sigma_2/dH_T}$$

anti-k_T 0.5

Towards a common jet in ATLAS and CMS

Towards a common definition

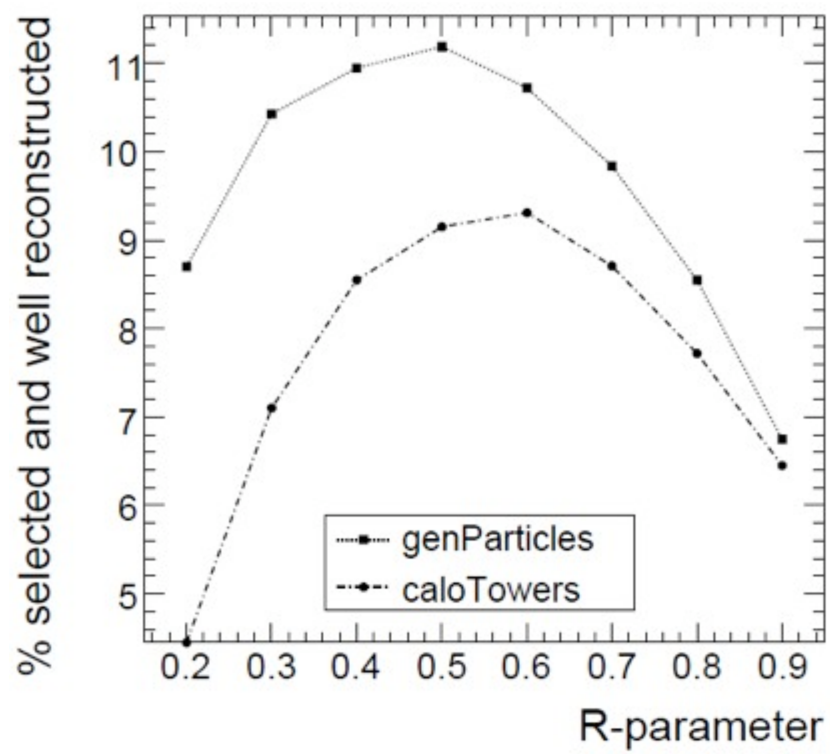
- Why do we want to discuss this issue?
  - Tevatron experiments did not have one common choice
  - Often regretted, especially by (QCD) theorists, but also experimentalist, since it did not allow any direct comparison.
  - The sooner we agree the better (if we ever do)
  - There will be analyses in ATLAS and CMS using jets that would greatly benefit from having one common point of comparison: QCD especially

CMS Issues and Proposal

- For CMS, switching from 0.5 to 0.4 would be a large load
  - Calculating from scratch calibrations and performance
- Mainly, physics needs one "small" and one "large" jet size
  - Small for multi-object physics and large for "di-object" physics (mostly heavy resonance searches).
- Changing 0.7 → 0.6 is not a good idea, leaves us with 0.5 and 0.6, too close to each other
- Changing 0.7 → 0.8 is possible, leaves us with 0.5 and 0.8
  - 0.7 or 0.8 would be good for dijets in both CMS & ATLAS
- A proposal could be that CMS and ATLAS use a common large cone size even if we keep our different small cone sizes.
- A change of jet size though, is only feasible if another major change happens, like a new version of GEANT or something of that magnitude

k_T Algorithm

- Scan the R-parameter for kT algorithm shows that it is hard to conclude an ideal cone size, no sharp performance peak is found
- Performance most likely be dependent on the set of inputs used for the jet and the analysis, making the R-parameter hard to set



Top Physics study, in which "well reconstructed" means well separated jets

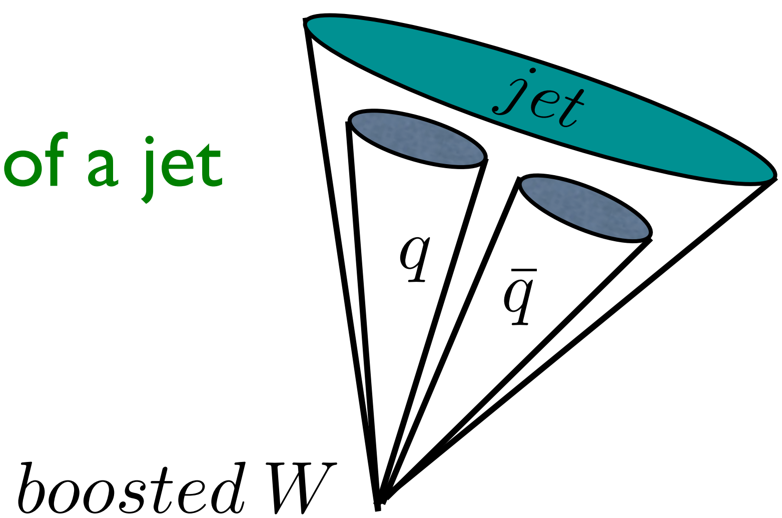
Physics concerns/needs

- Other concerns/needs that physics groups have with respect to jet algorithms:
  - Find the best algorithm that would allow efficient jet vetoes in particular in high pile-up conditions
  - Data driven background estimations (e.g. V+jets) might be naive in their assumptions and extrapolate the behavior from a control region.

Jet substructure

Jet Substructure

- Many physics scenarios often include heavy resonances, that can give rise to boosted hadronic objects
- The "classic" reconstruction techniques would fail identifying them due to merging of jets
- How can we look for them?
    - Looking at the substructure of a jet



Jet Pruning

- Jet Pruning from Ellis et al. (arXiv:0903.5081) on Cambridge-Aachen jets with R=0.8 is used for W-tagging
  - Removes soft, large angle particles
- Basic description:
  - It first clusters all inputs and obtains: m_{jet} and $p_{T_{jet}}$
  - Each jet is reclustered requiring each recombination (1+2→p) meet :

$$\frac{\min(p_{T1}, p_{T2})}{p_{Tp}} > 0.1 \qquad \Delta R_{12} < 0.5 \times \frac{m_{jet}}{p_{T_{jet}}}$$

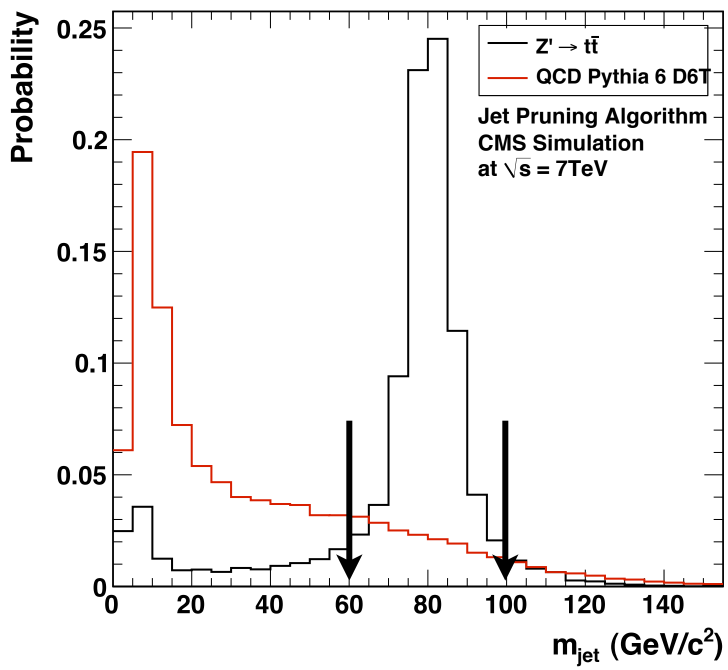
  - If the requirements are not met, the softest of the two, 1 or 2, is removed from the clustering sequence

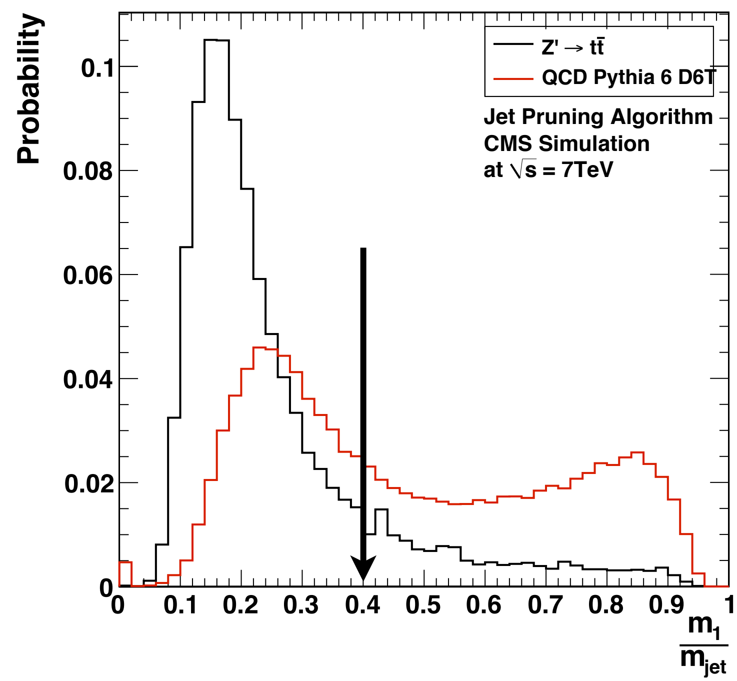
W tagging requirements

- For W tagging, we require the pruned jet mass to be centered in the W mass value and the hardest (m_1) of the two subclusters, from the last clustering iteration, to give a mass drop of 0.4:

$$60.0 < m_{jet} < 100.0\, GeV/c^2$$

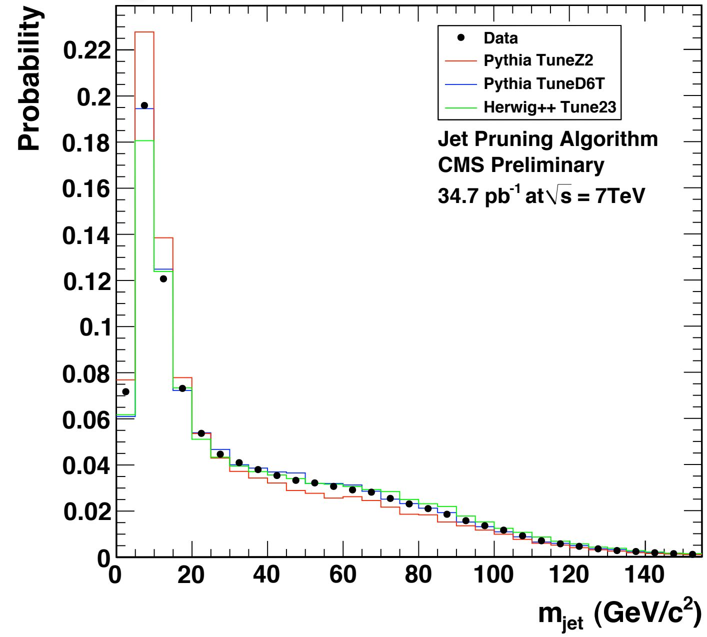
$$\frac{m_1}{m_{jet}} < 0.4$$

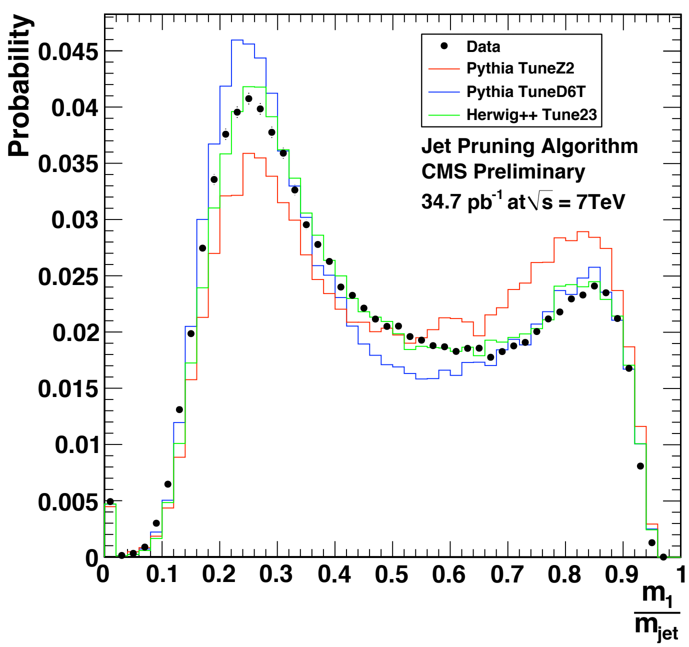


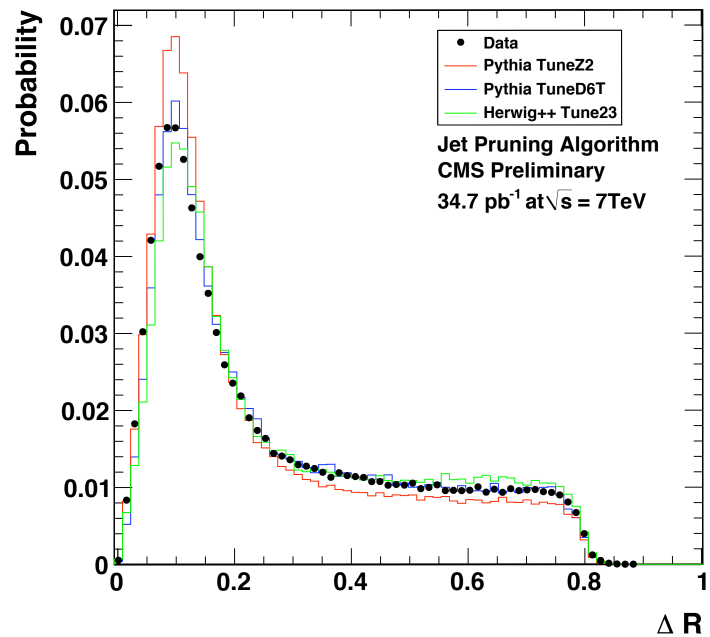


Pruned jets

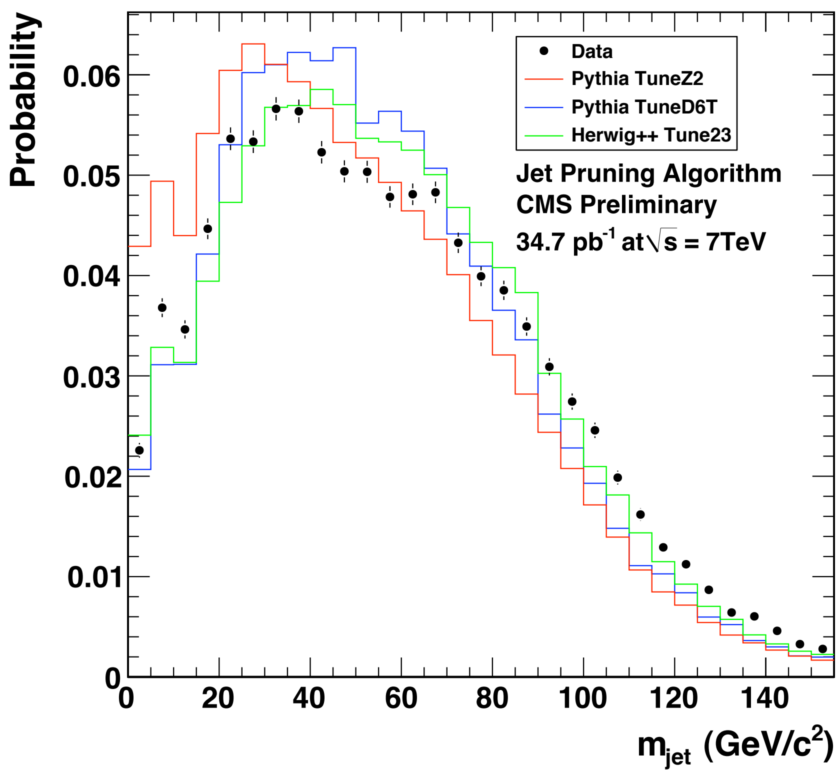
- Here we can see how several of the available Monte Carlos were compared to data

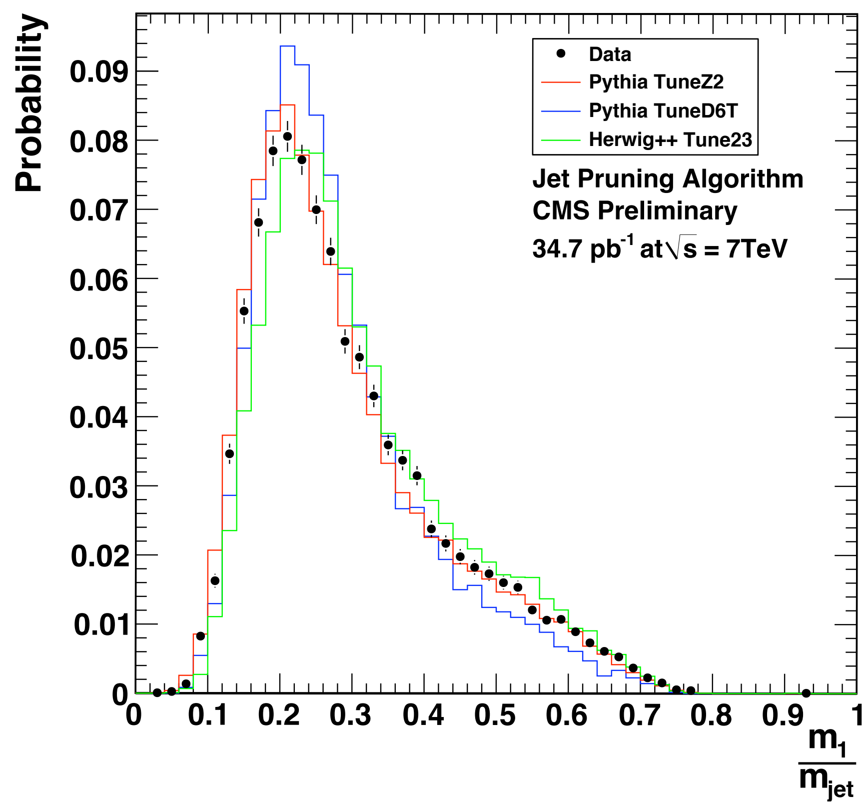






W tagging N-1





- Comparing data to different MC and MC tunes for all criteria applied except the one displayed, show some very promising results

Top Tagging Algorithm

- Based on Kaplan et al. (arXiv:0806.0848)
  - Cluster particle flow candidates using Cambridge-Aachen with R=0.8
- Reverse clustering sequence twice to find substructure
  - During the reversal of the sequence, subclusters found must satisfy two requirements:

$$p_T^{\text{subcluster}} > 0.05 p_T^{\text{cluster}}$$

$$\Delta R_{1,2} > 0.4 - 0.0004 p_T^{\text{cluster}}$$

- If one of the subclusters is too soft, it is removed, and the procedure is repeated on the remaining cluster
  - The grooming of the jet is done for last iterations of the clustering sequence only
- If the subclusters are too close to each other they don't get split

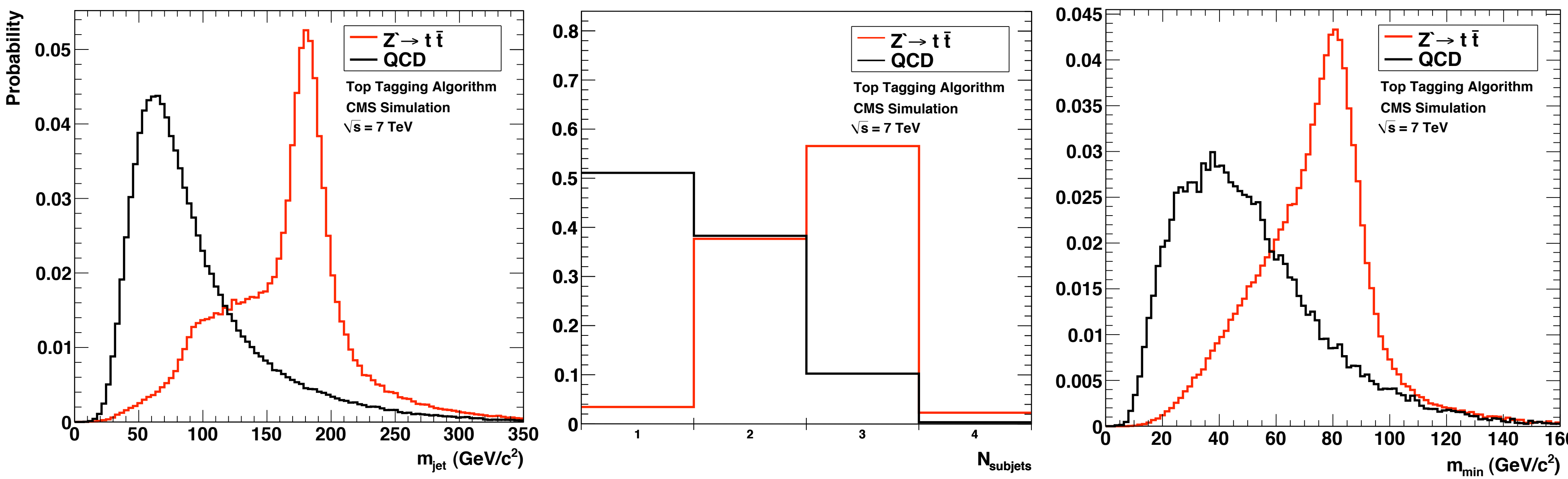
Top tagging requirements

$$N_{subjets} > 2$$

$$140.0 < m_{jet} < 250.0\, GeV/c^2$$

$$m_{\min} > 50.0\, GeV/c^2$$

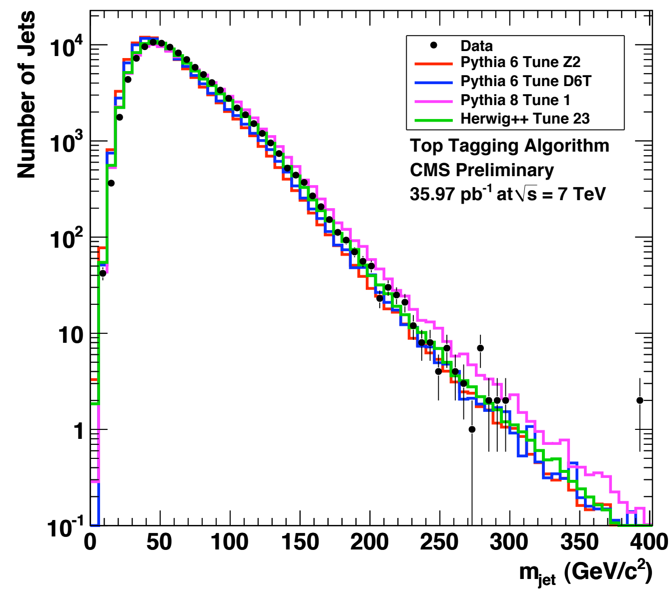
Simulation

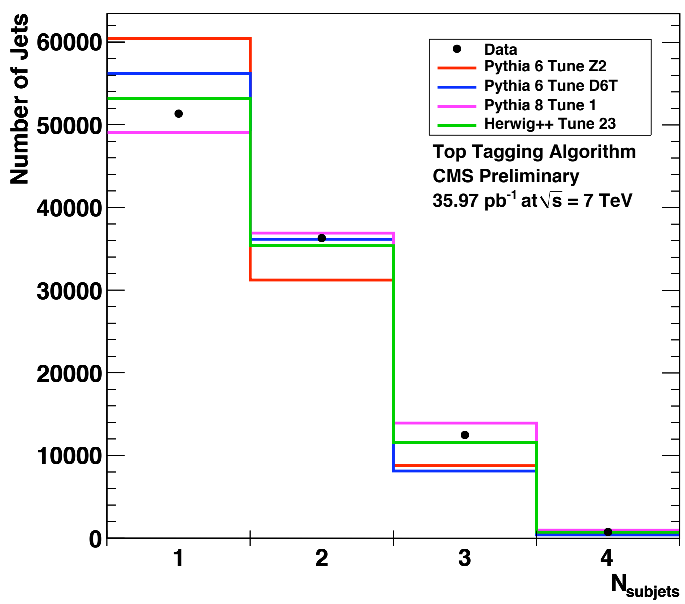


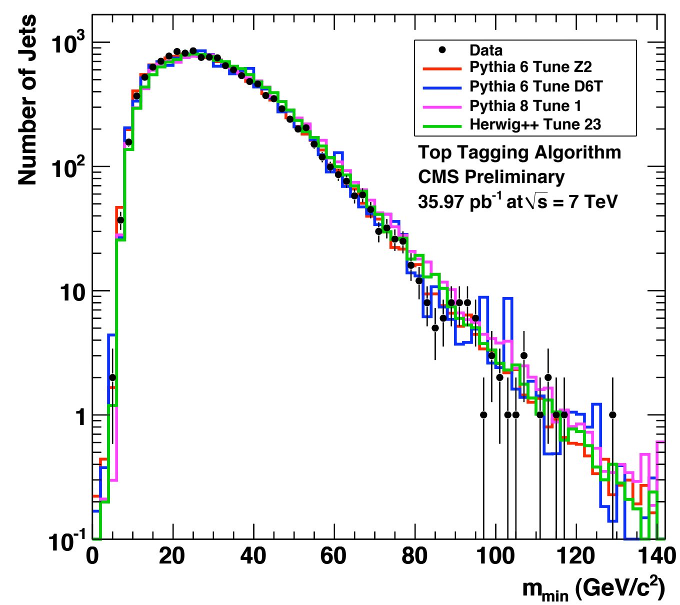
where: $m_{\min} = \min[m_{12}, m_{13}, m_{23}]$

Top tagging

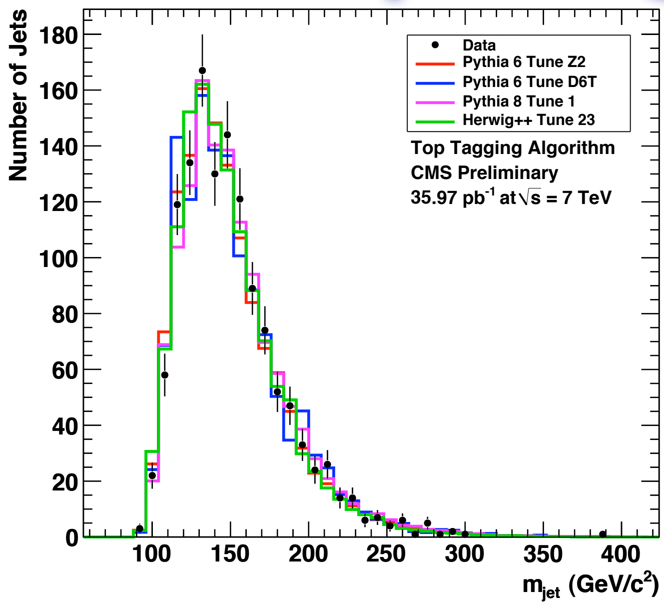
Comparison of top tagging variables with different Monte Carlos and MC tunes

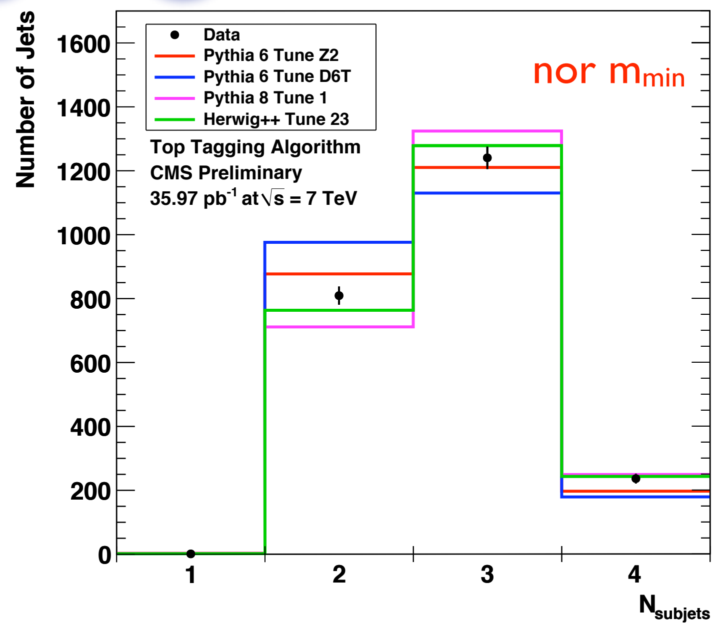


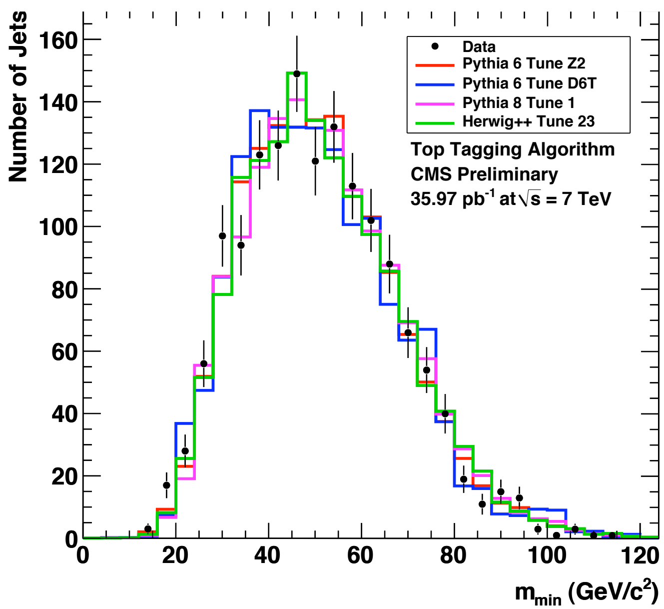




Top tagging N-1

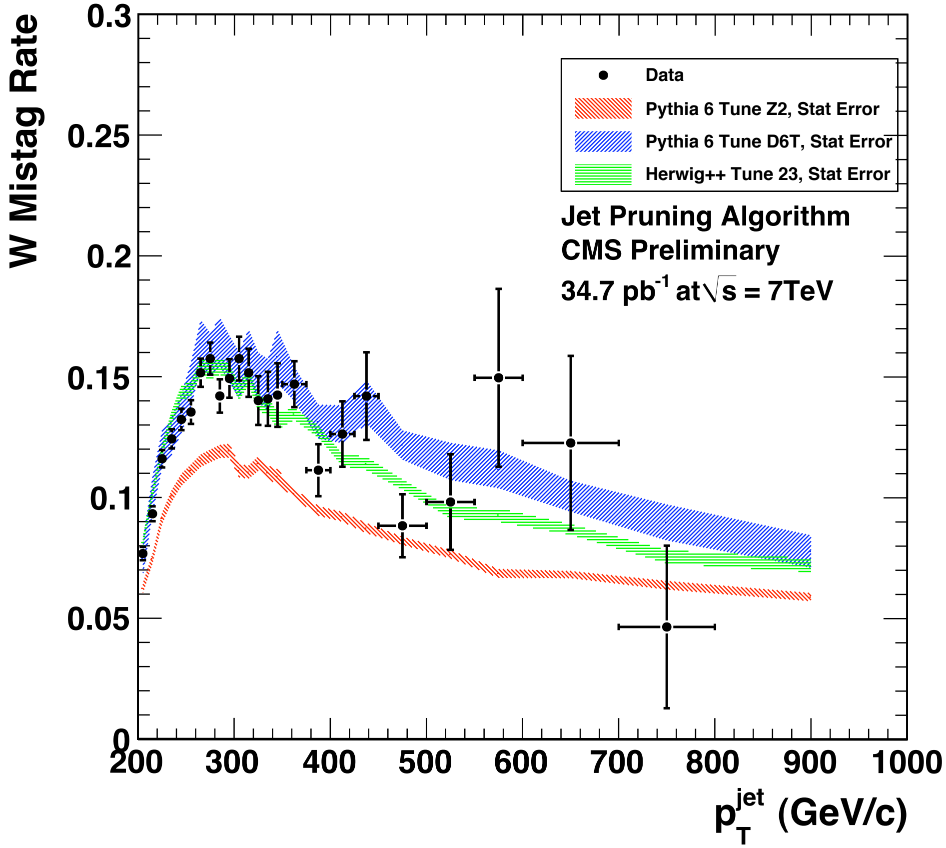


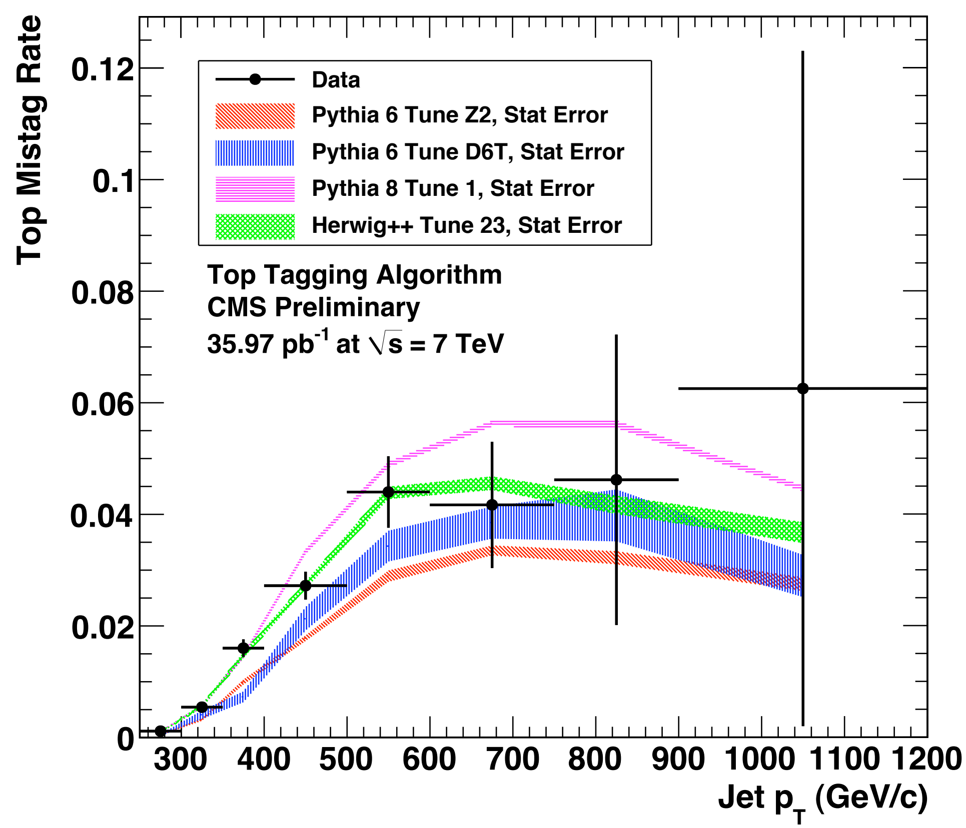




Here one can clearly see that the number of subjets variable has a very clear dependence on Monte Carlo and MC Tune used

Mistag measurements





- Dijet data were used to measure the mistag rates of both taggers and they were also compared to MC

Conclusions

- CMS predominantly uses anti-kt with cone size 0.5
  - changing this value right now is not possible because of the many implications/changes to the whole reconstruction/calibration/analysis chain
- We could imagine to discuss about a larger common cone size
- In any case, a change of algorithm parameters is only feasible in parallel with other major changes, e.g. some major new GEANT release in the future
- CMS has explored jet substructure techniques in data. They seem promising for the search of new physics!

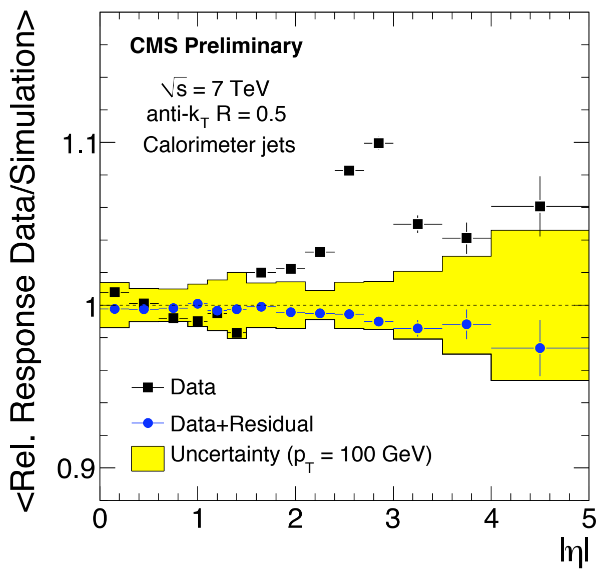
Steve Ellis' quote

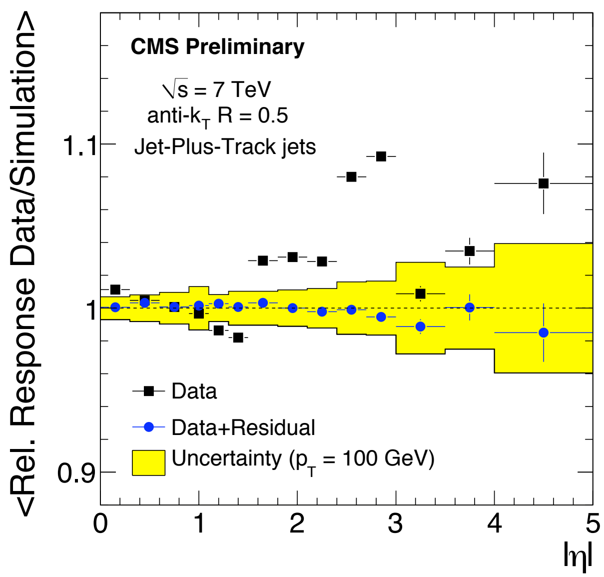
- "Since the underlying details of the jet analysis at CMS and ATLAS differ (e.g., particle-flow vs topo-clusters), it would be informative to have at least one jet distribution defined with exactly the same jet algorithm, such that, if both collaborations were correcting appropriately for the systematics, the distributions would agree within the residual uncertainties. N'est pas?"

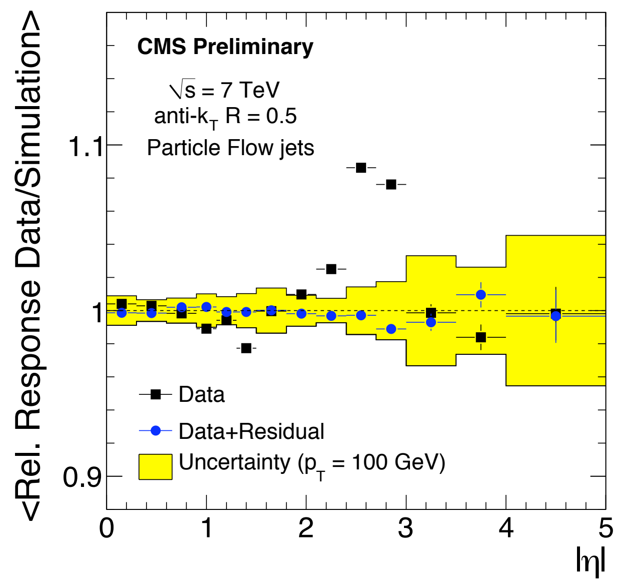
The end

Backup Slides

Jet Energy Corrections







- Three types of jets have been commissioned:
  - CaloJets, Jets-Plus-Tracks and Particle Flow Jets
- Good description of basic jet properties with respect to MC

CMS Searches with Jets

list not exhaustive

	ANALYSIS	Algo	"Size"
EXO-10-002	**Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.7
EXO-10-003	**Search for Stopped Gluinos in pp collisions at sqrt(s) = 7 TeV**	iterative cone	0.5
EXO-10-005	**Search for Pair Production of First-Generation Scalar Leptoquarks in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.5
EXO-10-007	**Search for Pair Production of Second-Generation Scalar Leptoquarks in pp Collisions at sqrt(*s*) = 7 TeV**	anti-k_T	0.5
EXO-10-010	**Search for Dijet Resonances in 7 TeV pp Collisions at CMS**	anti-k_T	0.7
EXO-10-017	**Search for Microscopic Black Hole Signatures at the LHC**	anti-k_T	0.5

- Some dijet searches use larger cone size to reduce jet energy loss from final state radiation

CMS Analyses with Jets

list not exhaustive

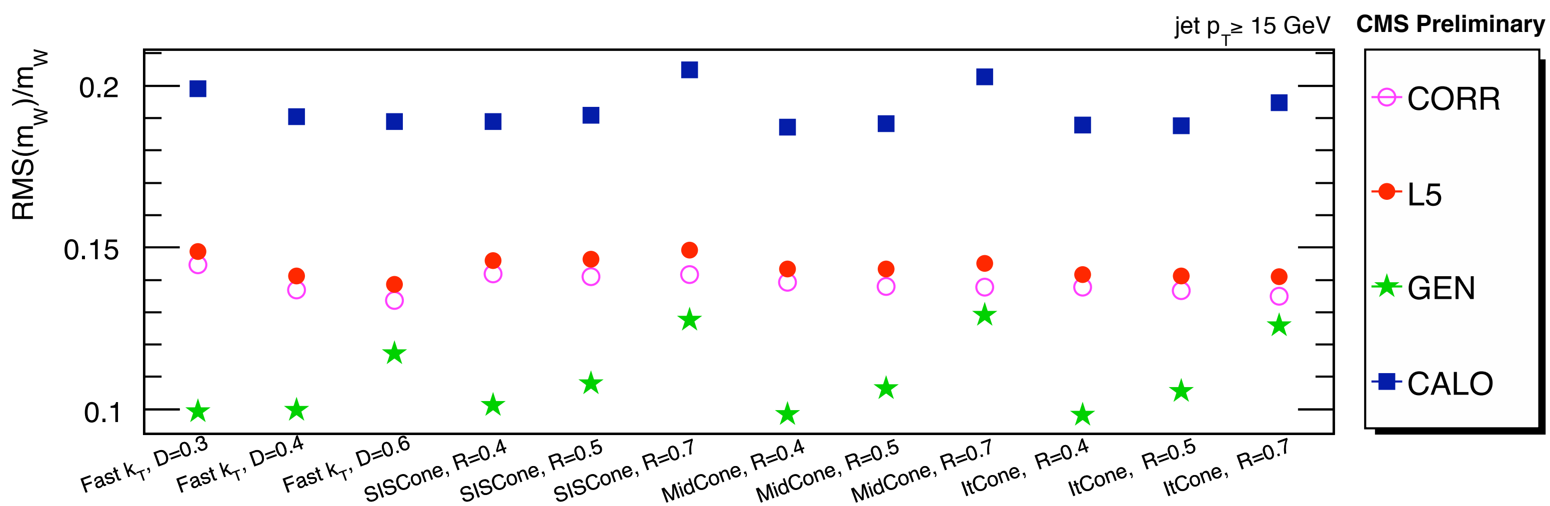
	ANALYSIS	Algo	"Size"
SUS-10-003	**Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy**	anti-k_T	0.5
QCD-10-026	**Dijet Azimuthal Decorrelations in pp Collisions at sqrt(s) = 7 TeV**	anti-k_T	0.5
QCD-10-016	**Measurement of Dijet Angular Distributions in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD-10-014	**Jet Transverse Structure and Momentum Distribution**	anti-k_T	0.5
QCD-10-013	**Hadronic Event Shapes in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD-10-012	**Measurement of the 3-jet to 2-jet Cross Section Ratio in pp Collisions at sqrt(s) = 7 TeV**	anti-k_T	0.5
QCD-10-011	**Measurement of the Inclusive Jet Cross Section in pp Collisions at 7 TeV**	anti-k_T	0.5
QCD-10-010	**Measurement of the Underlying Event Activity at the LHC**	SISCone	0.5

- Other analyses/searches that are not public yet are mostly using anti-k_T 0.5

Jet size studies

- In CMS, there have been some jet studies related to jet sizes
- Unfortunately, none of them include anti-kT

Relative width of m_W



W tagging details

In detail, the algorithm is as follows. The clustering sequence is reran for each jet and at every step of the clustering, when clusters i and j are being merged to a single cluster p, two additional requirements are made:

$$z_{\mathrm{ij}} \equiv min(p_{\mathrm{T},i}, p_{\mathrm{T},j})/p_{\mathrm{T},p} \quad < \quad z_{\mathrm{cut}} \tag{4}$$

$$\Delta R_{ij} \quad > \quad D_{\mathrm{cut}} = \alpha \times \frac{m_{\mathrm{J}}}{p_{\mathrm{T}}^{\mathrm{J}}} \tag{5}$$

If $(i + j \rightarrow p)$ step satisfies these two requirements, i and j are not merged and instead the softer of the two clusters is removed. The z_{ij} requirement ensures that soft particles are discarded, and the ΔR requirement ensures that wide-angle particles are discarded.

Top tagging variables

- **Jet Mass m_{jet}** - The mass of the four-vector sum of the constituents of the hard jet.
- **Minimum Pairwise Mass m_{min}** - The three highest p_{T} subjets are taken pairwise, and each pair's invariant mass is calculated via $m_{ij} = \sqrt{(E_i + E_j)^2 - (\vec{p}_i + \vec{p}_j)^2}$. m_{min} is the mass of the pair with the lowest invariant mass ($m_{\mathrm{min}} = \min[m_{12}, m_{13}, m_{23}]$). This variable is not defined for jets with less than three subjets.
- **Number of Subjets N_{subjets}** - The number of subjets found by the algorithm.

Top tagging details

1. The pairwise clustering sequence which was used to form the jet is examined in reverse order in order to find two subclusters.

2. If the two subclusters satisfy $\sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} > 0.4 - 0.0004 \times p_{\mathrm{T}}^{C}$, then continue to the next step. If not, the sub-clusters are too close and the decomposition fails. p_{T}^{C} is the transverse momentum of the initial object fed to the decomposition stage. The primary decomposition uses the jet p_{T} ($p_{\mathrm{T}}^{C} = p_{\mathrm{T}}^{\mathrm{jet}}$), while the secondary decomposition uses the p_{T} of the cluster found by the primary decomposition ($p_{\mathrm{T}}^{C} = p_{\mathrm{T}}^{A}$ or $p_{\mathrm{T}}^{C} = p_{\mathrm{T}}^{B}$).

3. If the two subclusters satisfy the momentum fraction criterion $p_{\mathrm{T}}^{\mathrm{cluster}} > \delta_{\mathrm{p}} \times p_{\mathrm{T}}^{\mathrm{hardjet}}$, then decomposition succeeds. The p_{T} requirement on the cluster serves to remove low-p_{T} clusters from consideration. The default value of $\delta_{\mathrm{p}} = 0.05$ was found to be optimal based on simulation studies.

4. If only one of the subclusters satisfies the criterion $p_{\mathrm{T}}^{\mathrm{cluster}} > \delta_{\mathrm{p}} \times p_{\mathrm{T}}^{\mathrm{hardjet}}$, then the decomposition process is repeated on the passed cluster, ignoring the constituents from the failed cluster. This decomposition is repeated until both clusters pass, both clusters fail, or the cluster consists of a single constituent.

5. If, after this iterative process, there is no cluster with $p_{\mathrm{T}}^{\mathrm{cluster}} > \delta_{\mathrm{p}} \times p_{\mathrm{T}}^{\mathrm{hardjet}}$, or the cluster is a single constituent, the decomposition fails.