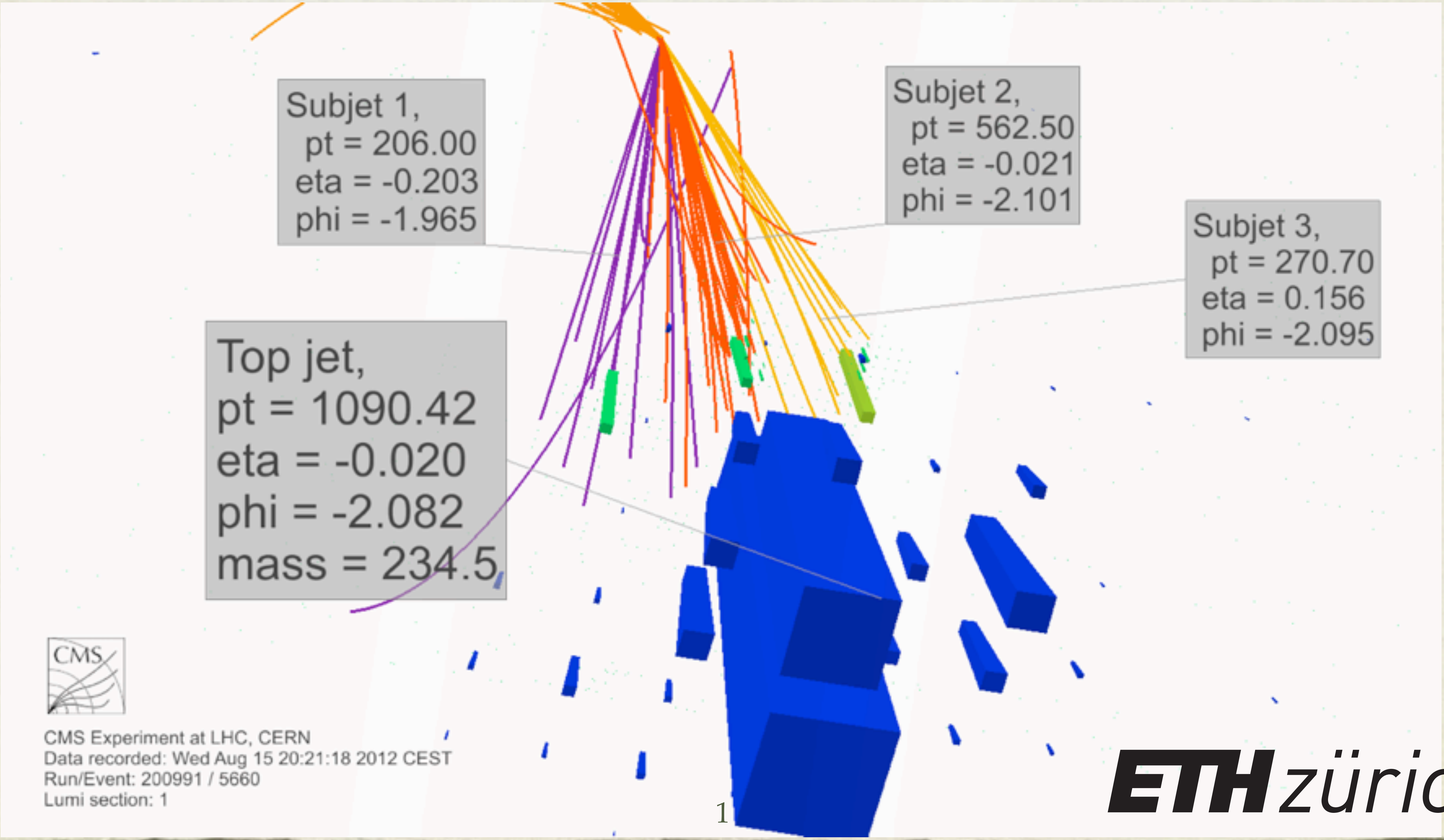
BOOSTED TOPS

Gregor Kasieczka
On behalf of the ATLAS and CMS Experiments

SM@LHC, Madrid - 10. April 2014



ETH zürich

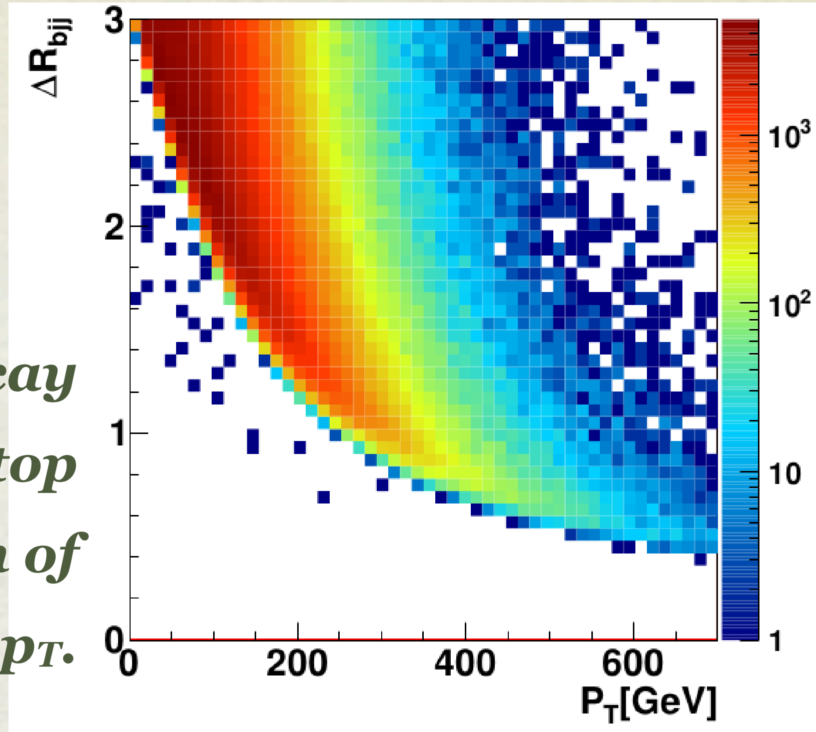
BOOSTED TOPOLOGIES

- Many models of new physics couple to the top quark
- At the LHC top-quarks are produced at high boost ($p_T/m > 2$)
- For these top quarks the decay products will become more collimated...

Separation for a two-body decay:

$$\frac{m_j^2}{p_{T,j}^2} \sim z(1-z)\, \Delta R_{j_1 j_2}^2 \qquad \text{with} \quad z = \frac{\min p_{T,j_i}}{p_{T,j}}$$

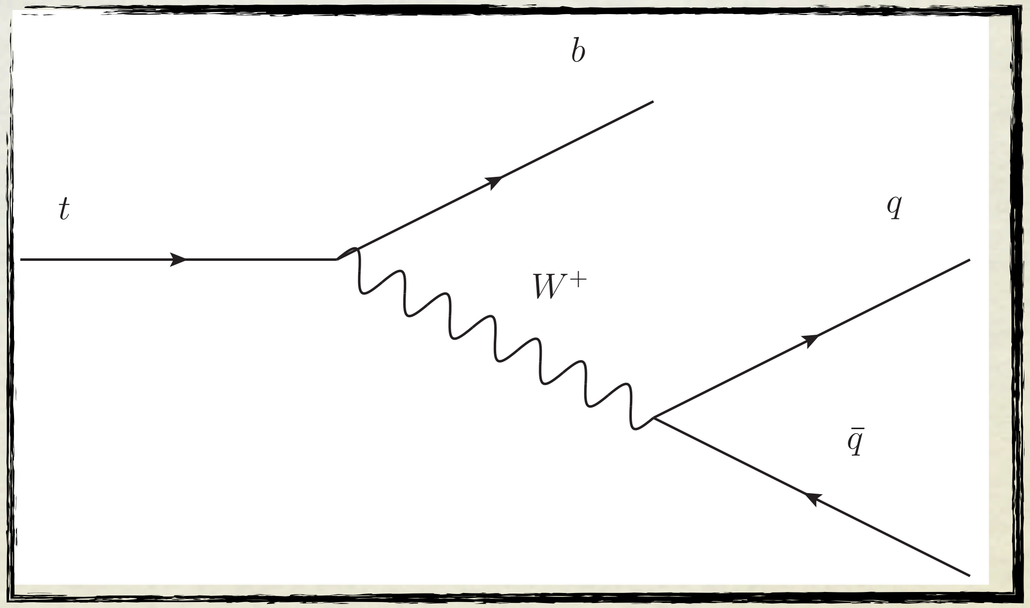
Distance of the decay products of a top quark as function of the p_T.



- ... and can/will merge into a single jet in the detector
- The fraction of such events will increase for the 13 TeV LHC run

ANALYSIS APPROACH

- Start with **large-R** (R=0.8 .. 1.8) and analyze the substructure
- General purpose tools (= *ingredients*):
  - **Find** the **hard** substructure: trimming, mass-drop, filtering, pruning, ...
  - **Discriminate** top vs non-top: masses, splitting-scales, n-subjettiness, ...
- **Taggers** are (usually) **combination** of these tools
  - example: mass-drop + filtering + masses (+ some combinatorics) = HEPTopTagger
- Goals:
  - Improve **mass resolution** by reducing pile-up, underlying event, soft radiation
  - **Identify** hadronic top candidate



PERFORMANCE STUDIES

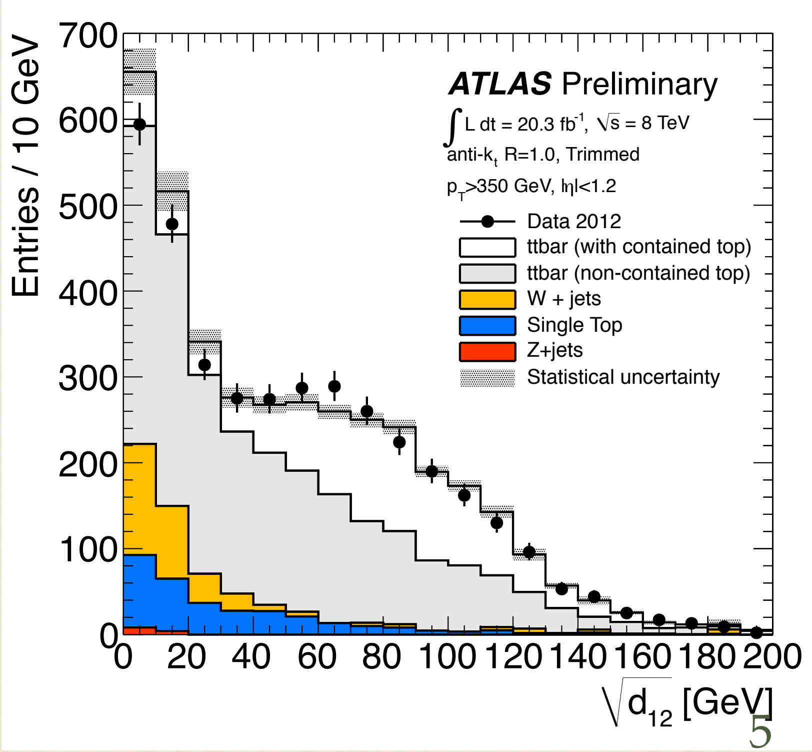
KT SPLITTING SCALES

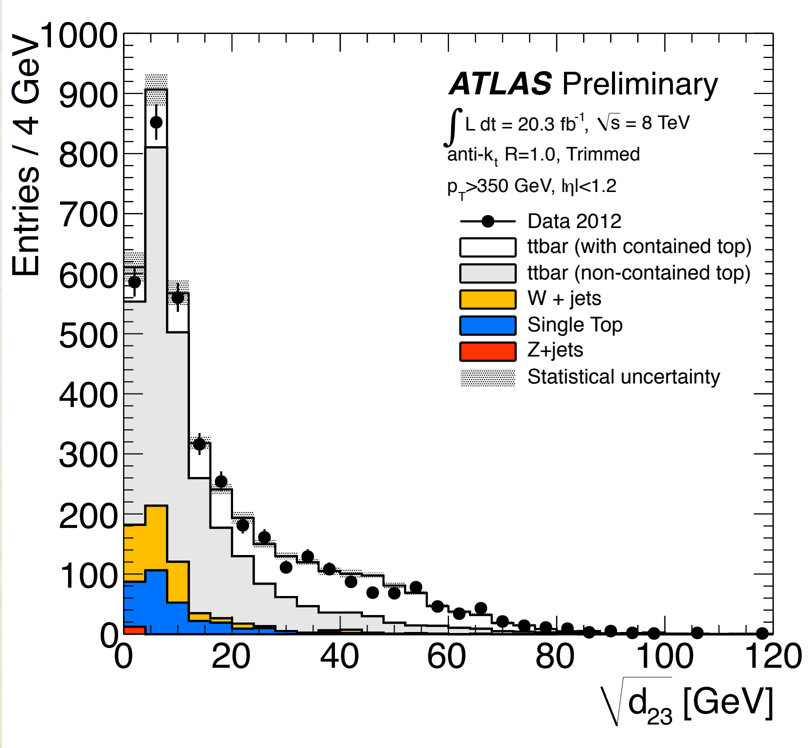
1306.4945

- **Re-cluster** the constituents of a jet using the **kT** algorithm
- Look at distance measure (*splitting scale)* of last or second-to-last step

$$\sqrt{d_{ij}} = \min(p_{Ti}, p_{Tj}) \times \Delta R_{ij},$$

- Distinguishes symmetric heavy particle decays from QCD radiation:
  - For a heavy-particle decay: $\sqrt{d_{12}} \approx m_{top/2}$ $\sqrt{d_{23}} \approx m_{W/2}$
  - QCD radiation: smaller



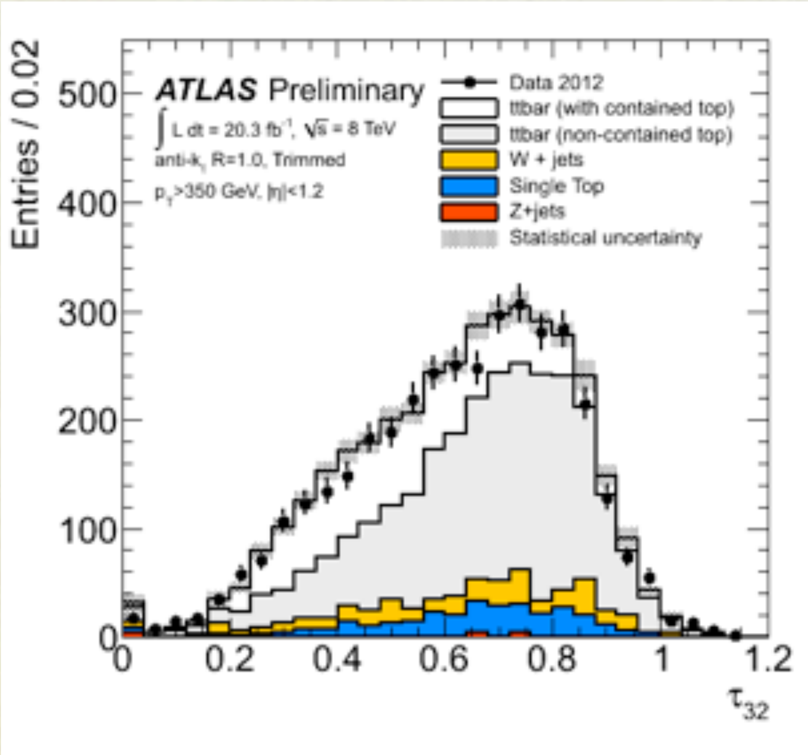


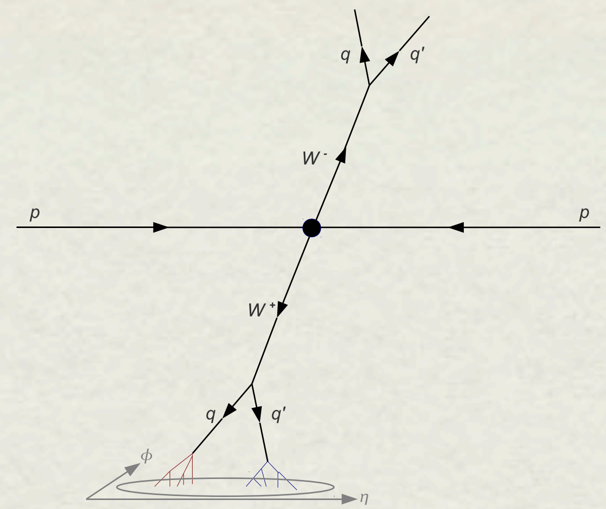
N-SUBJETTINESS

1011.2268 **CMS PAS JME-13-007** **ATLAS-CONF-2013-084**

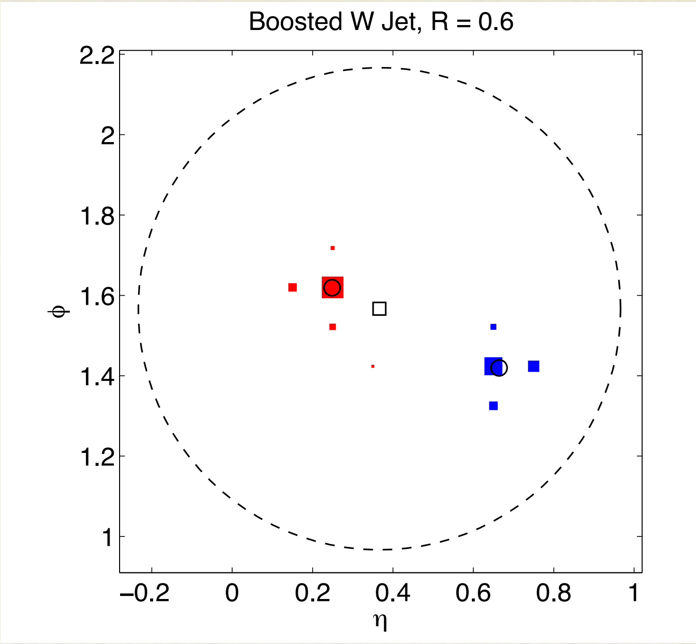
$$\tau_N = \frac{1}{d_0}\sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\} \qquad d_0 = \sum_k p_{T,k} R_0,$$

- Idea: Quantify compatibility with 1/2/3-subjet hypothesis
- Lower τ_n = better agreement
- Ratio τ_3/τ_2 for top tagging

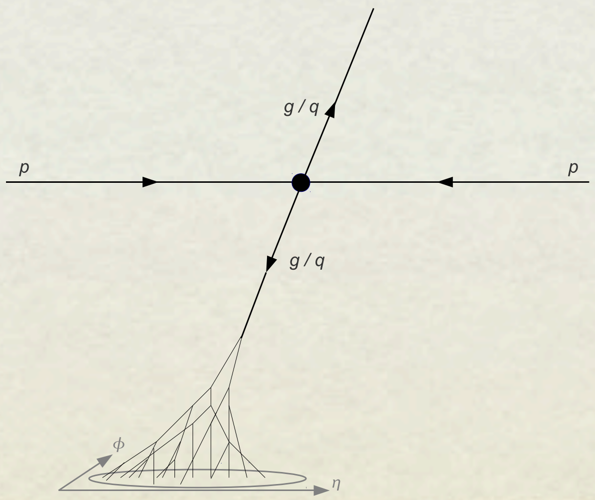


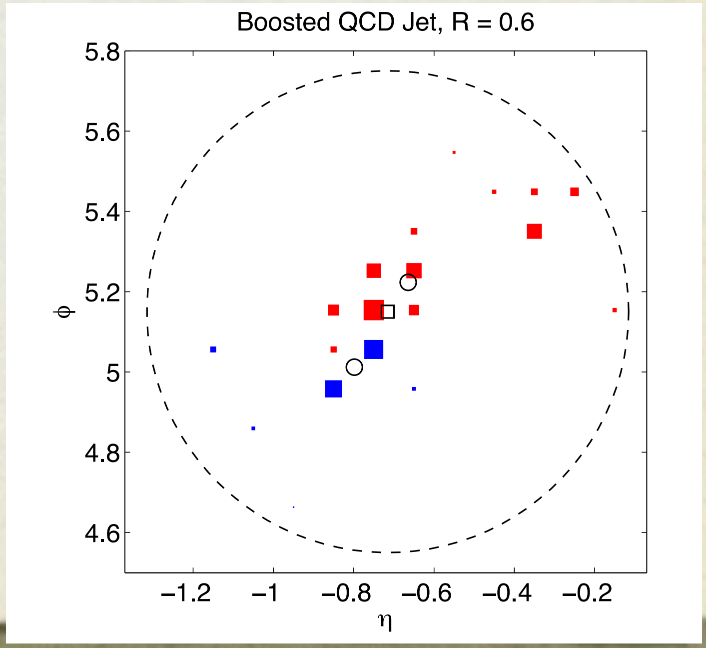


(a)



(b)





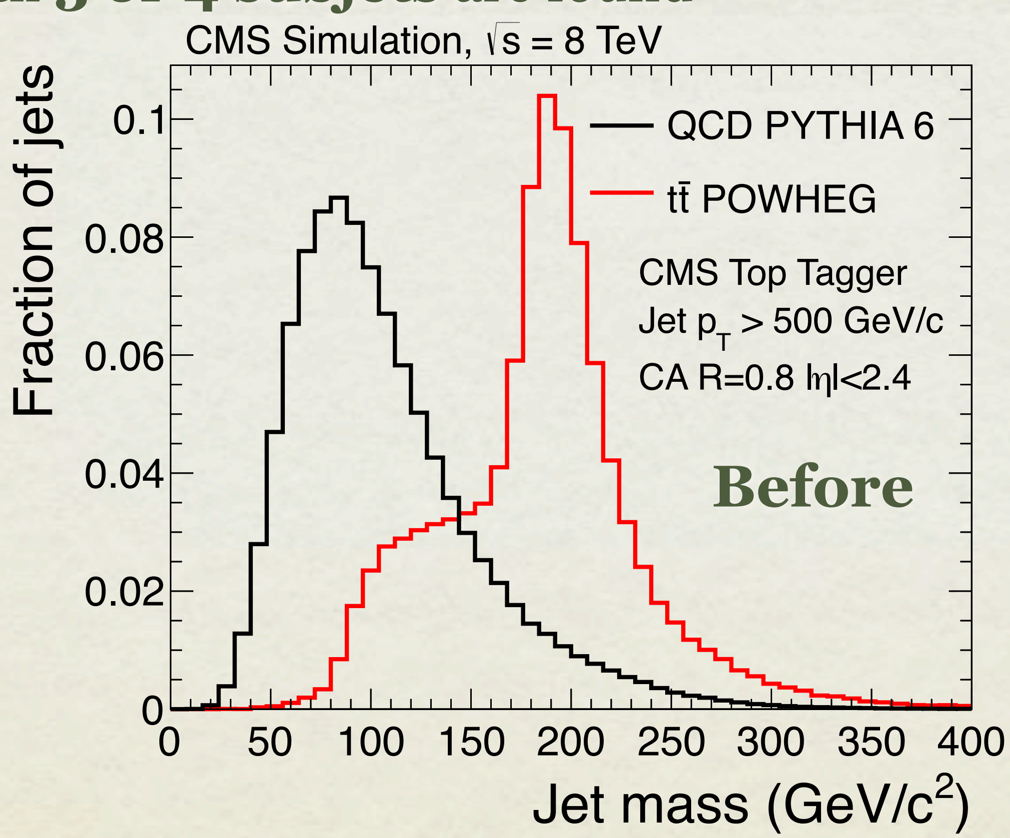
CMSTOP TAGGER

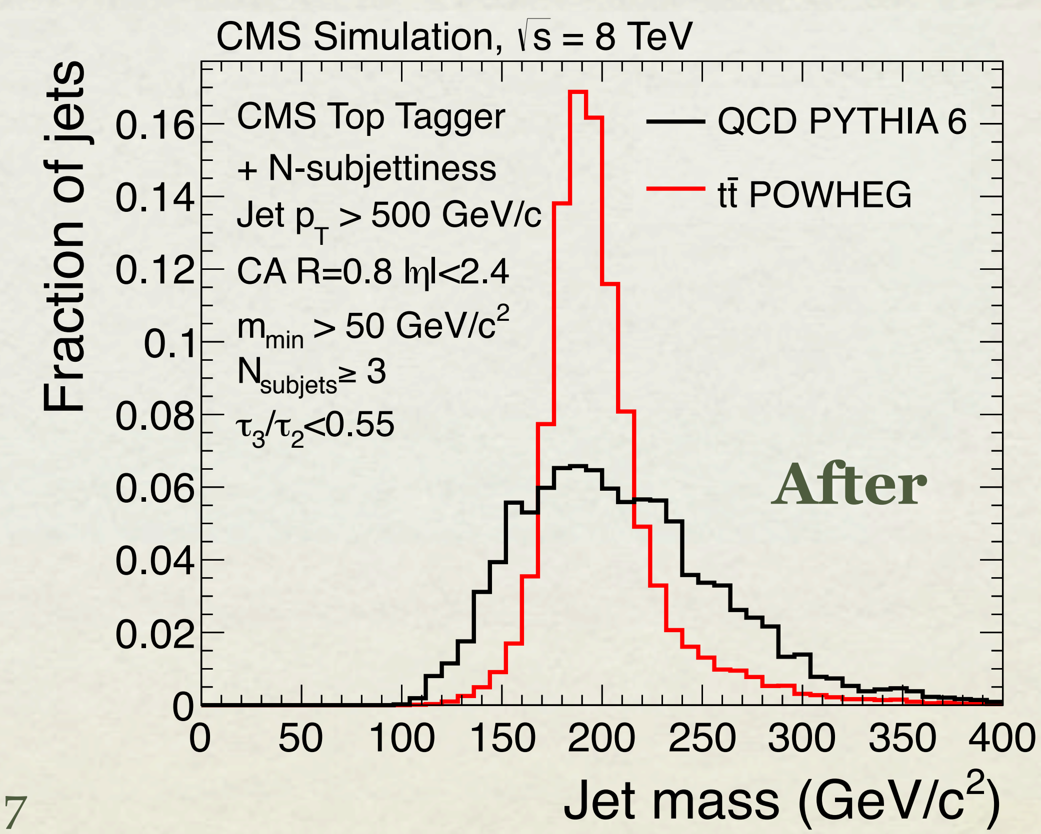
0806.0848

1204.2488

CMS PAS JME-13-007

- Start with **CA, R=0.8** large-R jets (p_T > 350 GeV)
- Repeatedly **uncluster** large-R jet
  - drop low-p_T (wrt/ large-R jet) subjets and close-by subjets
  - until **3 or 4 subjets** are found
- Selection criteria:
  - **mass** of sum has: [140,250] GeV
  - **minimal pairwise mass** of three leading subjets > 50 GeV
  - (tau_3/tau_2 <0.55)



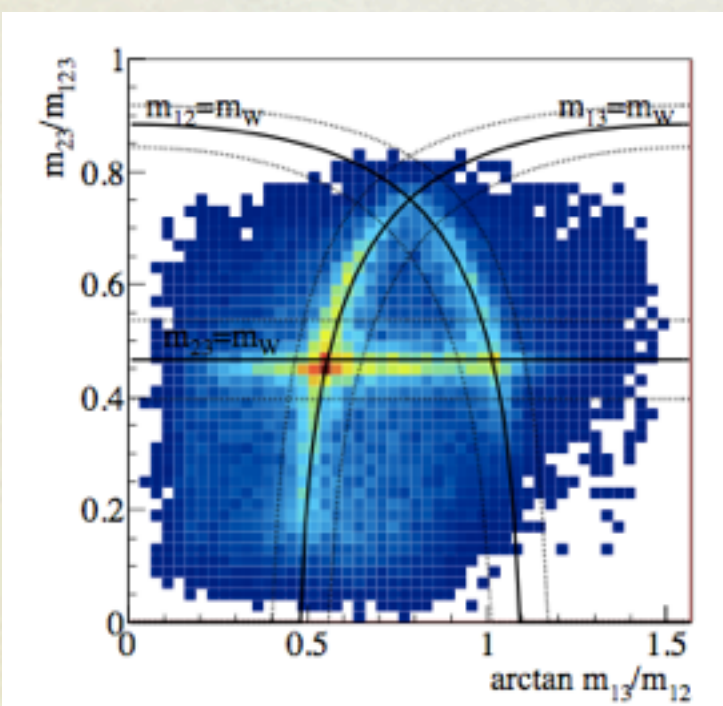


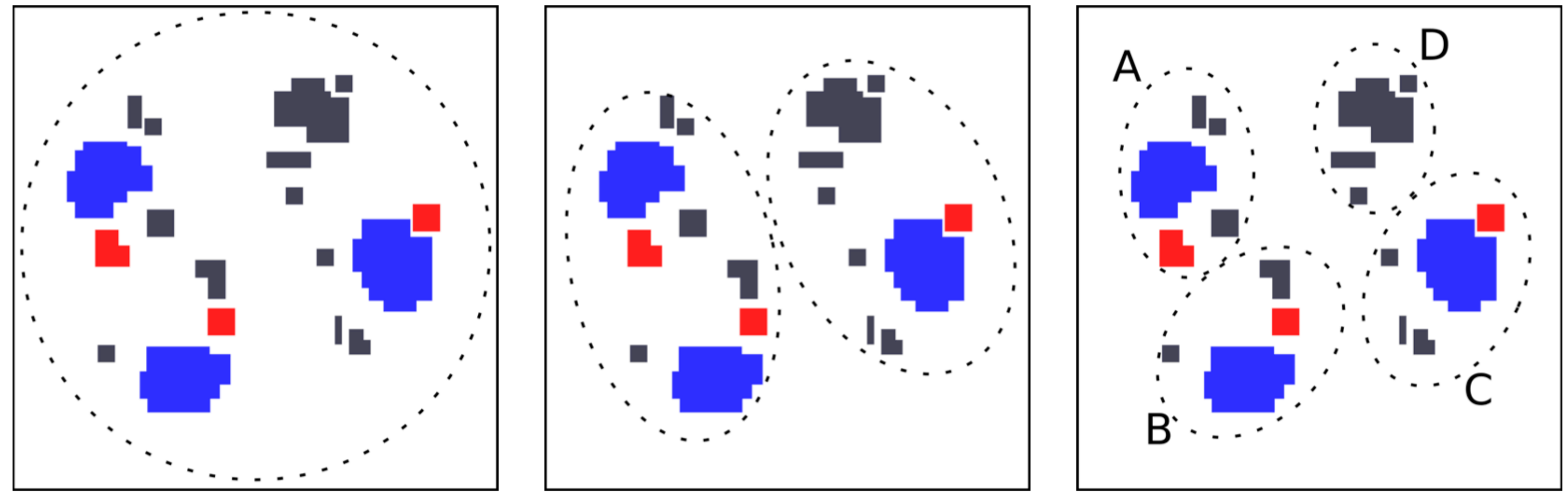
HEPTOPTAGGER

0910.5472

1006.2833

- Identify hadronically decaying top quarks with $p_T >$ 200 GeV
- Start with **CA, R=1.5** jets
- Identify top quark decay via mass ratios:

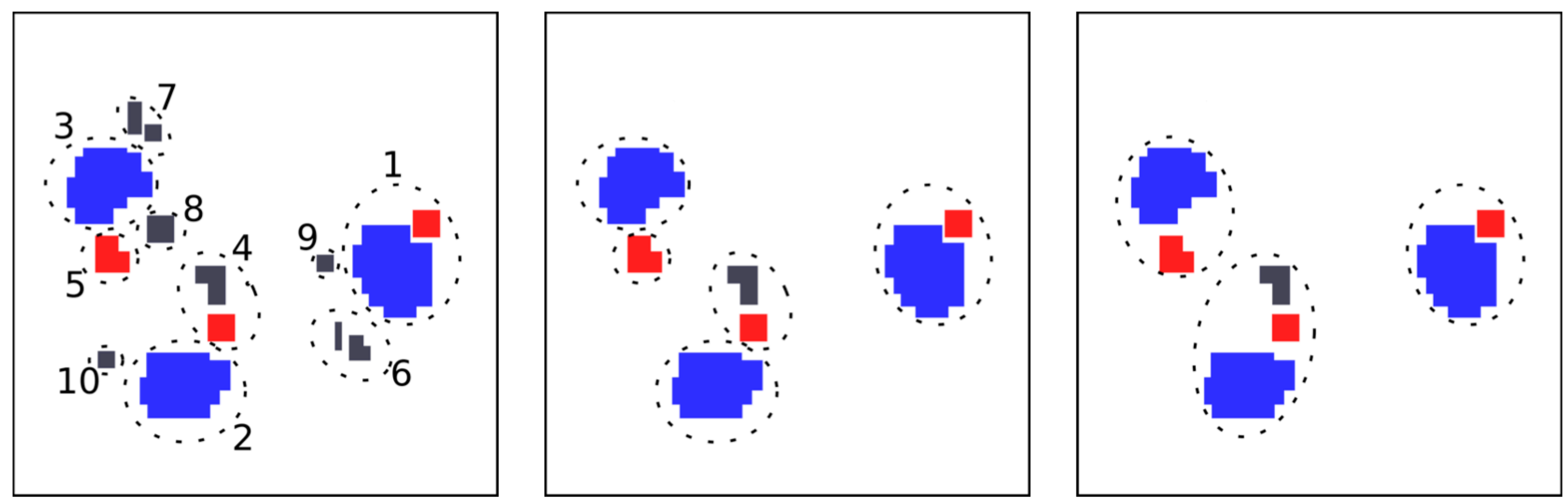




(a) initial fat jet

(b) fat jet after first unclustering step

(c) fat jet after the last unclustering step



(d) re-clustered constituents of the triplet of jets A, B and C from the previous step using a smaller distance parameter

(e) the five leading subjets found in the previous step

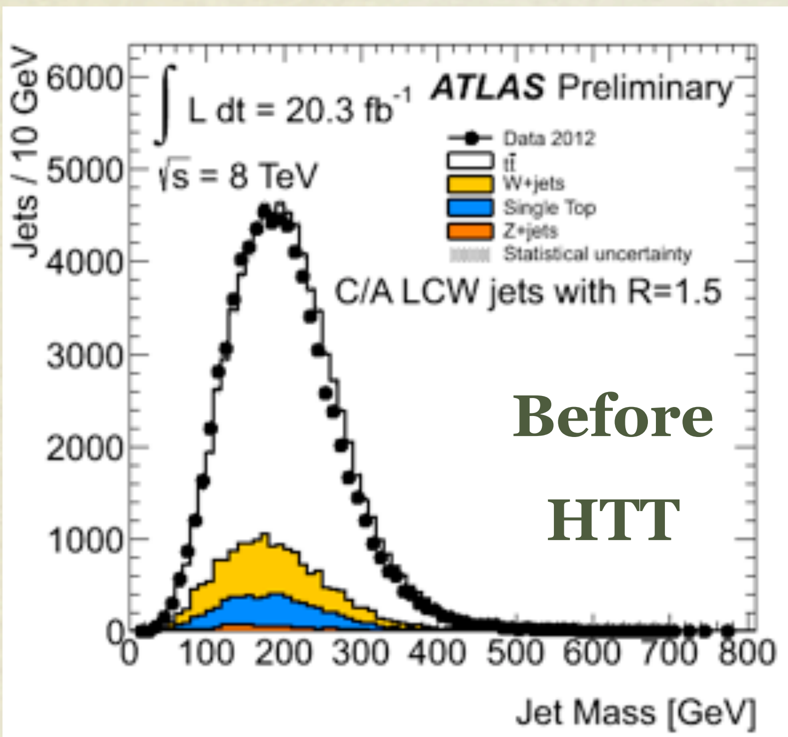
(f) constituents of the five jets from the previous step re-clustered into exactly three jets for testing mass ratios

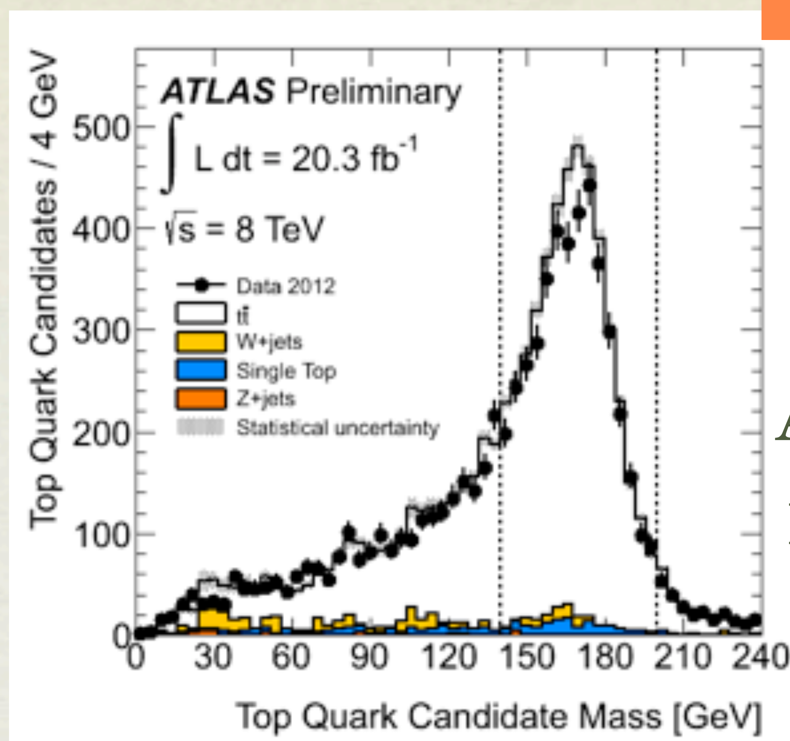
Monday, April 7, 14

CMS & HEP TAGGERS IN DATA

ATLAS-CONF-2013-084

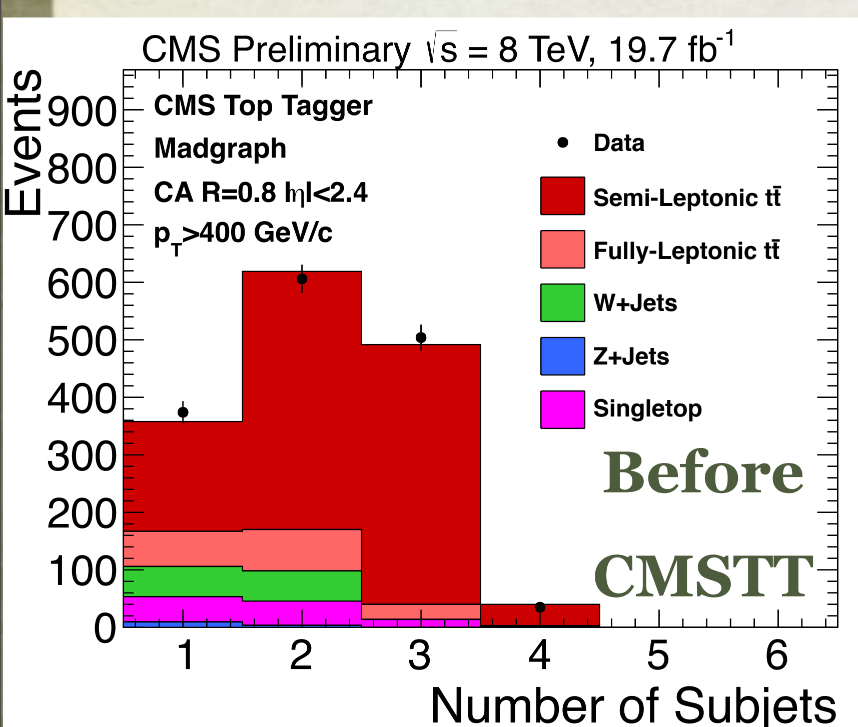
CMS PAS JME-13-007

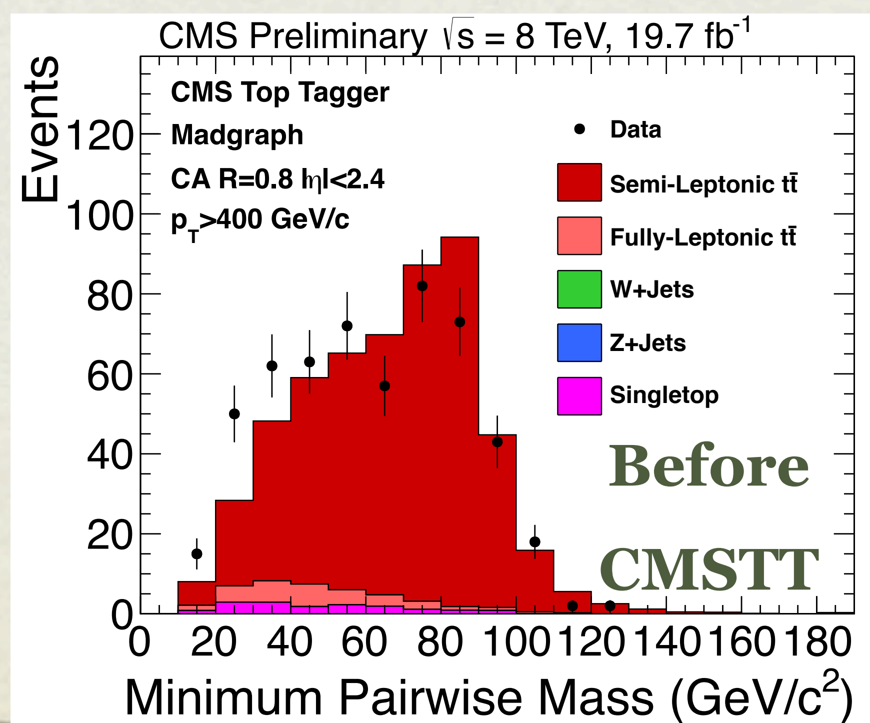


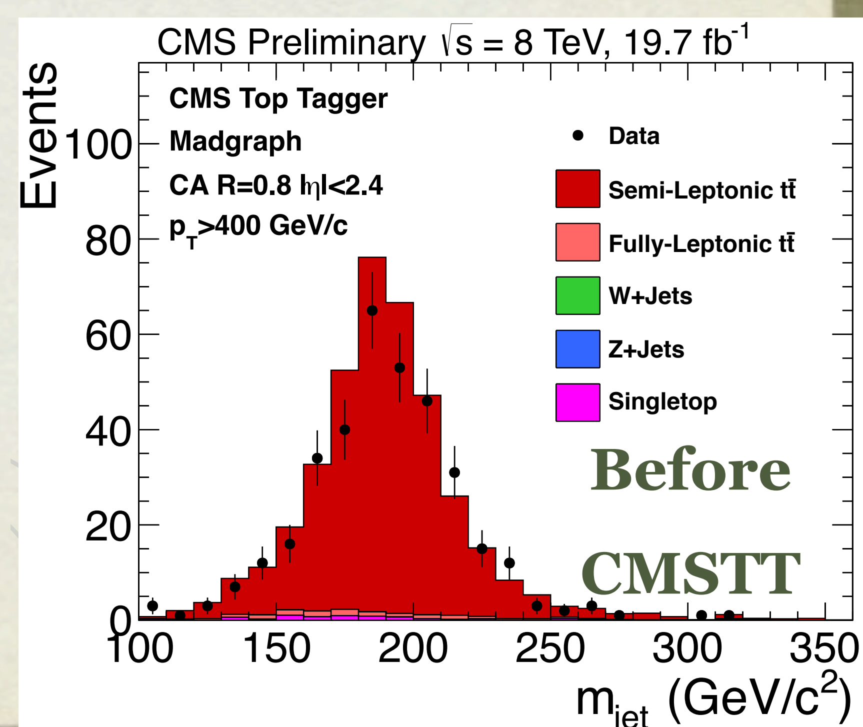


After

HTT





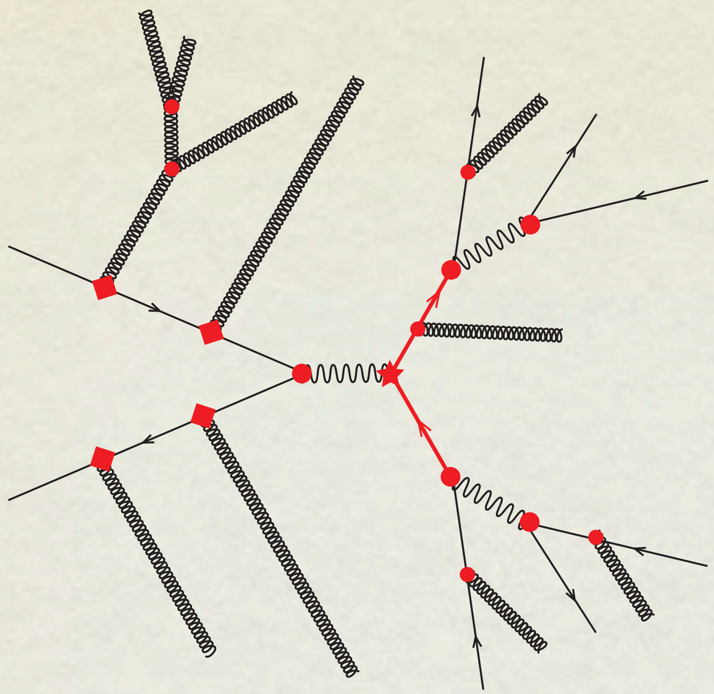


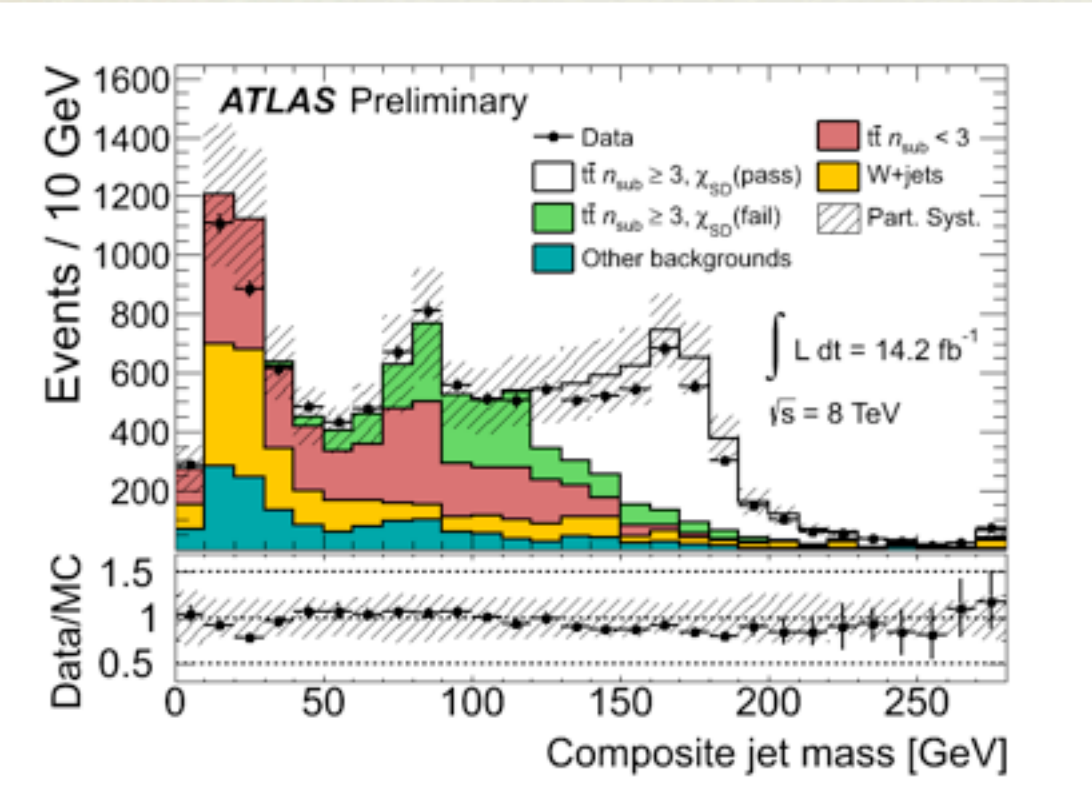
SHOWER DECONSTRUCTION

ATLAS-CONF-2014-003

1211.3140

1402.1189





- Assign signal and background probabilities to large-R jet:

$$\mathcal{P}(\{k\}_N, \text{model}) = \sum_h \mathcal{P}(\{k\}_N, \text{model}, h)$$

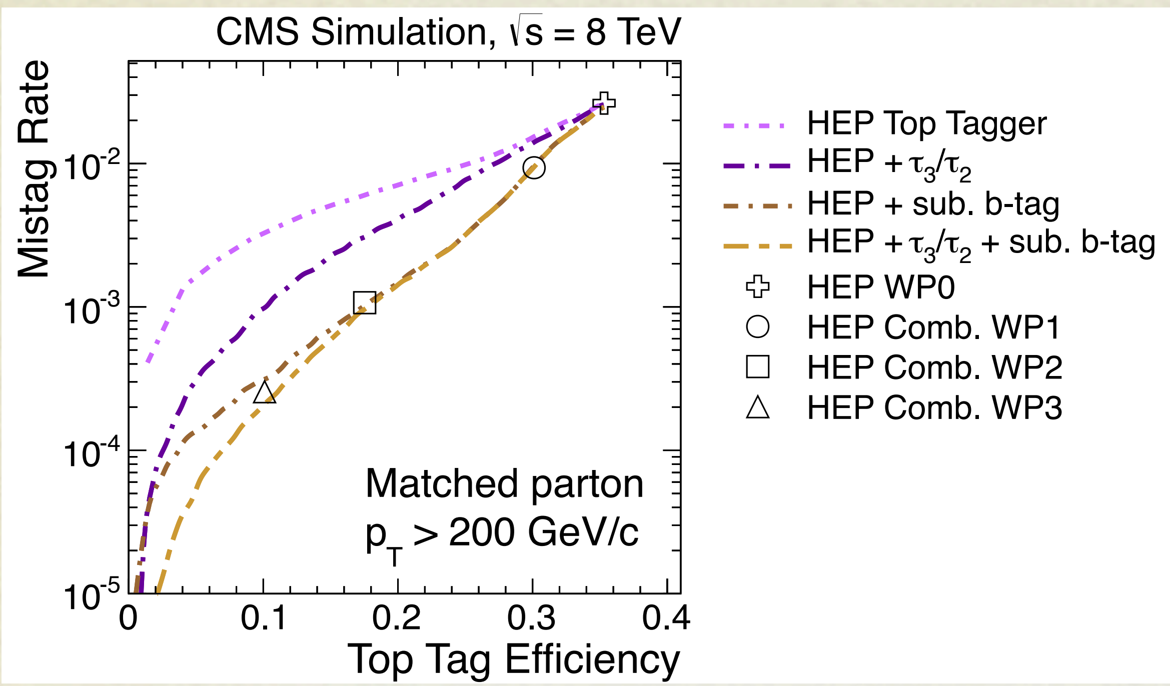
- Obtain likelihood:

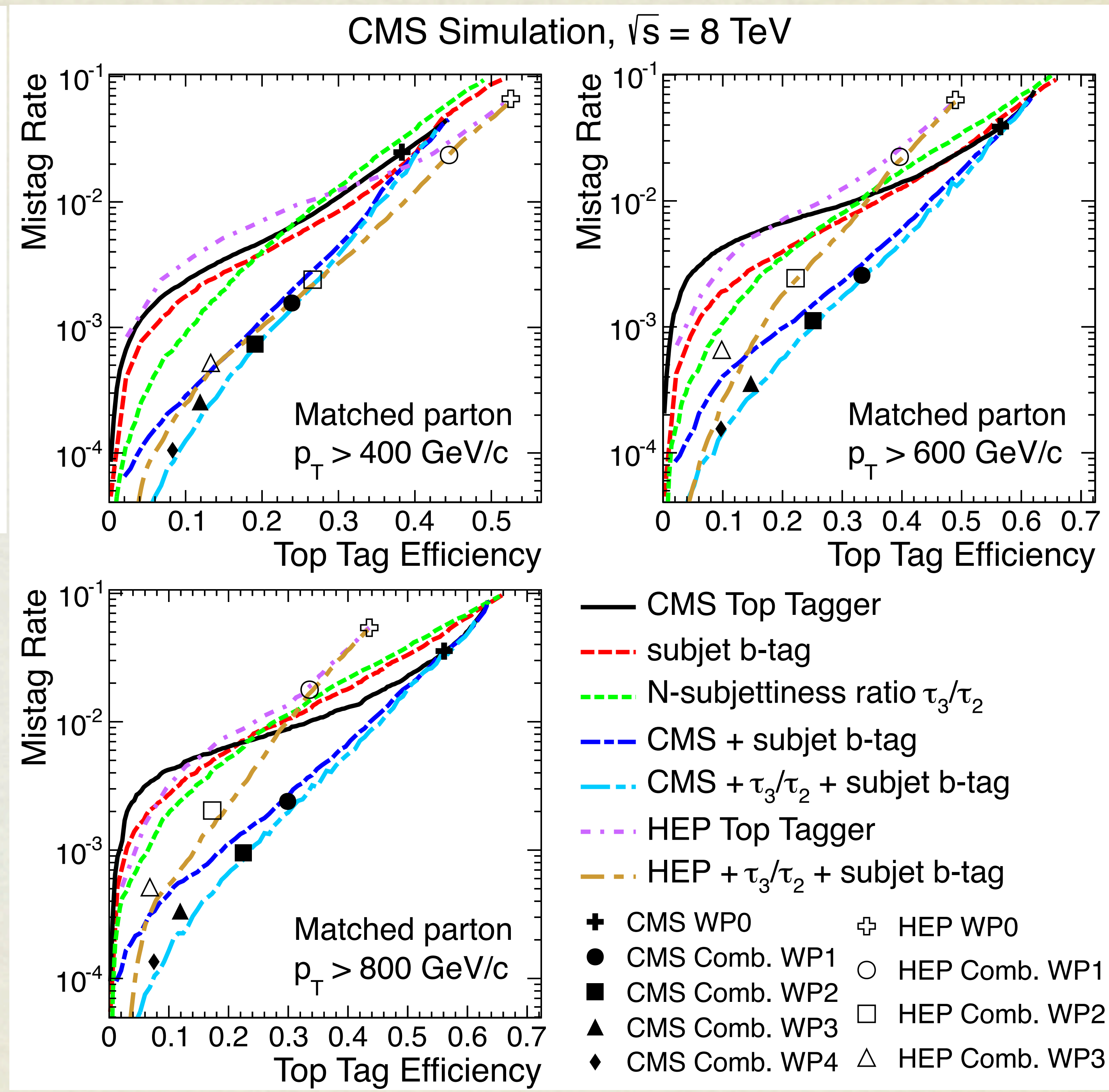
$$\chi = \frac{\mathcal{P}(\{k\}_N, \text{signal})}{\mathcal{P}(\{k\}_N, \text{background})}$$

- Start with AntiKt R=1.0 large-R jet and re-cluster with CA, R=0.2 (Microjets)
- Use Matrix element + parton shower with Sudakov factors to generate possible event histories

CMS ROCS

CMS PAS JME-13-007





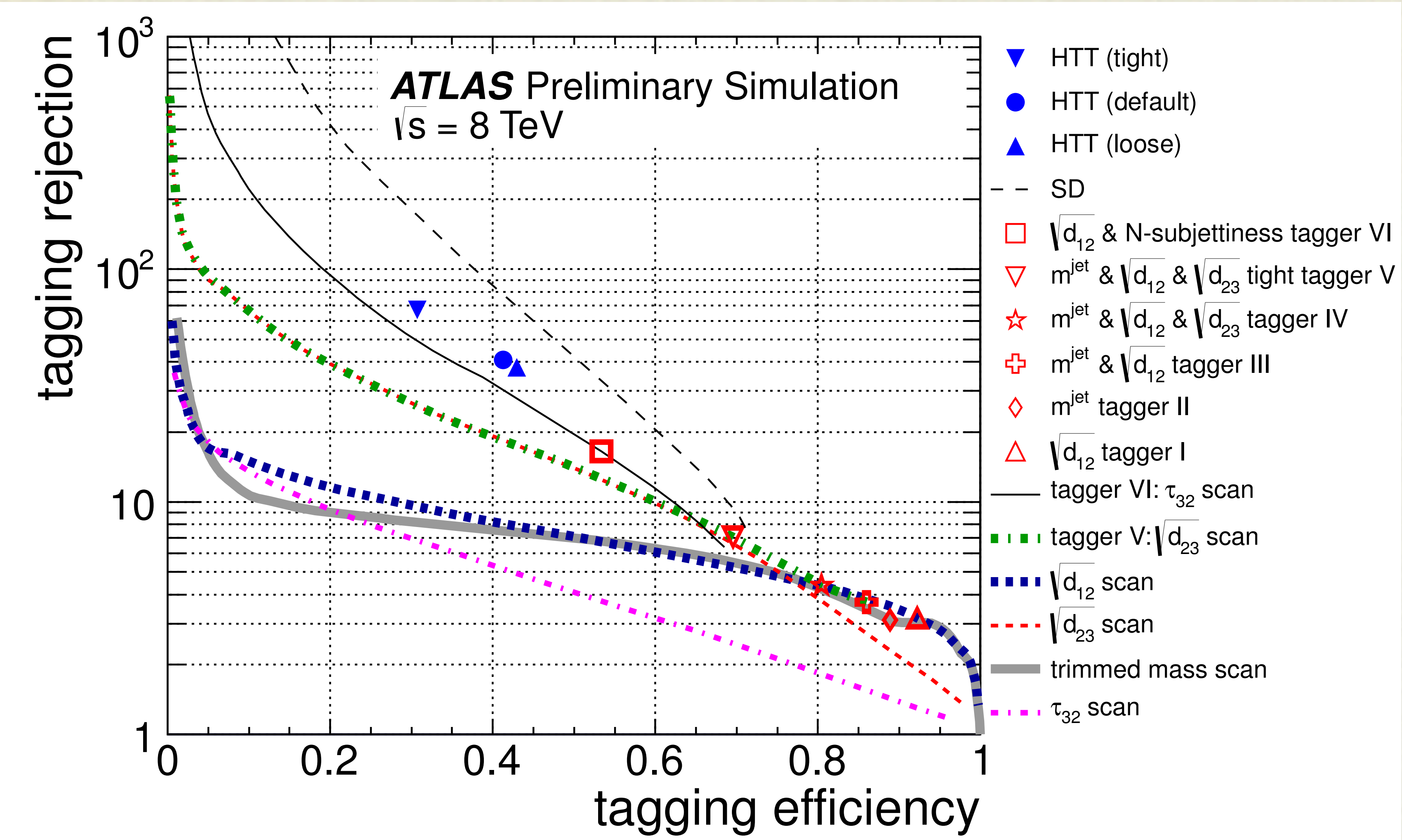
Denominator: jets matched to true top/ quark/gluon

Numerator: jets from Denominator that fulfill tagging criteria

ATLAS ROCS

ATLAS-CONF-2013-084

ATLAS-CONF-2014-003



Only using large-R jets with p_T > 550 GeV

Figure 7: Comparison of expected top jet tagging efficiency and light quark/gluon jet rejection. All substructure taggers and scans use trimmed anti-k_t R = 1.0 jets, except the HEPTopTagger (HTT) that uses C/A R = 1.2. The same $Z' \rightarrow t\bar{t}$, $m_{Z'}$ = 1.75 TeV signal samples and multijet background samples and selection are used for all taggers. Systematic uncertainties are not considered for any of the algorithms.

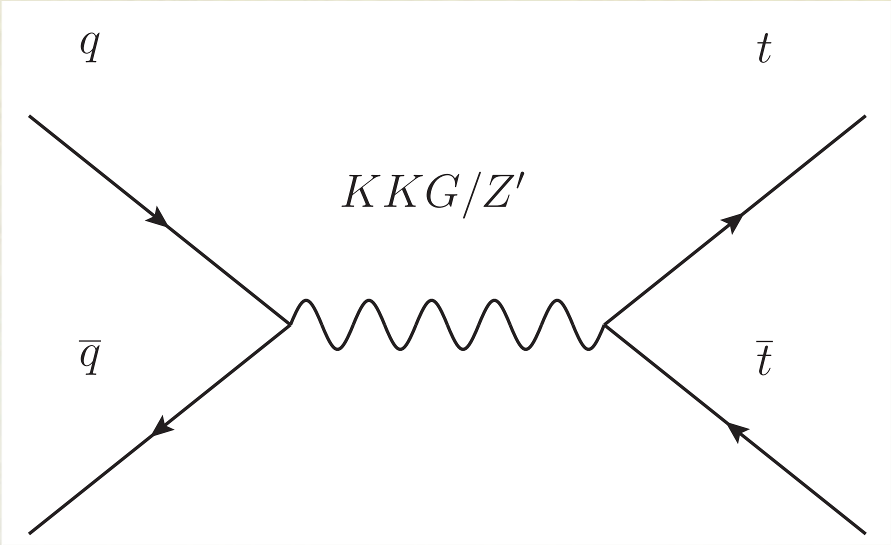
PHYSICS RESULTS

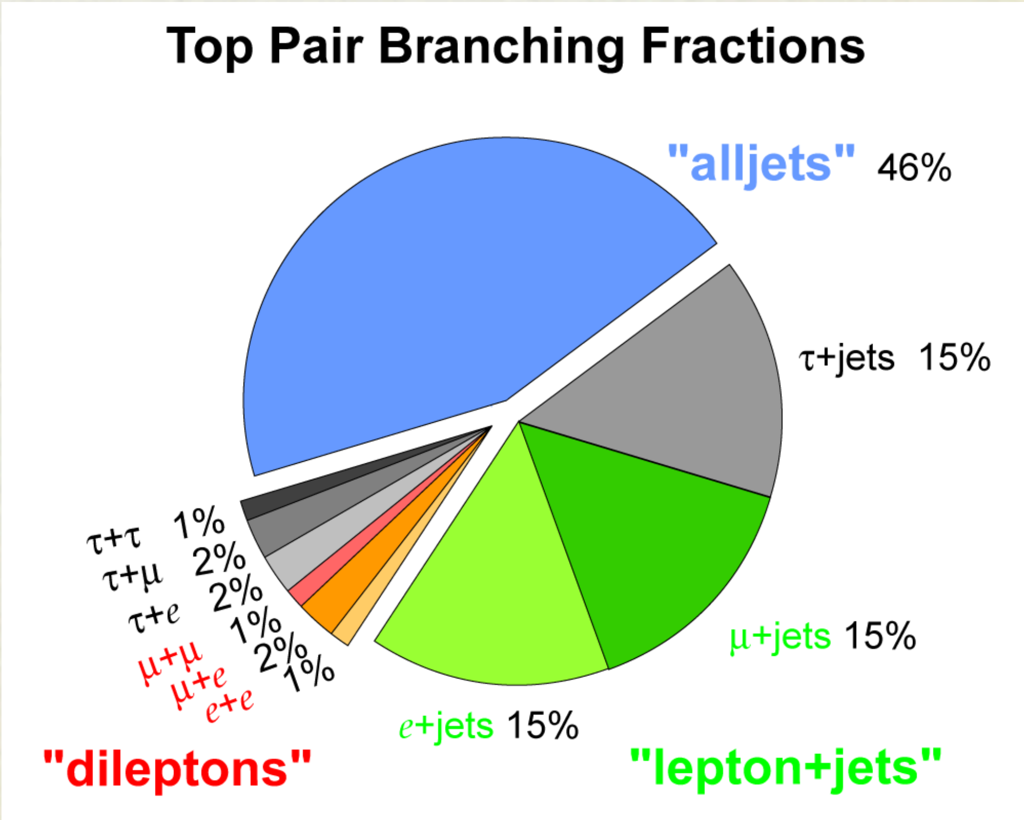
RESONANCES

1309.2030

ATLAS-CONF-2013-052

- Search for new particles decaying to a **pair** of **top** quarks
- Look for an **excess** over SM production in the **di-top** invariant **mass** distribution
- Benchmark models:
  - **Z' boson**
    - Leptophobic topcolor model
    - *Phys. Rev. D 49 (1994) 4454*
  - **Kaluza-Klein Gluon (KKG)**
    - Colored resonance
    - Randall-Sundrum extra dimension models
    - *Phys. Rev. D 76 (2007) 115016*





RESONANCES II

1309.2030

ATLAS-CONF-2013-052

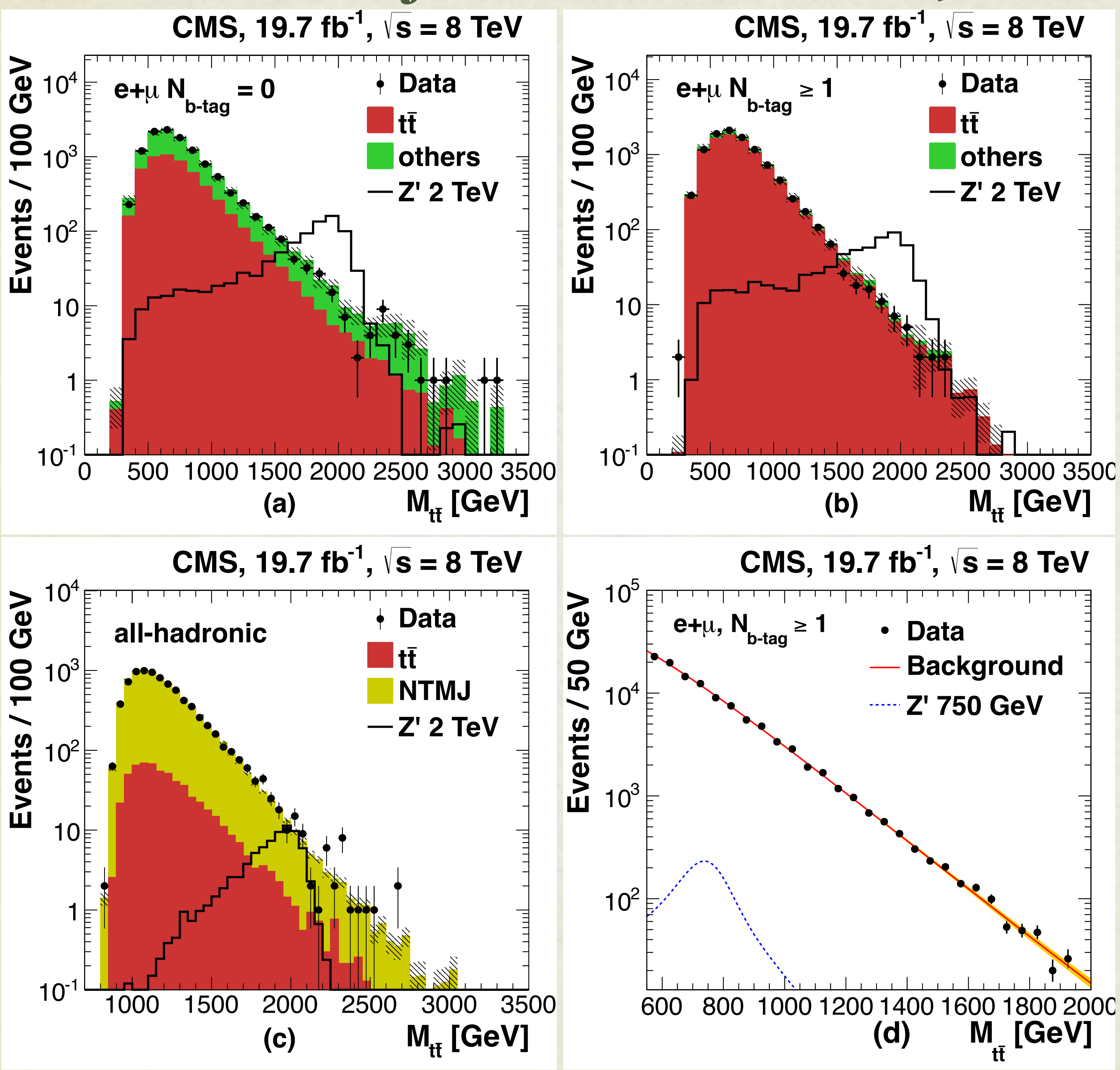
- CMS **Lepton+Jets** (**full 2012** data: 1309.2030)
  - **Resolved**: electron/muon + missing energy + 2 jets (AntiKt, R=0.5) + *b*-tags
  - **Boosted**: electron/muon + missing energy + 1 CMS-top-tagged jet + optional *b*-tags
  - Only combined Lepton+Jets Boosted/Resolved + All Hadronic limit
- CMS **All Hadronic** (**full 2012** data: 1309.2030)
  - 2 CMSTopTagger candidates
- ATLAS **Lepton+Jets** (**14 fb^{-1}**: ATLAS-CONF-2013-052)
  - **Resolved:** electron/muon + missing energy + 3-4 jets (AntiKt, R=0.4) +1 *b*-tag
  - **Boosted:** electron/muon + missing energy + 1 jet (AntiKt, R=0.4) + 1 trimmed AntiKt R=1.0 jet with cuts on kT-splitting scale + 1 *b*-tag
  - Only combined Lepton+Jets Boosted/Resolved limit
- ATLAS **All Hadronic** (**full 2011**:1211.2202, 2012 results not public yet)
  - 2 HEPTopTagger candidates or 2 template-method candidates + 2 *b*-tags

RESONANCES III

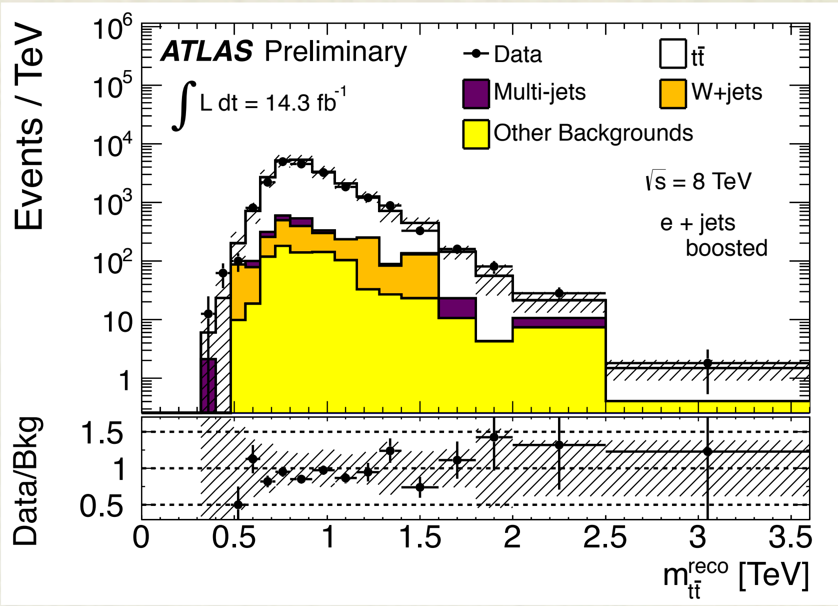
1309.2030

ATLAS-CONF-2013-052

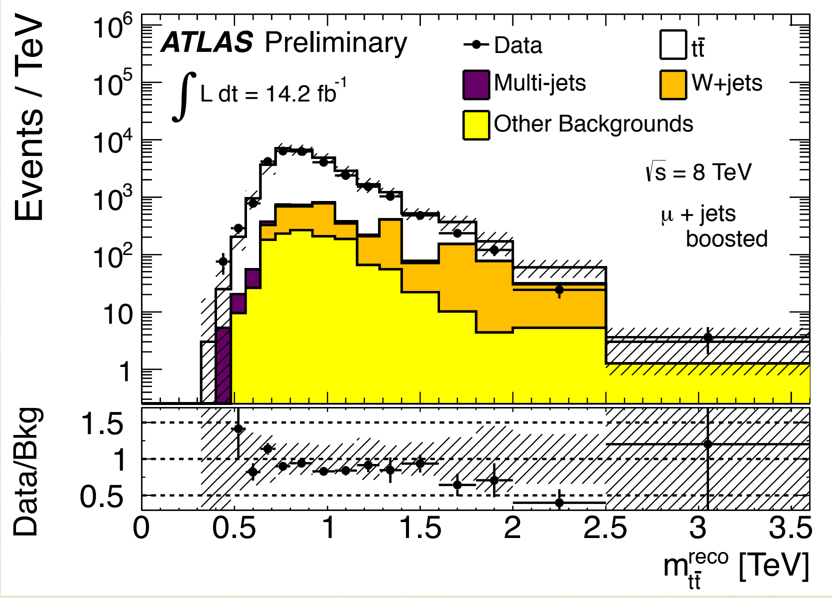
CMS, full 2012 (boosted lepton +jets and all hadronic)



ATLAS, 14 fb^{-1}, lepton+jets



(c) *e*+jets channel, boosted selection.



(d) *μ*+jets channel, boosted selection.

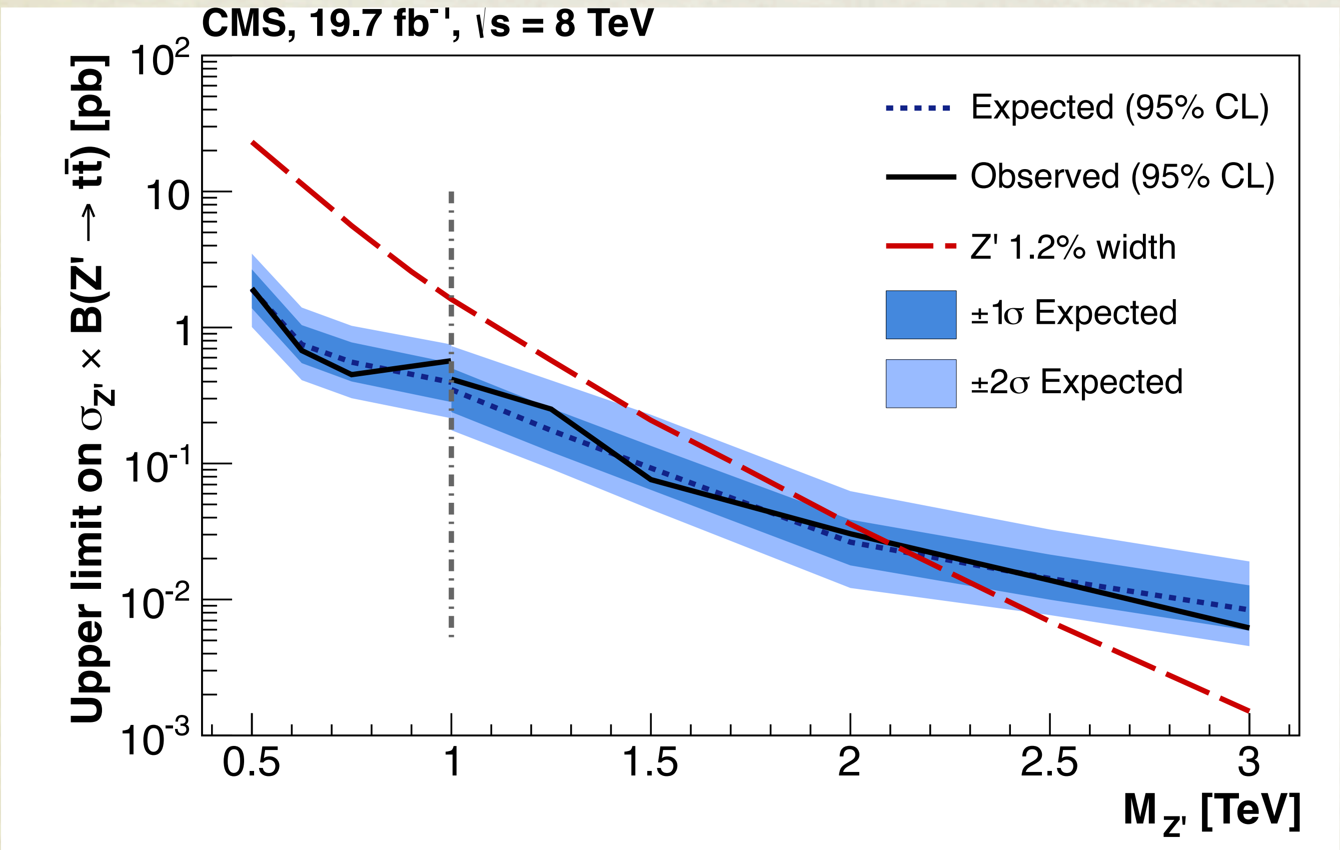
Monday, April 7, 14

RESONACES IV

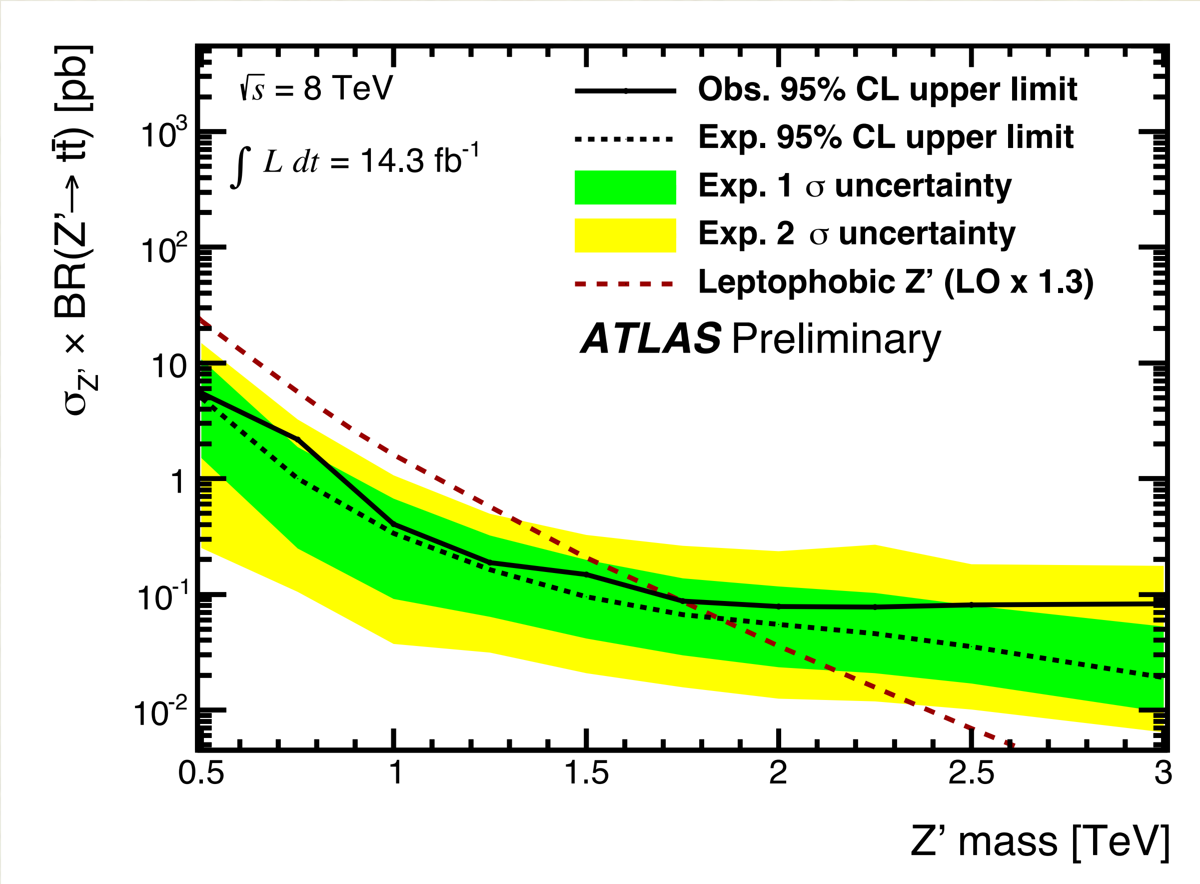
1309.2030

ATLAS-CONF-2013-052

CMS, full 2012 (lepton+jets and all hadronic)



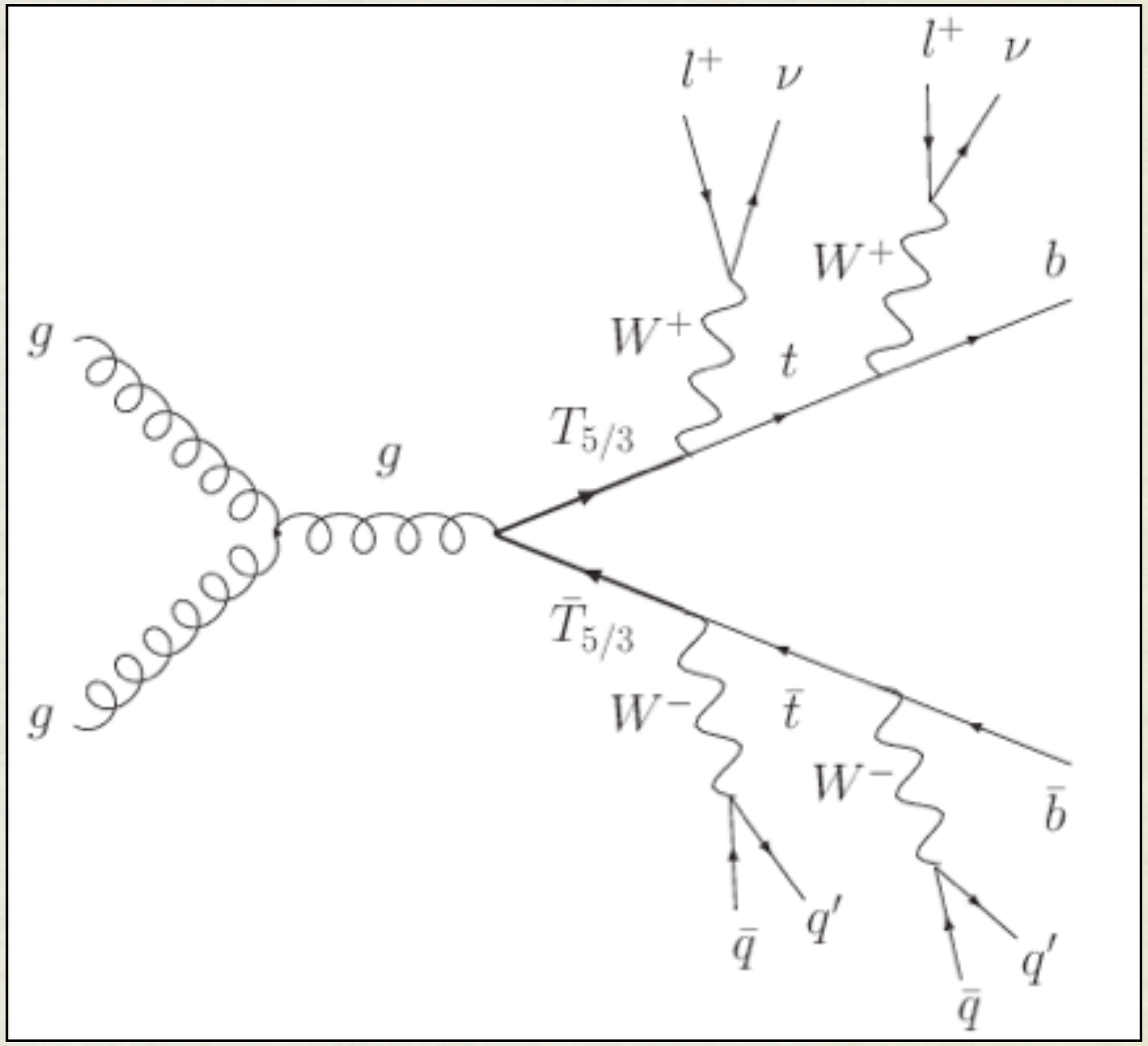
ATLAS, 14 fb^{-1}, lepton+jets



TOP-QUARK PARTNERS

1312.2391

- Di-top **resonance** searches are the main **testing ground** for boosted-top techniques
- Example for another analysis using boosted tops:
  - Search for exotic **top-quark partners** with **charge 5/3** in the same-sign dilepton final state
  - **Pair** production
  - Decay to **tW**

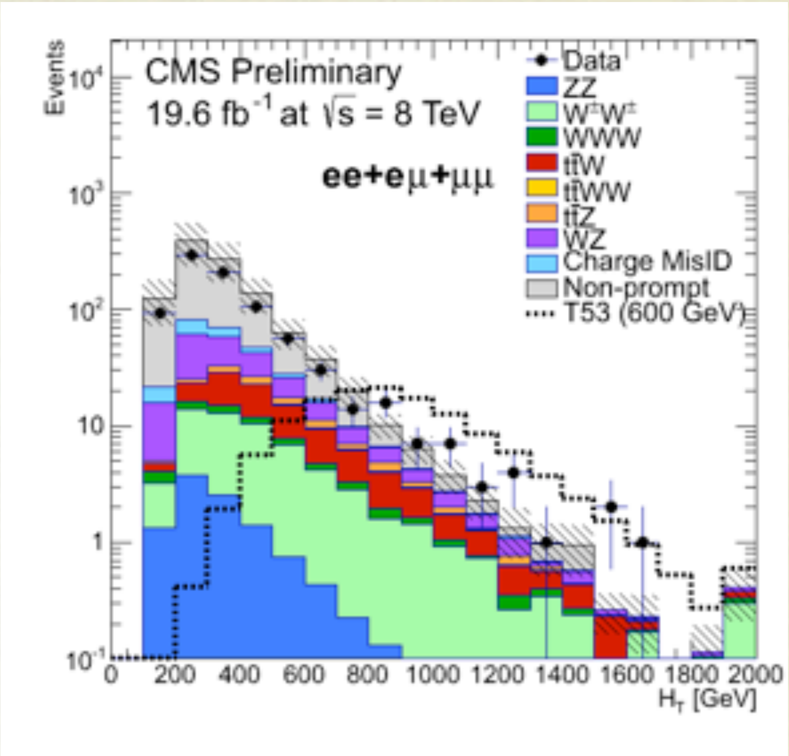


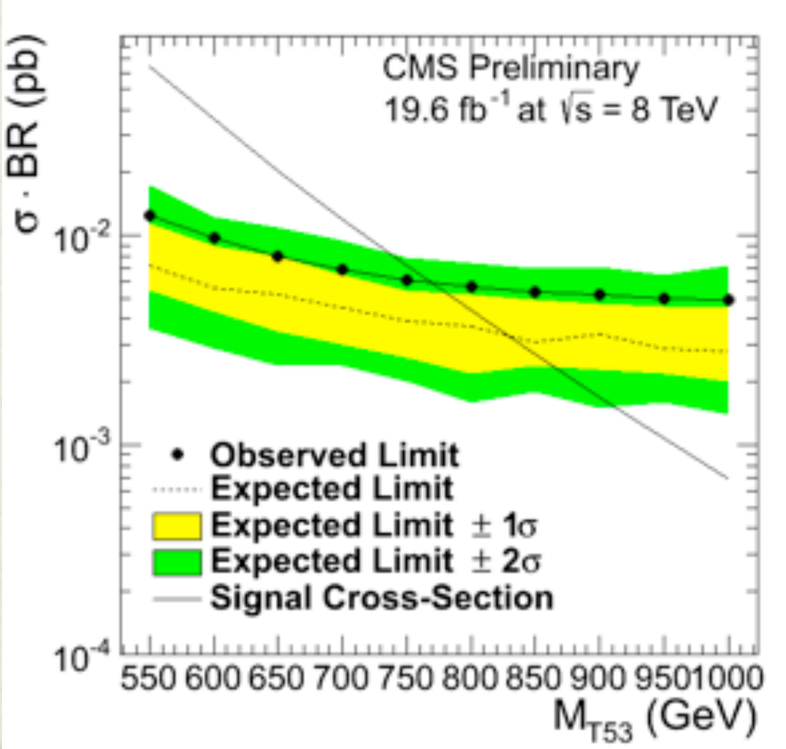
Monday, April 7, 14

TOP-QUARK PARTNERS II

1311.7667 1312.2391

- The **main** analysis **variable** is **H_T** = sum of selected jet and lepton p_T
- **Defaults** jets are reconstructed with **AntiKt (R=0.5)**
- If there are **CA (R=0.8)** top (CMSTopTagger) or W jets these are used for calculation of H_T **instead**
- Improves limit by **10-20%** for masses from **800 GeV to 1 TeV**
- The CMSTopTagger is also used as input for a BDT in a search for vector-like T quark with charge 2/3



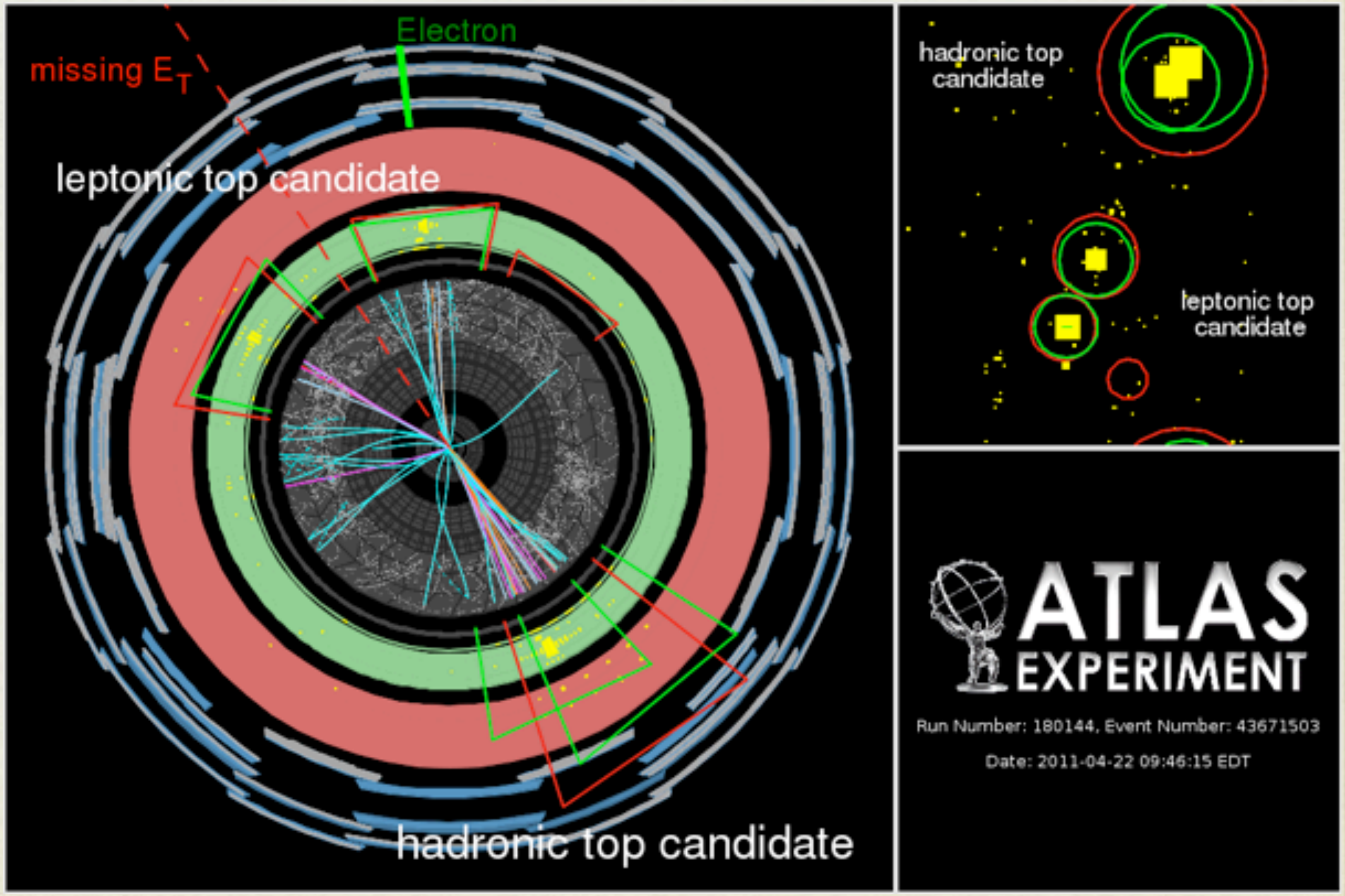


CONCLUSIONS

- Boosted reconstruction techniques aim to improve the **mass resolution** and **identify** top-quark decays using the **substructure** of large-R jets

- **Many different** boosted-top reconstruction techniques are being **tested** by ATLAS and CMS

- Top-tagging is essential for **boosted resonances** and is starting to play a role for other searches for BSM physics



BACKUP

JET CLUSTERING I

1112.4441

- Two approaches:
  - **Cone based** - Find stable axes of particles flow (ie. SisCone: 0704.0292)
  - **Sequential** - Pairwise combination of clusters according to a distance measure:

k_T	$d_{j_1 j_2} = \frac{\Delta R^2_{j_1 j_2}}{D^2} \min\left(p^2_{T,j_1}, p^2_{T,j_2}\right)$	$d_{j_1 B} = p^2_{T,j_1}$
Cambridge/Aachen	$d_{j_1 j_2} = \frac{\Delta R^2_{j_1 j_2}}{D^2}$	$y_{j_1 B} = 1$
anti-k_T	$d_{j_1 j_2} = \frac{\Delta R^2_{j_1 j_2}}{D^2} \min\left(\frac{1}{p^2_{T,j_1}}, \frac{1}{p^2_{T,j_2}}\right)$	$d_{j_1 B} = \frac{1}{p^2_{T,j_1}}$.

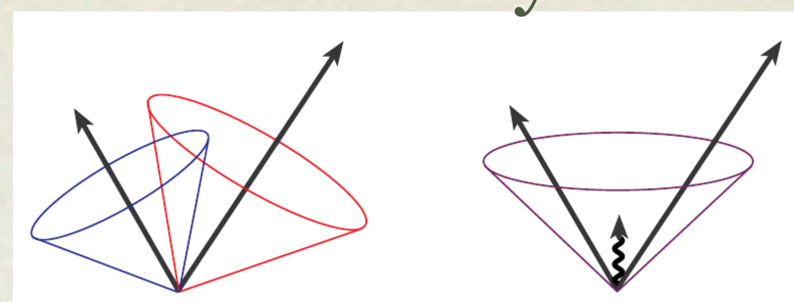
JET CLUSTERING II

hep-ex/0005012

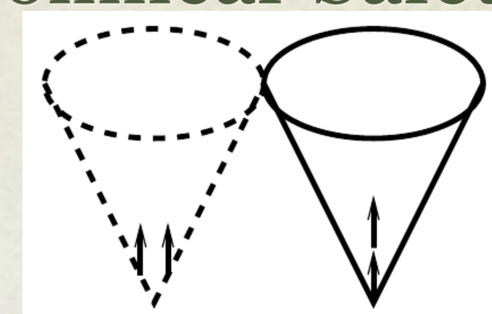
0802.1189

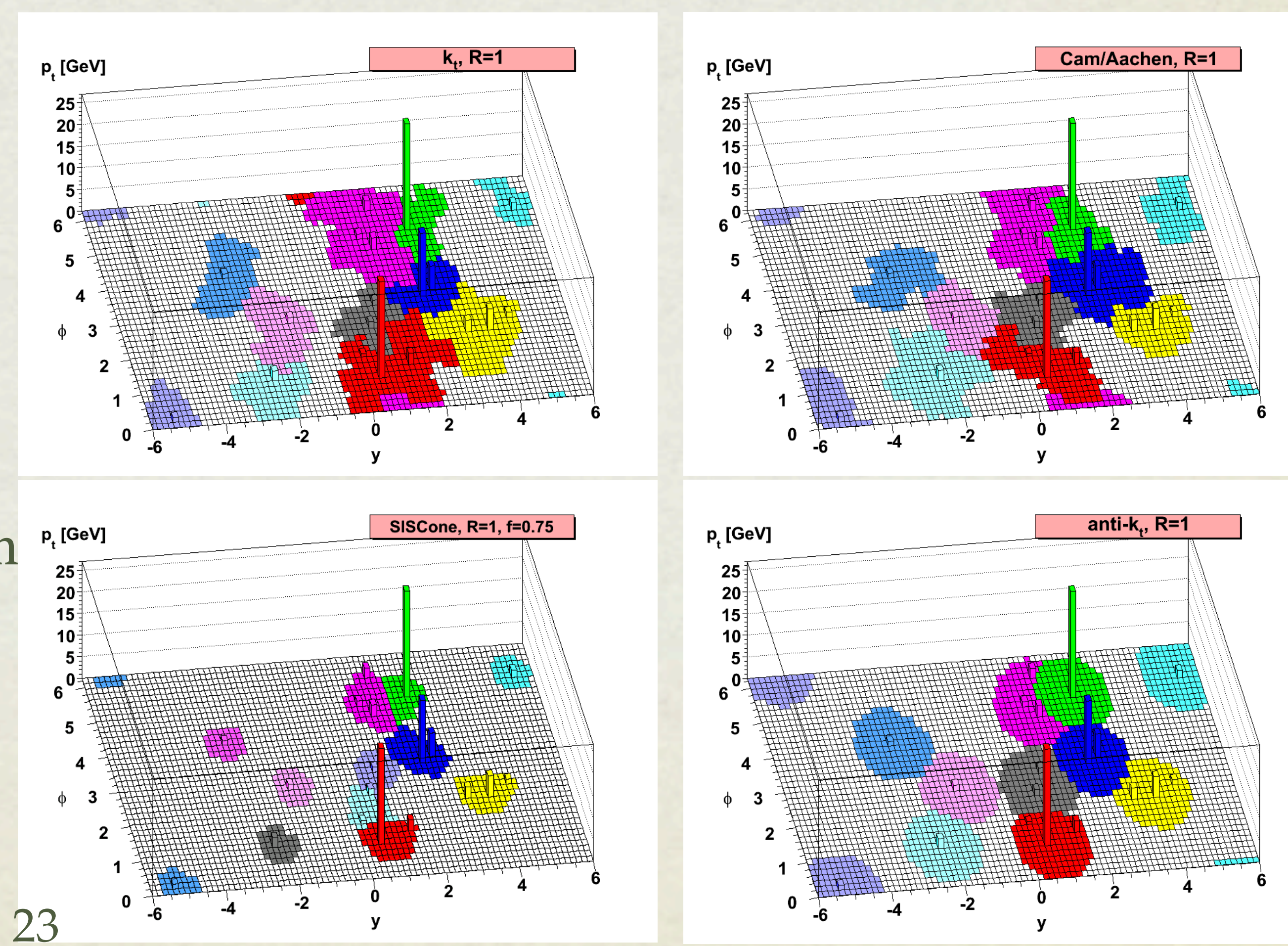
- Important Properties:
  - Infrared safety
  - Collinear safety
- Most commonly used is **AntiKt**
  - R=0.5 and 0.7 for CMS
    R=0.4 and 0.6 for ATLAS
  - Resilient, circular boundaries
- For large-R jet studies **CA** and **kT** are very comm
  - Substructure can be better interpreted
    in terms of QCD

IR safety



Collinear Safety

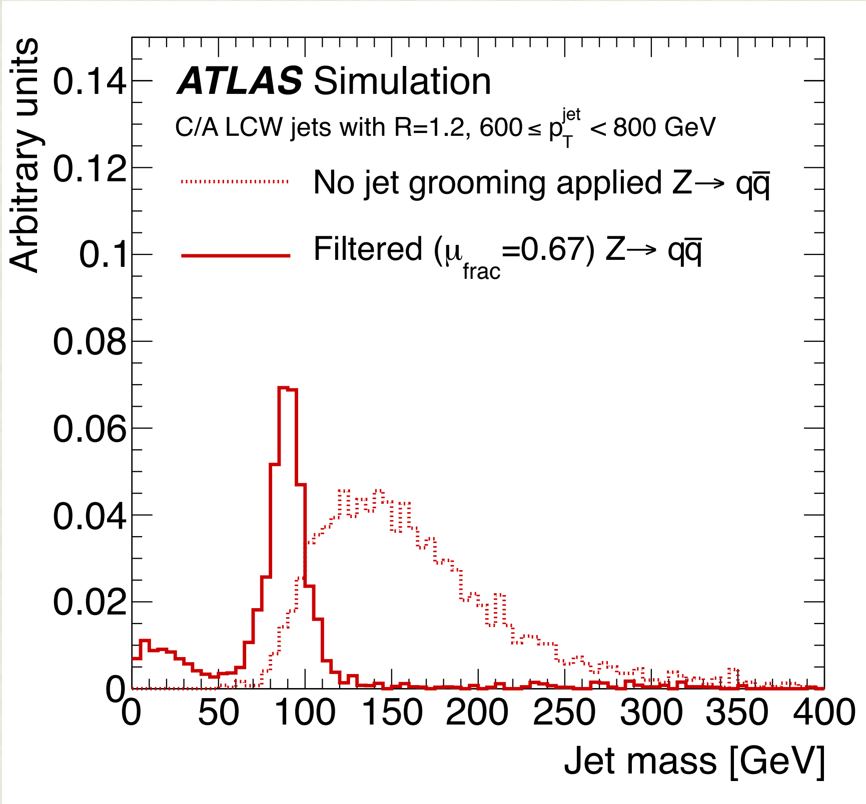




GROOMING

1306.4945

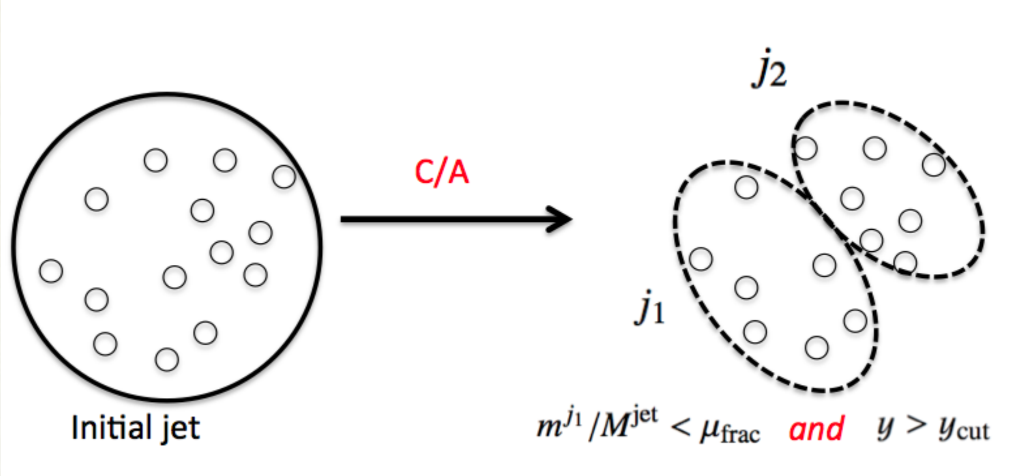
- Idea: Start with **large-R jet** (R=0.8-1.8) and **remove** constituents from
  - **pile-up**
  - **rest** of the event
- Arrive at the **mass** of the **initial** hard particle (Z, t, Higgs,..)
- Different techniques:
  - Mass Drop
  - Filtering
  - Trimming
  - Pruning



MASS DROP

1306.4945

0802.2470



- Mass Drop:
  - **Undo** last **clustering** step (j ->j1 + j2, mj1>mj2)
  - If there was a **large mass drop** and the splitting is **not** too **asymmetric** (default: mu=0.67, y=0.09):

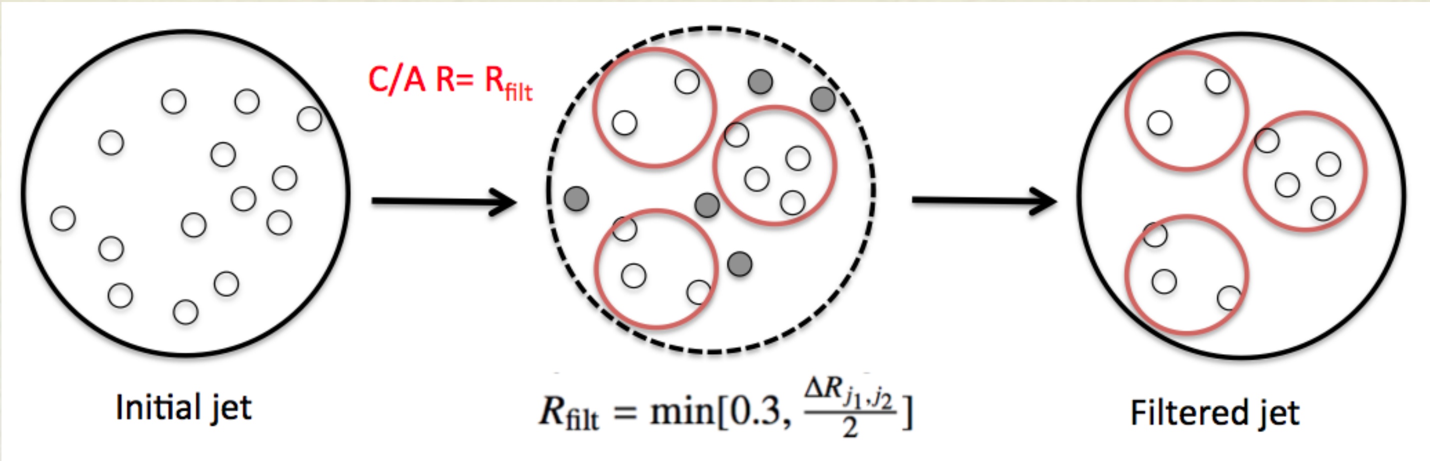
$$m^{j_1}/m^{\text{jet}} < \mu_{\text{frac}} \qquad \frac{\min(p_{tj_1}^2, p_{tj_2}^2)}{m_j^2} \Delta R_{j_1,j_2}^2 > y_{\text{cut}}$$

  - Then consider j1, j2 as interesting substructure and **exit** the loop
  - Else: **Repeat** with j1

FILTERING

1306.4945

0802.2470



- Take constituents and **re-cluster** using the CA algorithm and a **smaller** distance parameter:

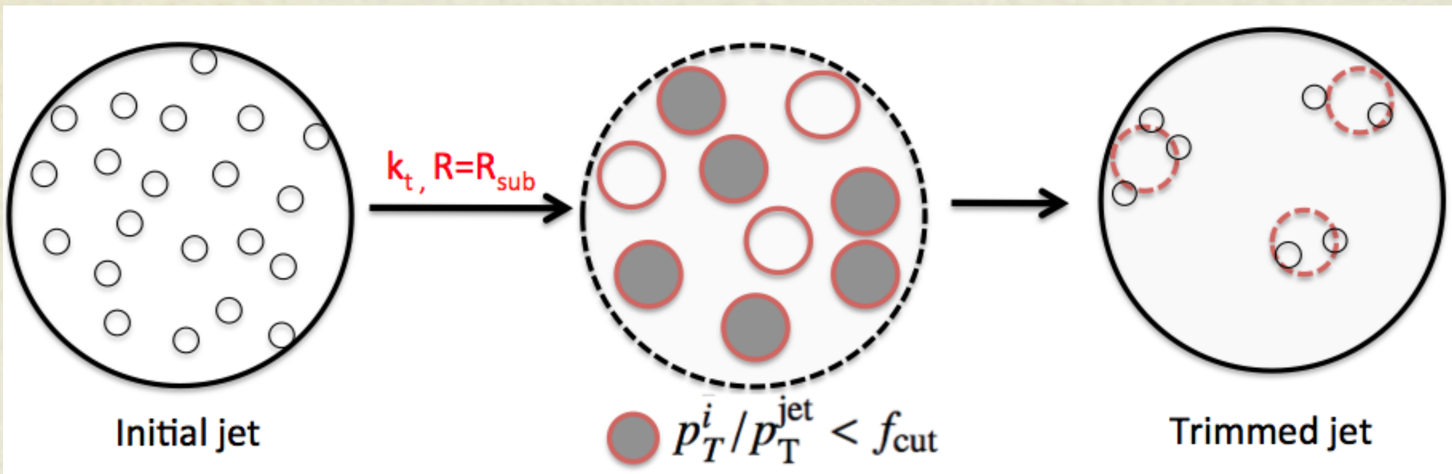
$$R_{\text{filt}} = \min[0.3, \Delta R_{j_1,j_2}/2]$$

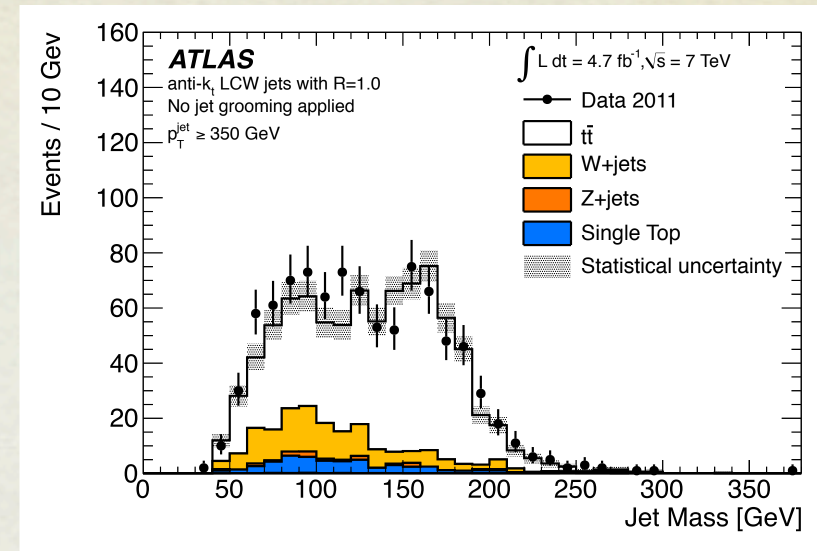
- Usually use j1 and j2 resulting from **mass-drop** as input
- Keep **n-leading** (in pT) subjets (typically 3..5)
- Dynamic filtering radius increases sensitivity to collimated decays (high pT)

TRIMMING

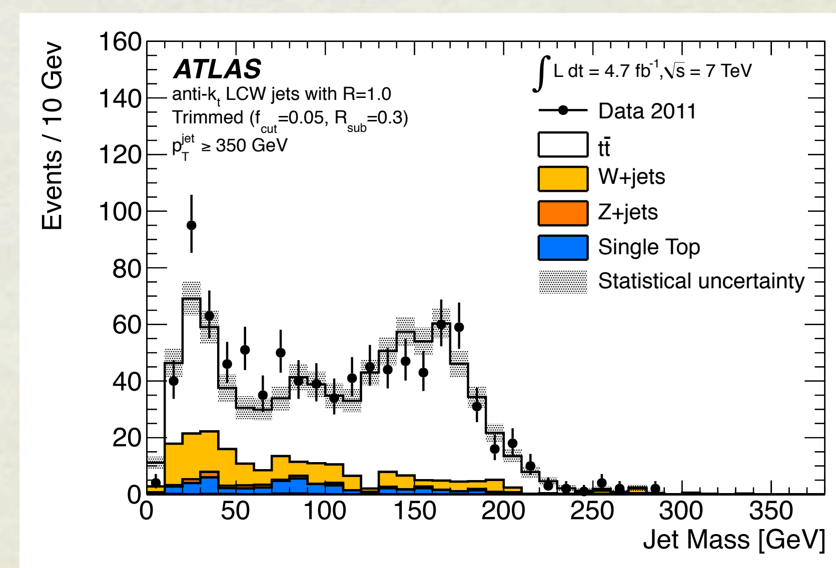
1306.4945

0912.1342





(a) anti-k_t, $R = 1.0$, ungroomed



(b) anti-k_t, $R = 1.0$, trimmed

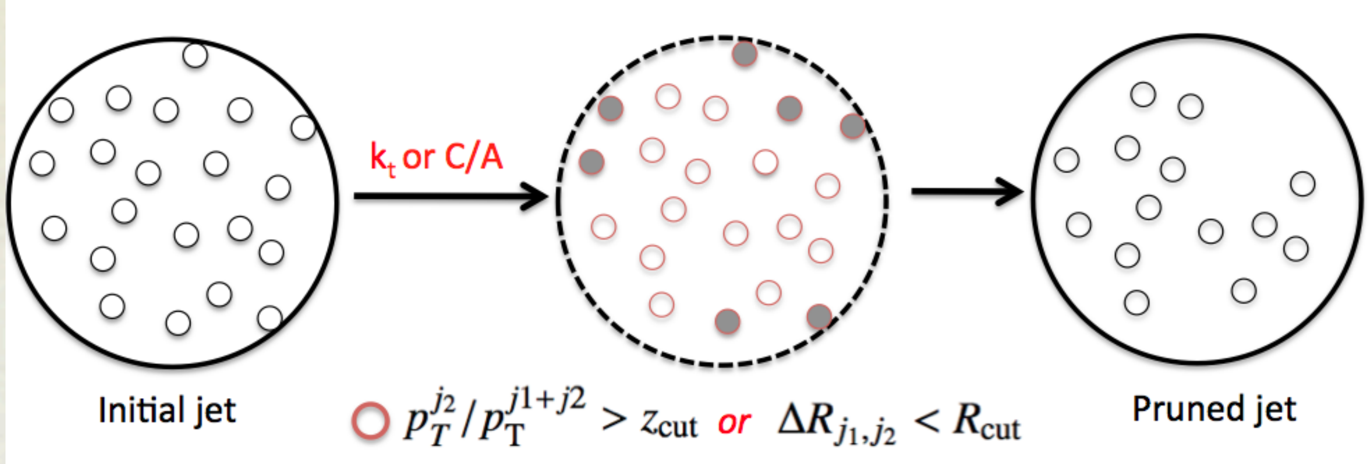
- **Radiation** from..
  - pile-up
  - multi-parton interactions
  - initial-state radiation
- ..is typically **softer** than decay products of hard interaction
- **Recluster** using kT algorithm with R=Rsub (eg. 0.2)
- **Remove** all subjets with pT/pT_Jet < f_cut (eg. 0.03)

PRUNING

1306.4945

0903.5081

0912.0033

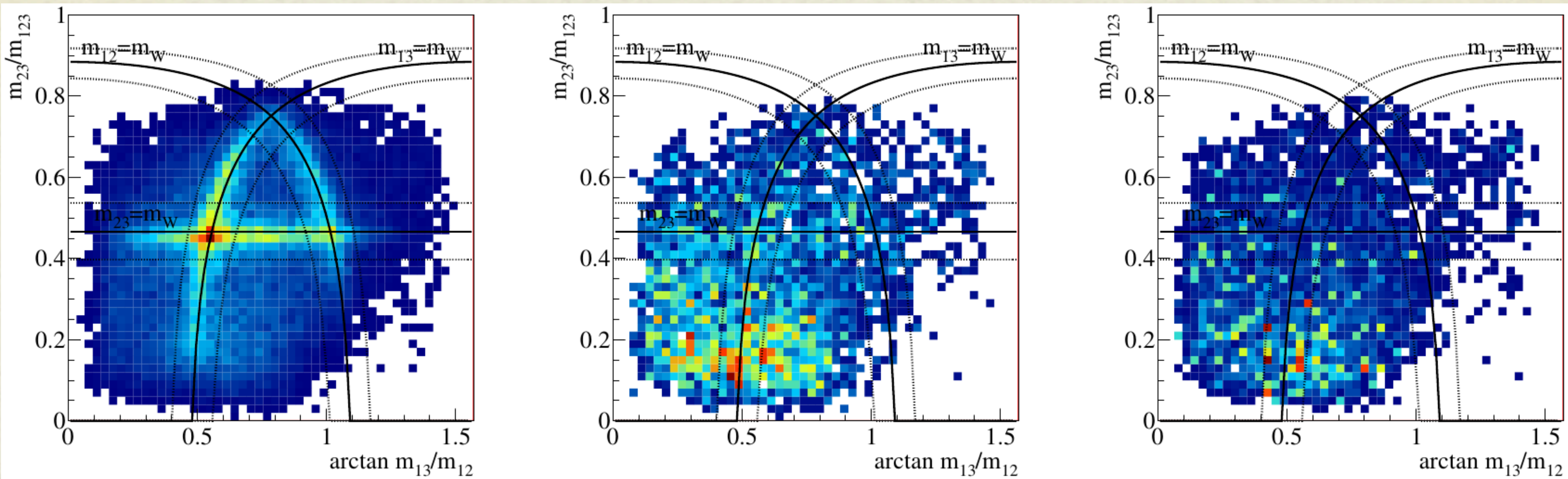


- Similar to trimming, also **removes small pT** constituents
- Additional **veto** on **wide-angle** radiation
- Rerun kT/CA clustering and apply criteria at each stage
- If one of $p_T^{j_2}/p_T^{j1+j2} > z_{cut}$ or $\Delta R_{j_1,j_2} < R_{cut} \times (2m^{jet}/p_T^{jet})$ is satisfied: **merge**
- Otherwise: **discard** j2 and **continue**

HTT ALGORITHM

0910.5472

1006.2833



Top W QCD

$$0.2 \arctan \frac{m_{13}}{m_{12}} < 1.3 \text{ and } R_{min} < \frac{m_{23}}{m_{123}} < R_{max}$$

$$R_{min}^2 \left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{max}^2 \left(1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35$$

$$R_{min}^2 \left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{max}^2 \left(1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35.$$

TEMPLATE OVERLAP METHOD

1211.2202 **1006.2035**
1212.2977 **1112.1957**

- Calculate templates (sets of "n" four-momenta which satisfy the kinematic constraints of the decay products of a boosted massive jet)
- Test the agreement of the jet constituents with each signal template
- Important difference to MEM: no background hypothesis

$$Ov^{(F)}(i,j) = max_{TS}\, exp\left[-\sum_f \frac{1}{2\sigma_f^2}\left(\sum_j (E_j - E_f)F(f,j)\right)^2\right]$$

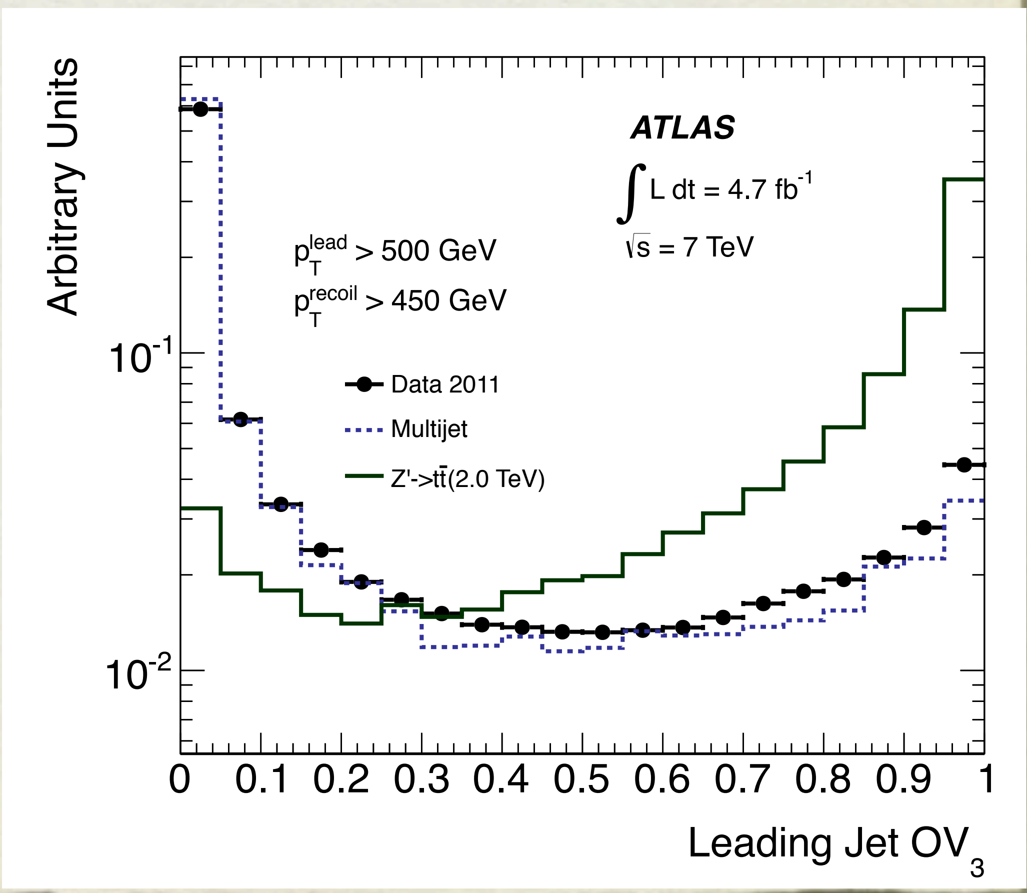
template resolution (typically E(f)/3)

j - jet "constituent"

TS - template "space"

f - template momentum

The kernel F restricts the angular region around each template momentum



From Mihailo Backovic's talk at Boost 2013

ATLAS RESONANCES

ATLAS-CONF-2013-052

Systematic Uncertainties	Resolved selection yield impact [%] total bkg.	Resolved selection yield impact [%] Z'	Boosted selection yield impact [%] total bkg.	Boosted selection yield impact [%] Z'
Luminosity	2.9	4	3.3	4
PDF	2.9	5	6	2.9
ISR/FSR	0.2	–	0.7	–
Parton shower and fragm.	5	–	4	–
$t\bar{t}$ normalization	8	–	9	–
$t\bar{t}$ EW virtual correction	2.2	–	4	–
$t\bar{t}$ Generator	1.5	–	1.6	–
W+jets $b\bar{b}$+$c\bar{c}$+c vs. light	0.8	–	1.0	–
W+jets $b\bar{b}$ variation	0.2	–	0.4	–
W+jets c variation	1.1	–	0.6	–
W+jets normalization	2.1	–	1.0	–
Multi-Jet norm, e+jets	0.6	–	0.3	–
Multi-Jet norm, μ+jets	1.8	–	0.3	–
JES, small-radius jets	6	2.2	0.7	0.5
JES+JMS, large-radius jets	0.3	4	17	3.3
Jet energy resolution	1.6	0.4	0.6	0.7
Jet vertex fraction	1.7	2.3	2.1	2.4
b-tag efficiency	4	1.8	3.4	6
c-tag efficiency	1.4	0.3	0.7	0.9
Mistag rate	0.7	0.3	0.7	0.1
Electron efficiency	1.0	1.1	1.0	1.0
Muon efficiency	1.5	1.5	1.6	1.6
All systematic uncertainties	14	9	22	9

CMS RESONANCES

1309.2030

Table 1: Constraints used in the likelihood maximization. The $M_{t\bar{t}}$ distributions of the boosted channels are combined into a single joint likelihood, imposing consistency of the various background and signal components.

	Resolved Semi-Leptonic	Boosted Semi-Leptonic	Boosted All-Hadronic
Constraints on normalization			
Luminosity [41]	2.6%	2.6%	2.6%
Pileup	6%	6%	6%
$t\bar{t}$ [42]	–	15%	15%
Parton distribution functions [43]	1σ	1σ	1σ
Single top	–	50%	–
W+light-flavor jets	–	50%	–
W+heavy-flavor jets	–	100%	–
Z+ jets	–	100%	–
Lepton selection	0.5-3.0%	0.5-3.0%	–
Top-tagging efficiency	–	–	9%
Constraints on shape			
$t\bar{t}$ renormalization, factorization and matching scales	–	variation by $\times 2$ and $\times 0.5$	
Jet energy scale	1–6%	1–6%	1–6%
Jet energy resolution	8–10%	8–10%	8–10%
b-tagging efficiency	2–8%	2–8%	–
b-tagging mis-ID	–	20%	–
Top-tagging mis-ID	–	–	5–20%
Signal and background pdf	1σ	–	–