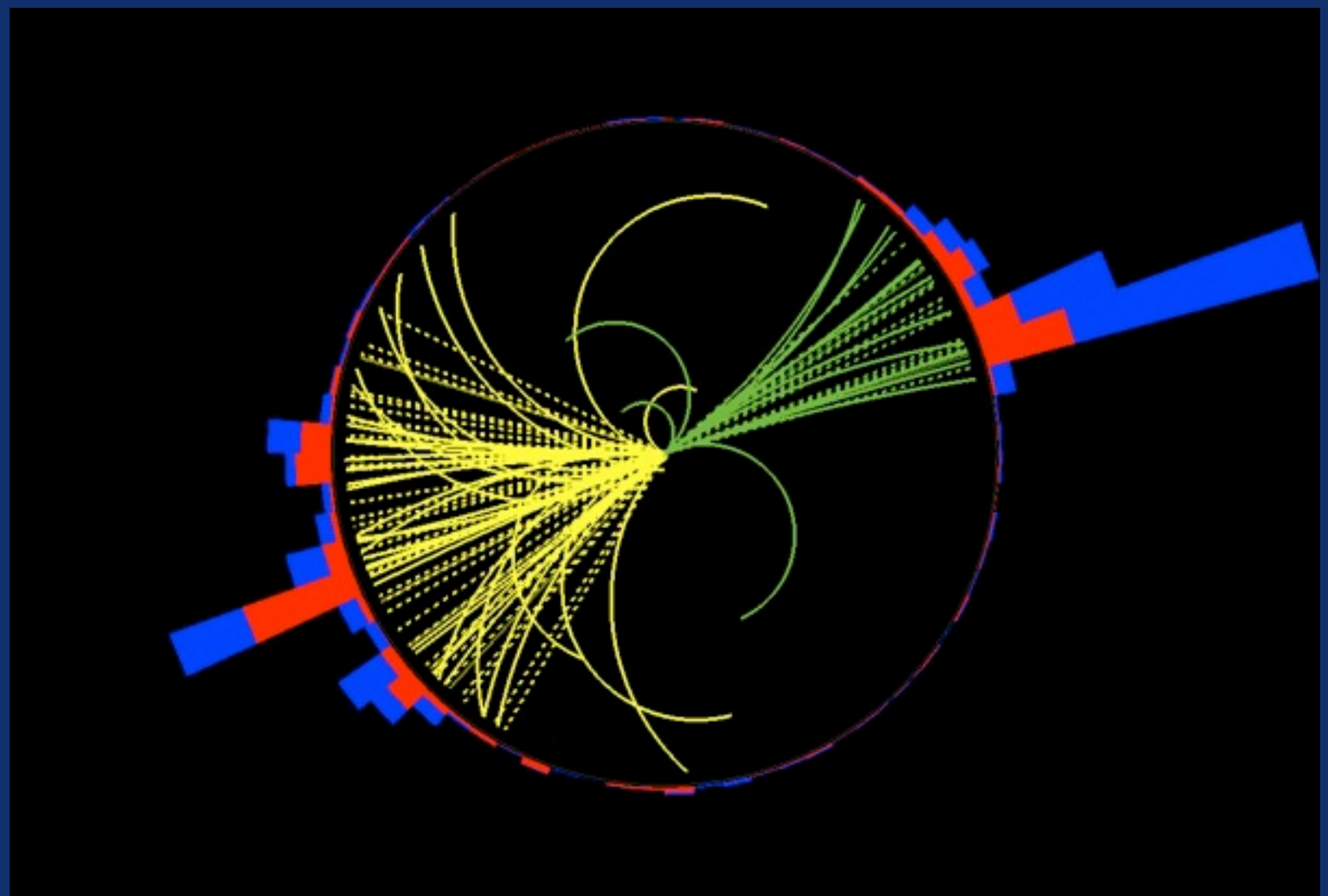
Jets and Jet Substructure

A Quick Overview

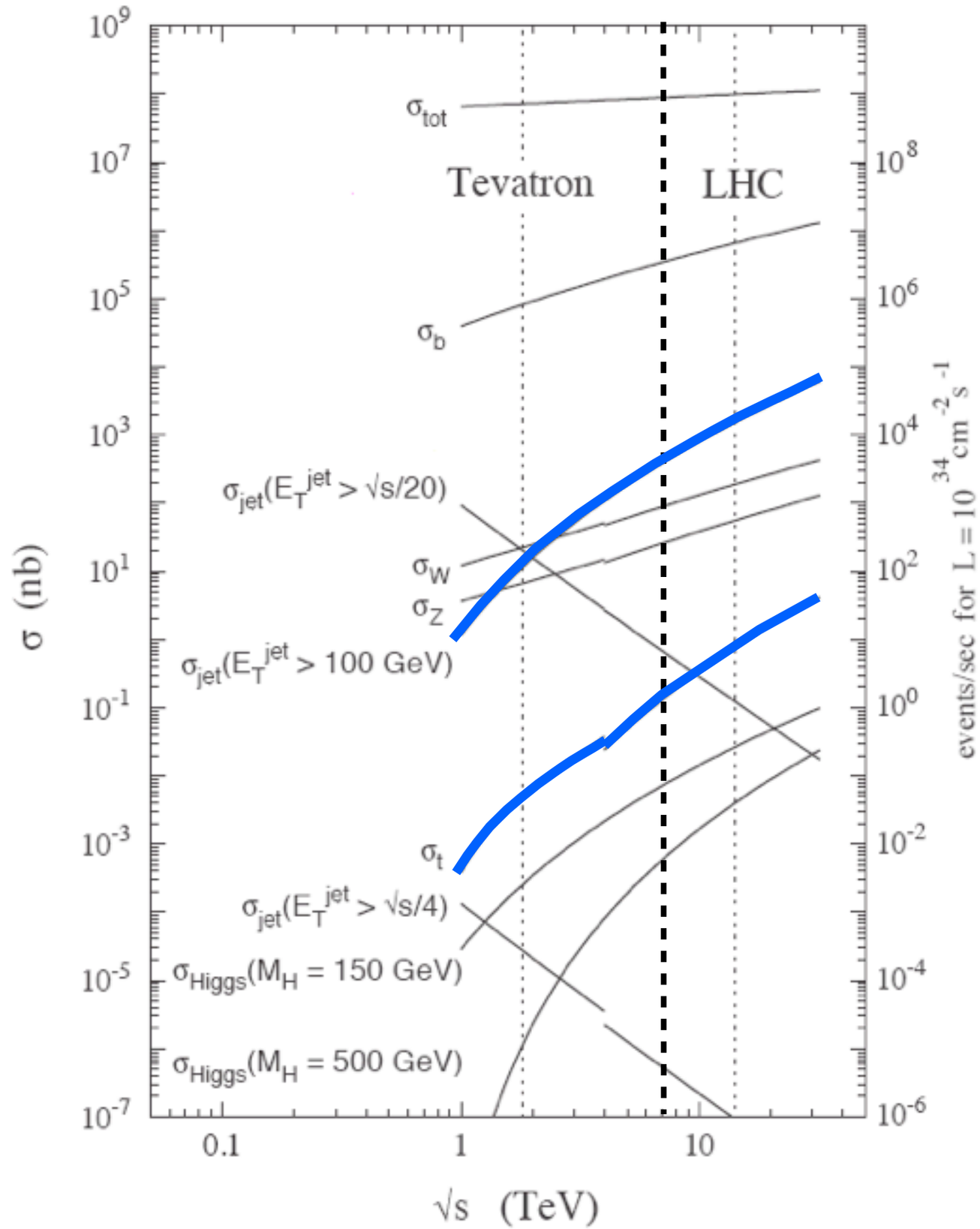
Salvatore Rappoccio (JHU)



26 July 2011

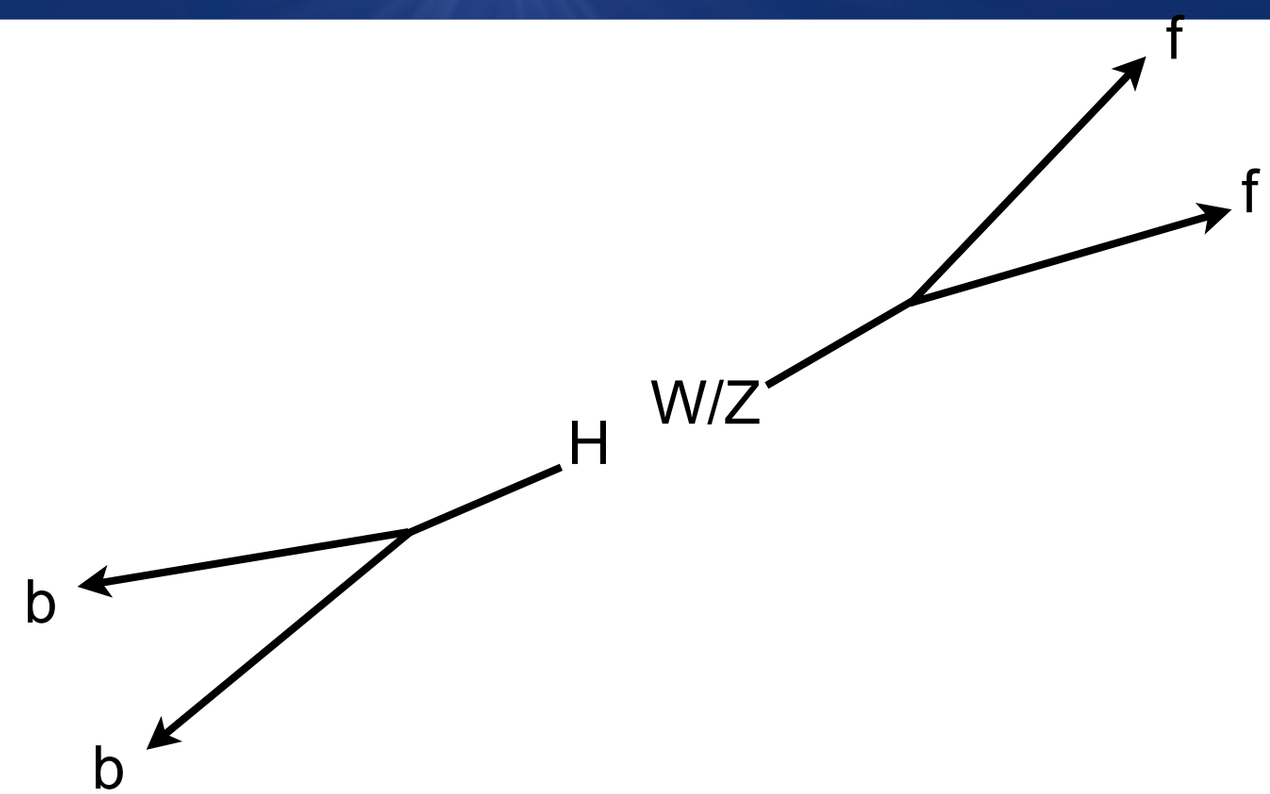
Jets, Jets, and Oh yeah.... More Jets

- Even more than at the Tevatron, the LHC is a very jetty place
  - Backgrounds contain jets
  - Signals contain jets
  - Pileup contains jets
- Understanding jets is a critical piece for all analyses



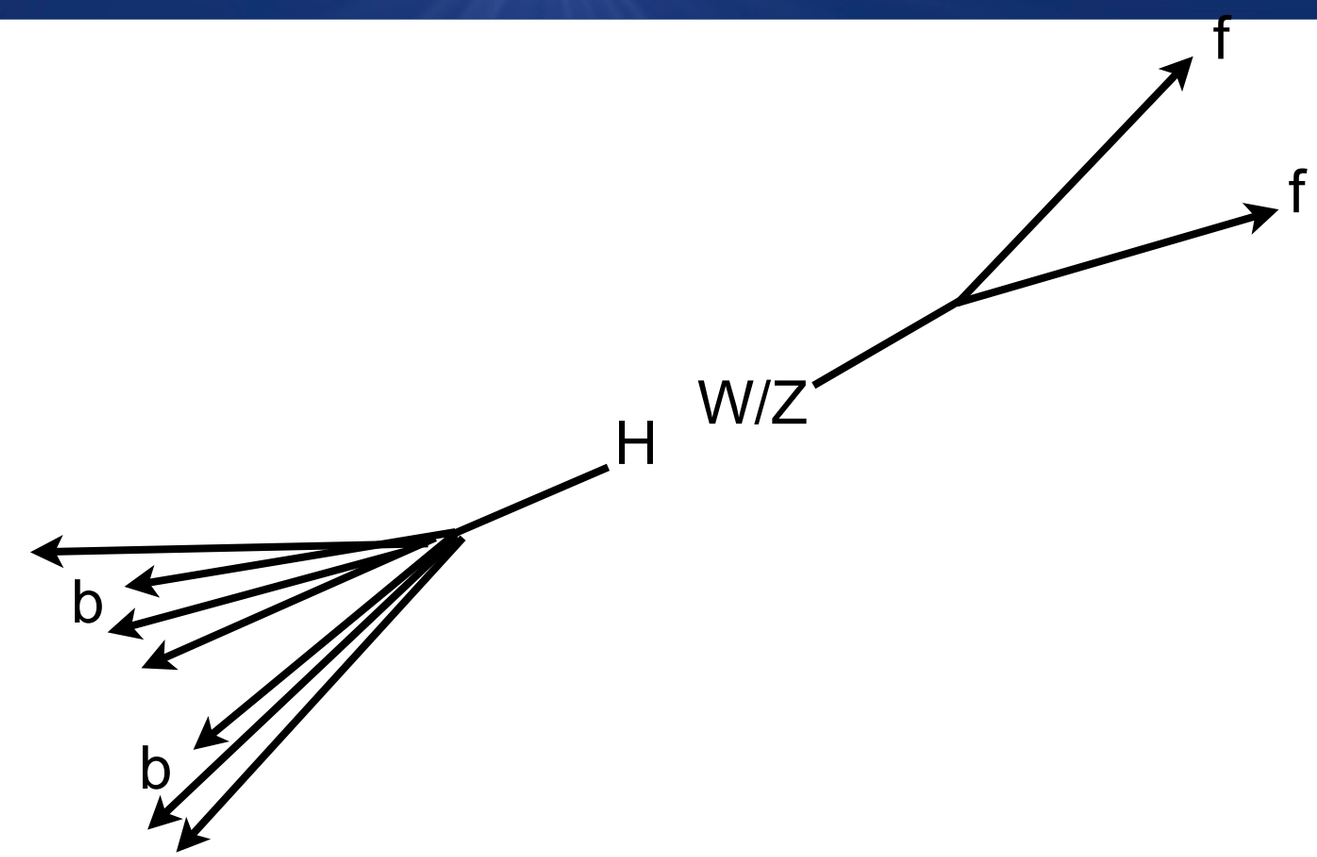
Particle Jets

- Nice overview: Ellis et al (arXiv:hep-ph/0412013v1)
- Simple picture of 1->2 decay is complicated in QCD!
- Spray of collimated particles
- A lot of complications!
  - Underlying event
  - Pileup
  - Initial + final state radiation



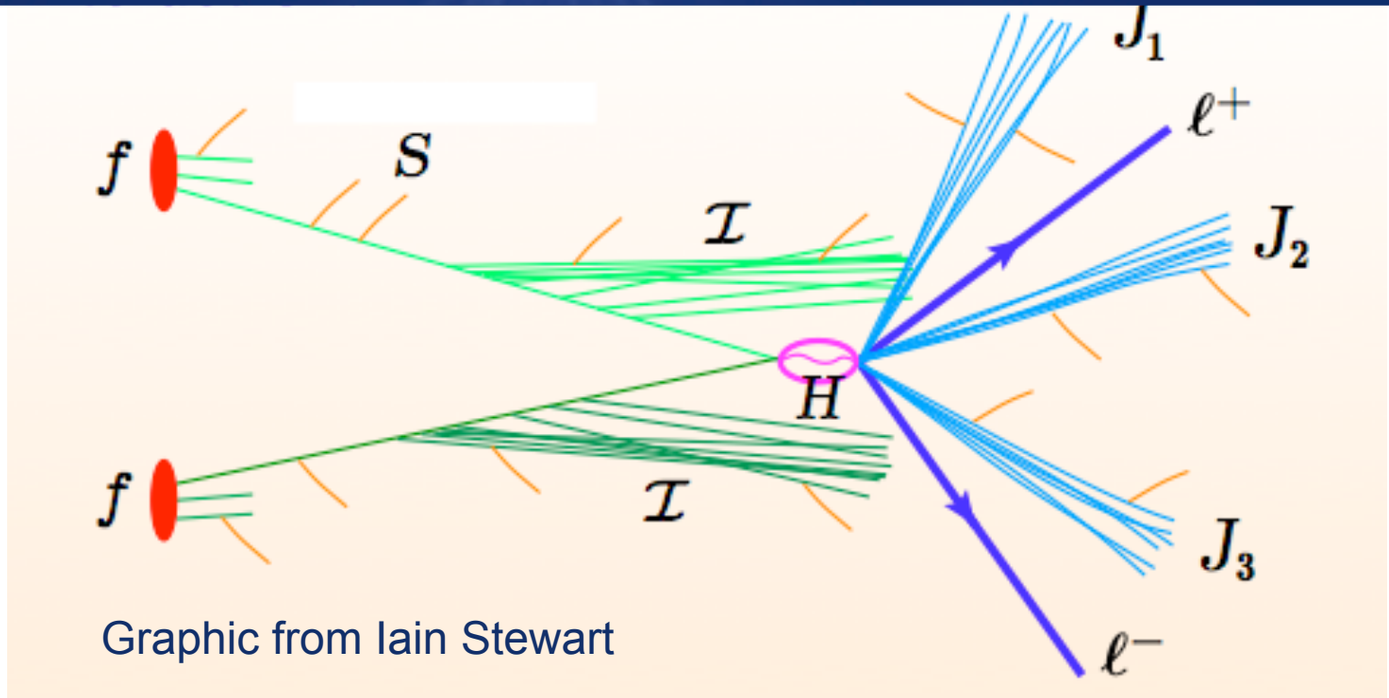
Particle Jets

- Nice overview: Ellis et al (arXiv:hep-ph/0412013v1)
- Simple picture of 1->2 decay is complicated in QCD!
- Spray of collimated particles
- A lot of complications!
  - Underlying event
  - Pileup
  - Initial + final state radiation



Particle Jets

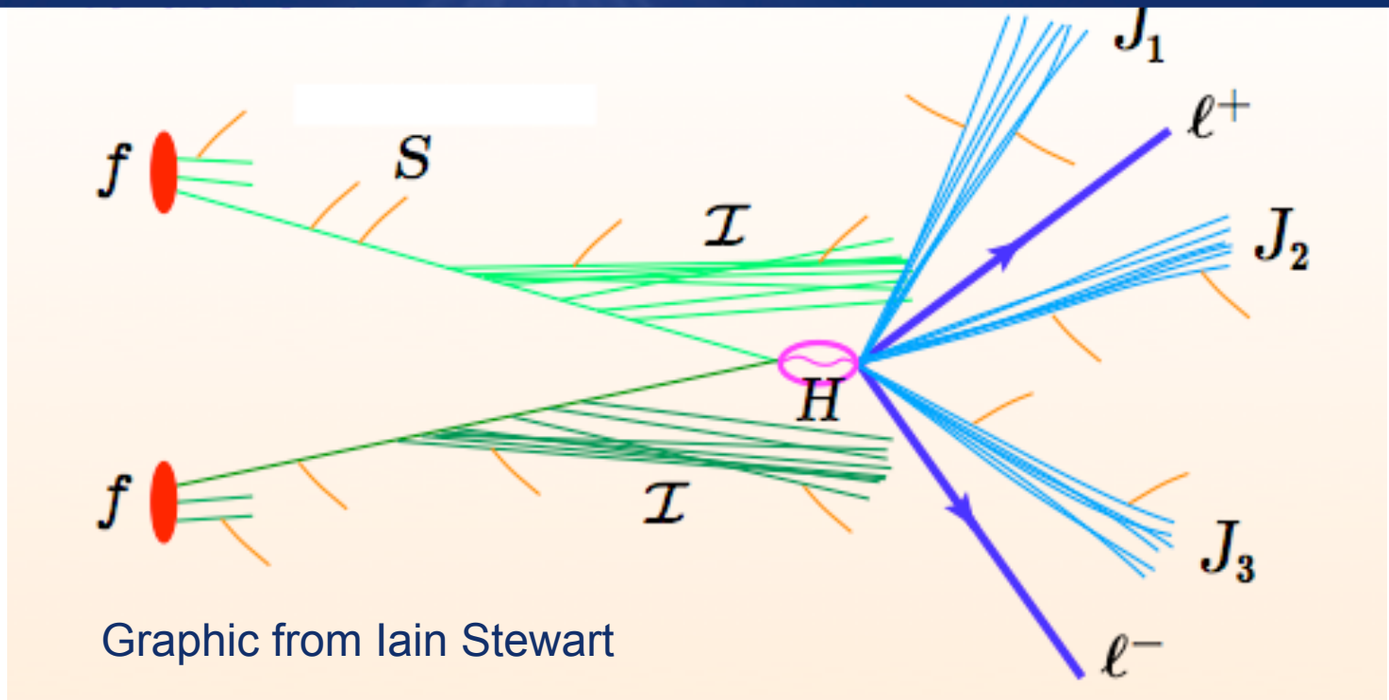
- Nice overview: Ellis et al (arXiv:hep-ph/0412013v1)
- Simple picture of 1->2 decay is complicated in QCD!
- Spray of collimated particles
- A lot of complications!
  - Underlying event
  - Pileup
  - Initial + final state radiation



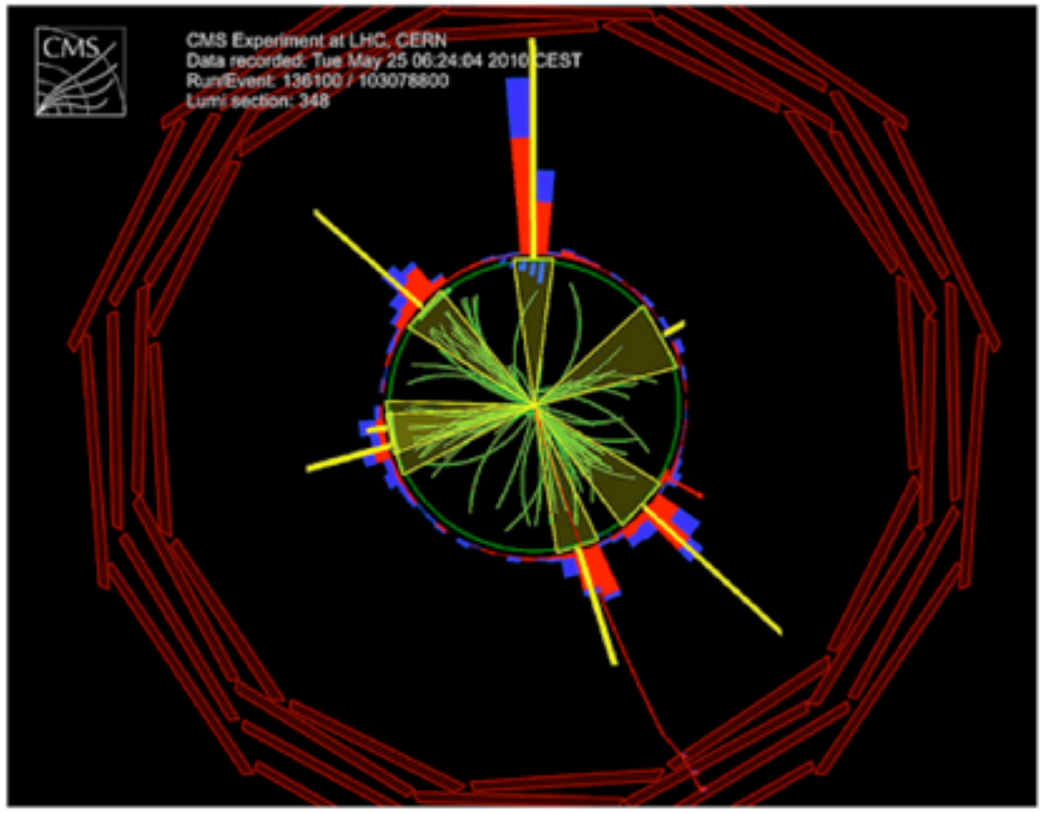
Graphic from Iain Stewart

Particle Jets

- Nice overview: Ellis et al (arXiv:hep-ph/0412013v1)
- Simple picture of 1->2 decay is complicated in QCD!
- Spray of collimated particles
- A lot of complications!
  - Underlying event
  - Pileup
  - Initial + final state radiation



Graphic from Iain Stewart



Caveat

- This talk isn't a history, but a snapshot

- Will focus on recent results with application to data
  - There are many other substructure tools available from the theoretical community
    - Will only discuss those that are already used by experiments
  - There are many studies from the experiments using MC
    - Will only discuss those that have data comparisons

- Investigate the principle for most of these tools in detail:
  - Sequential algorithms!

Sequential Clustering Algorithms

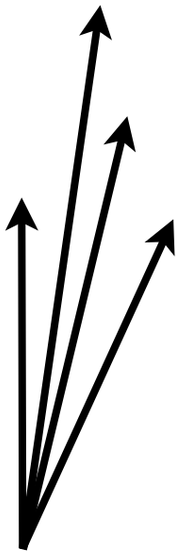
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential Clustering Algorithms

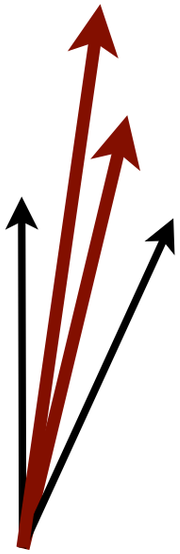
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential Clustering Algorithms

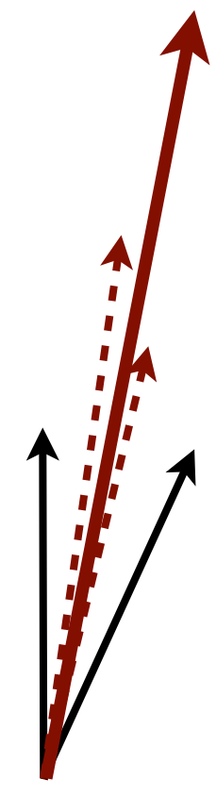
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential Clustering Algorithms

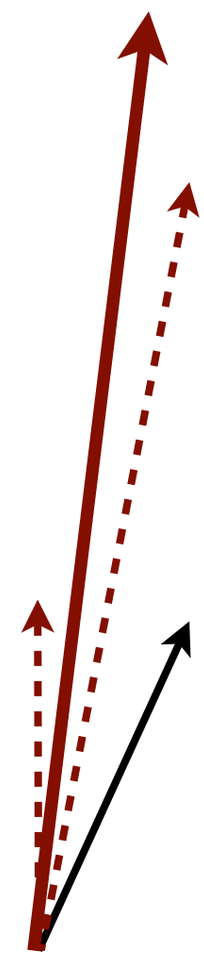
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential Clustering Algorithms

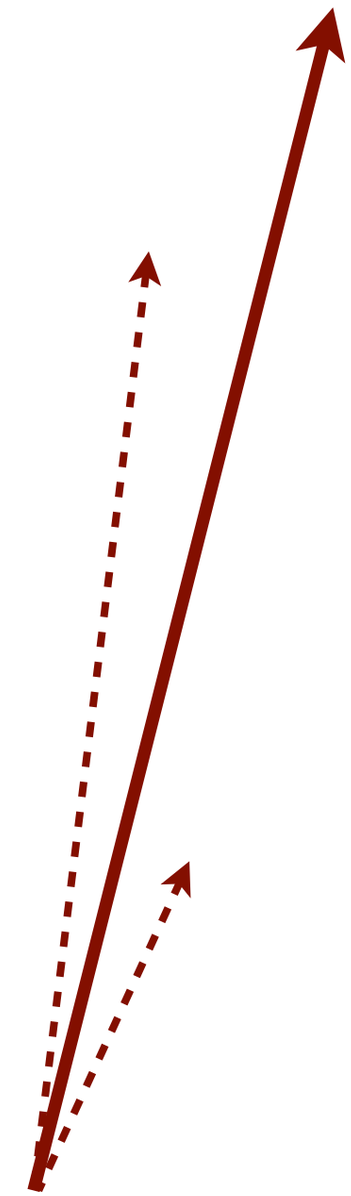
- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted



Sequential Clustering Algorithms

- Pairwise examination of input 4-vectors
- Calculate d_{ij}

$$d_{ij} = min(k_{ti}^n, k_{tj}^n)\Delta R_{ij}^2 / R^2$$

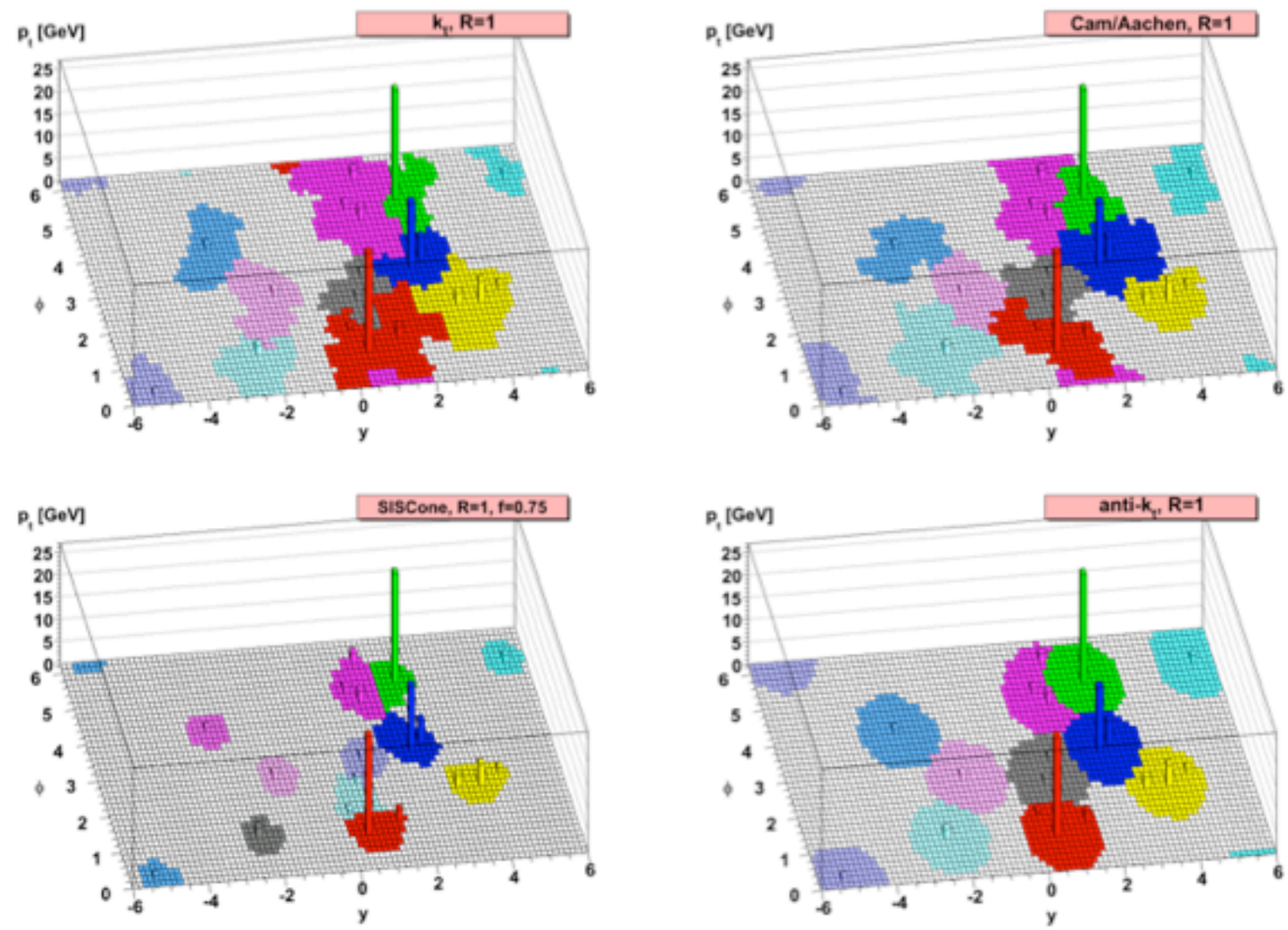
- Also find the "beam distance"

$$d_{iB} = k_{T,i}^n$$

- Find min of all d_{ij} and d_{iB}
  - If min is a dij, merge and iterate
  - If min is a d_{iB}, classify as a final jet
- Continue until list is exhausted

Sequential Clustering Algorithms

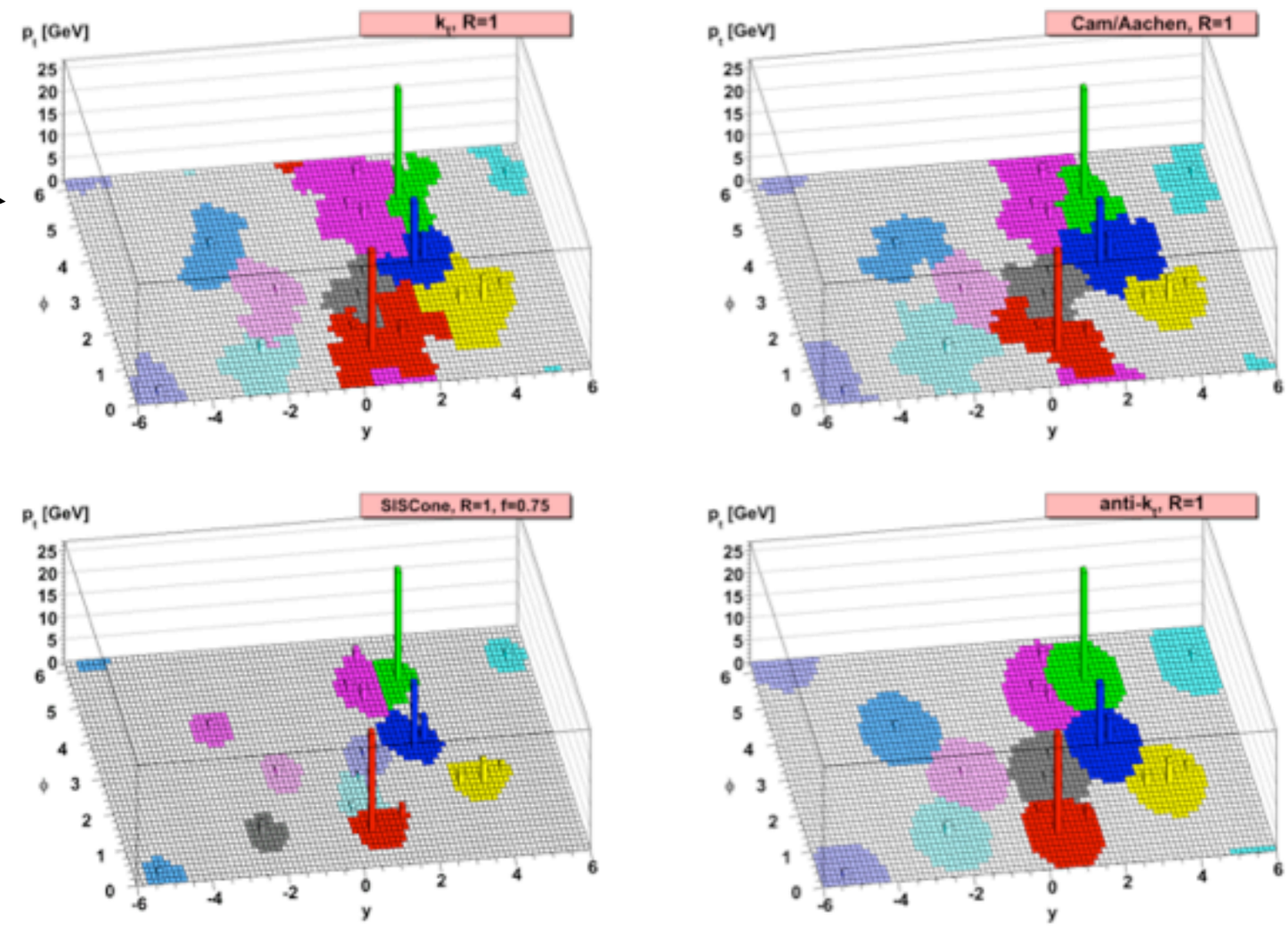
- Different types
  - N = 2: “kT”
    - “Irregular” jets, but good for low pt
  - N = 0 : “Cambridge-Aachen” (CA)
    - Also irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Sequential Clustering Algorithms

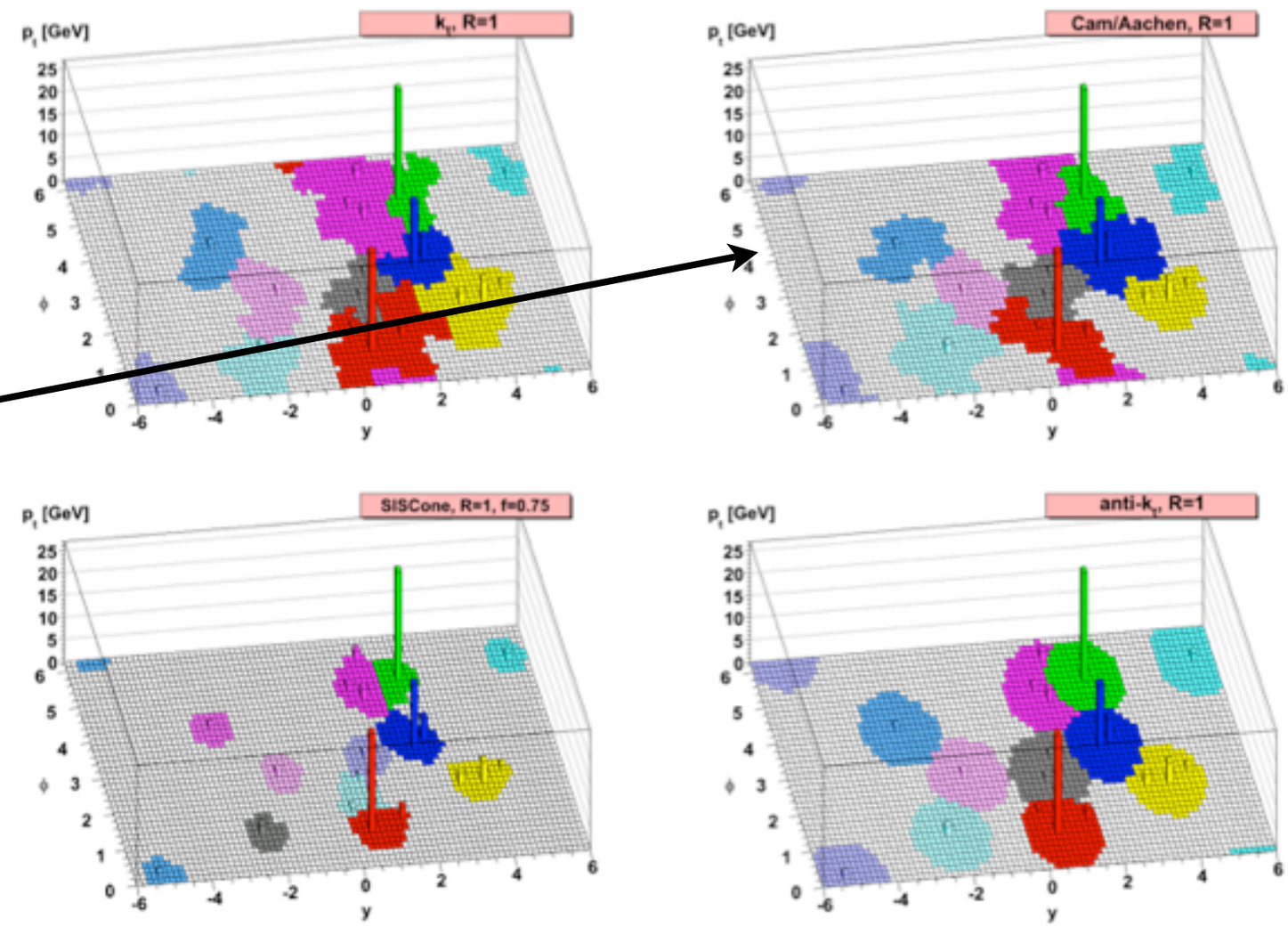
- Different types
  - N = 2: “kT”
    - “Irregular” jets, but good for low pt
  - N = 0 : “Cambridge-Aachen” (CA)
    - Also irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Sequential Clustering Algorithms

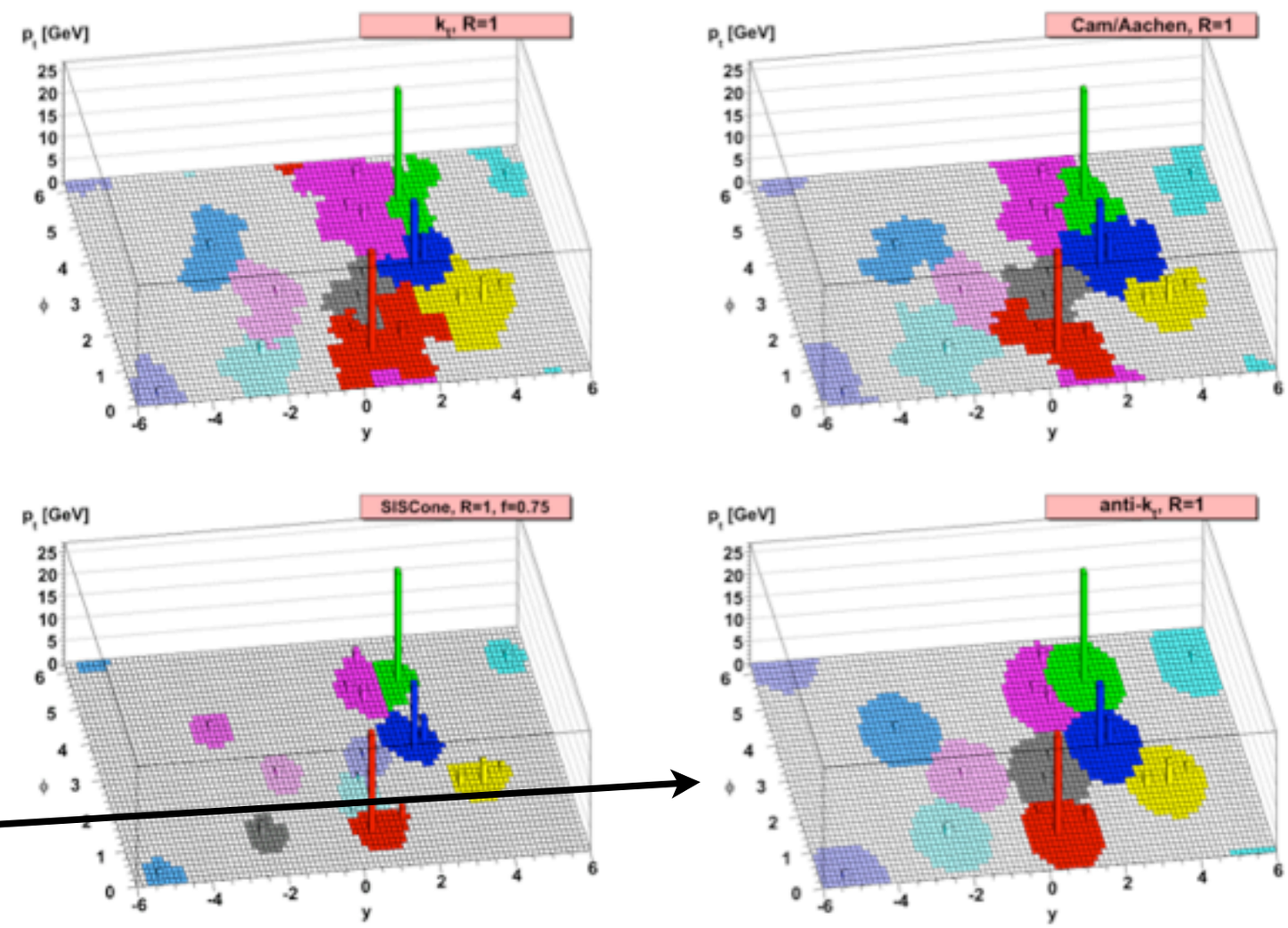
- Different types
  - N = 2: “kT”
    - “Irregular” jets, but good for low pt
  - N = 0 : “Cambridge-Aachen” (CA)
    - Also irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Sequential Clustering Algorithms

- Different types
  - N = 2: “kT”
    - “Irregular” jets, but good for low pt
  - N = 0 : “Cambridge-Aachen” (CA)
    - Also irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Jet Mass and Substructure

Jet Mass

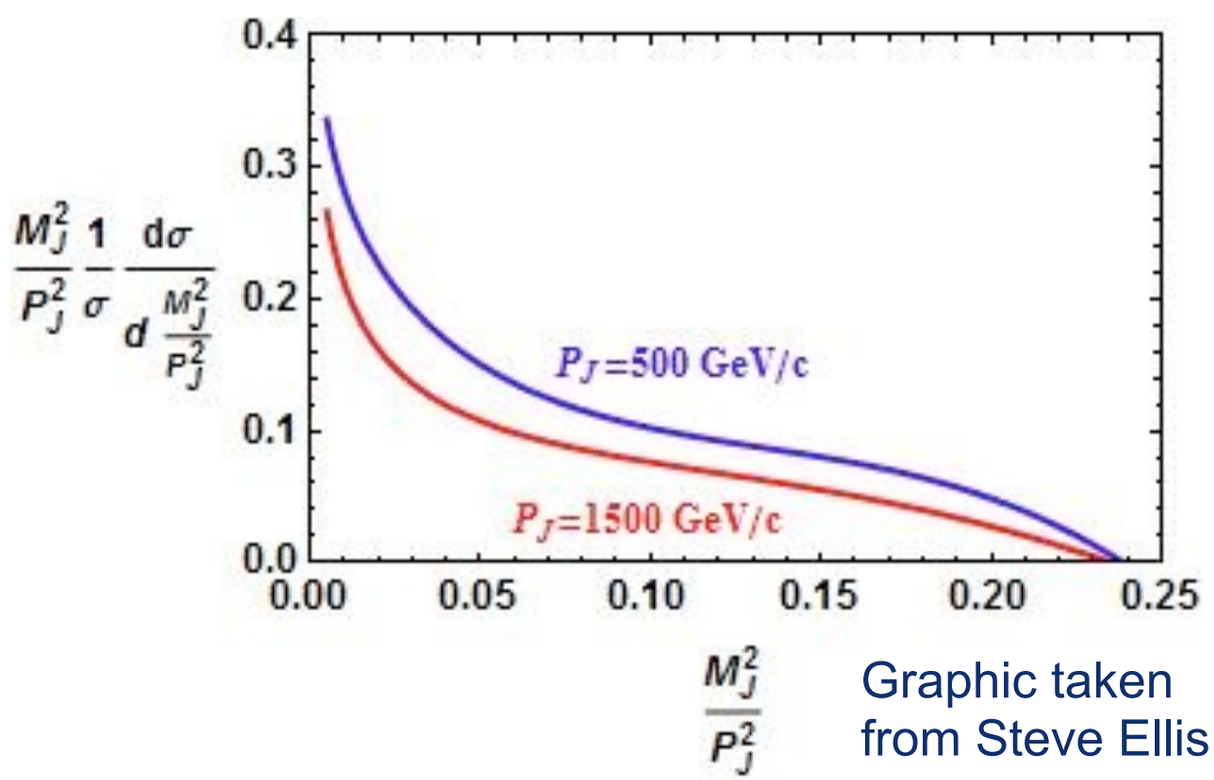
- What's the typical mass scale of QCD?
  - See e.g. Ellis et al (arXiv:0712.2447v1)

Finite-size effects from cutoff

Scales ~linearly with momentum

Log. divergence at low mass

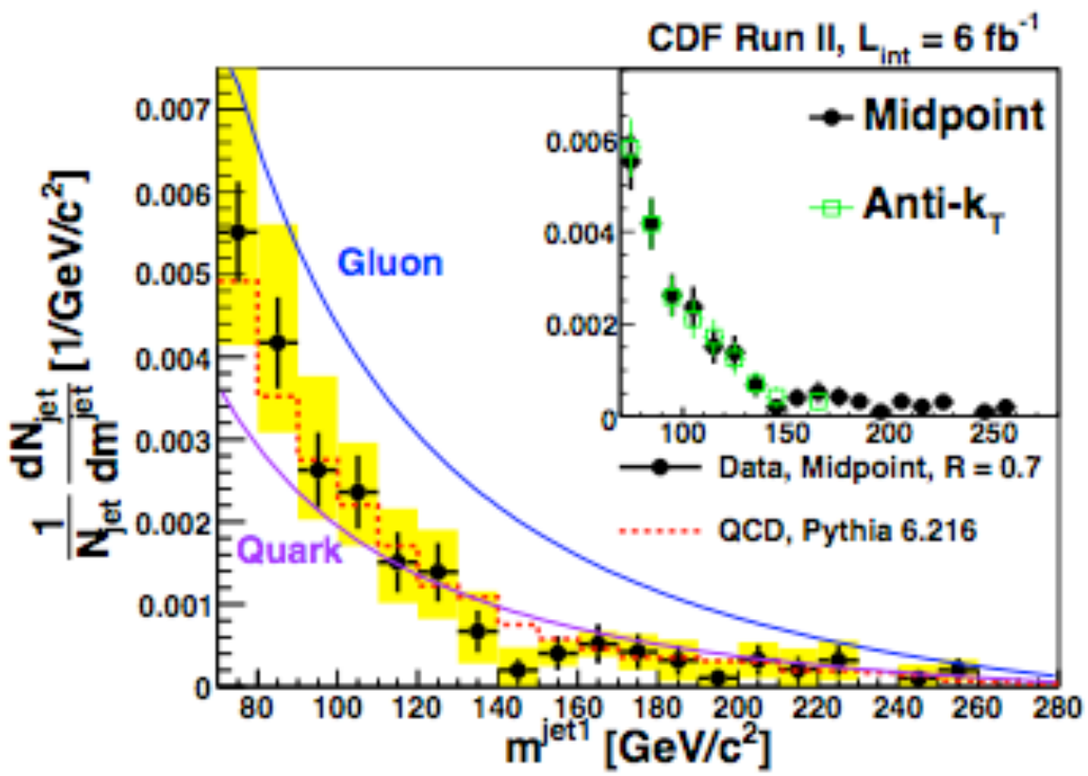
$$\langle M_J^2 \rangle_{NLO} \simeq \overline{C} \left(\frac{p_J}{\sqrt{s}} \right) \alpha_s \left(\frac{p_J}{2} \right) p_J^2 R^2,$$



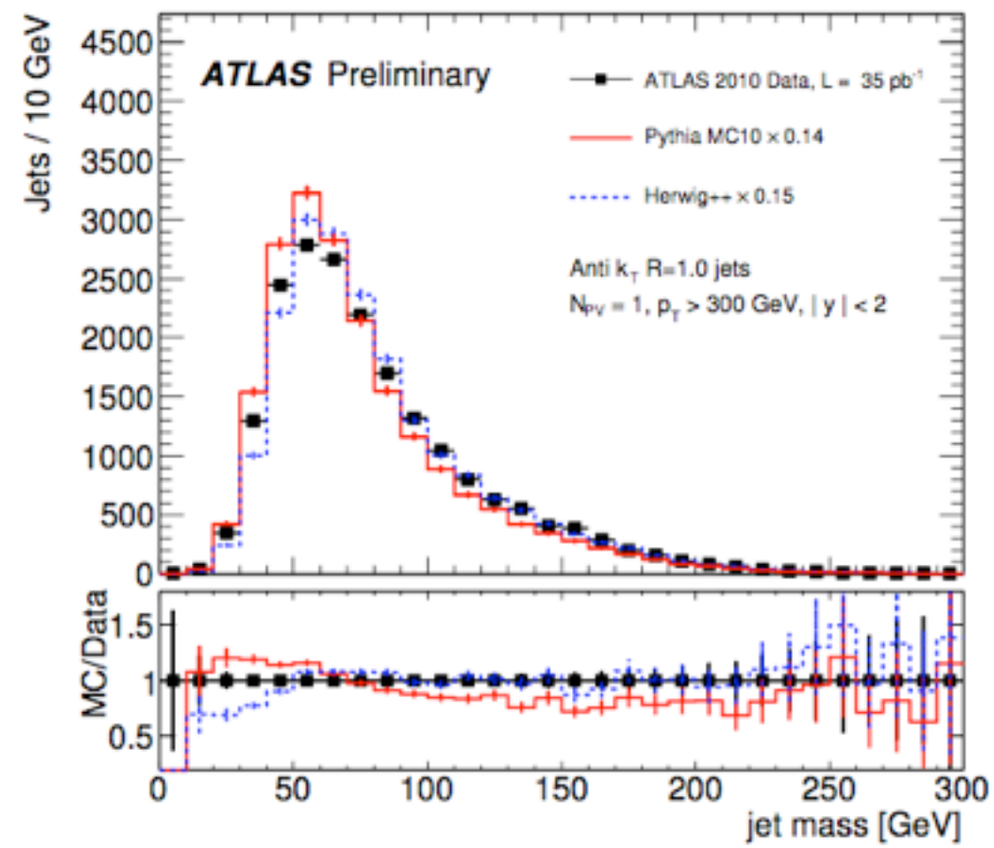
Graphic taken from Steve Ellis

Jet Mass

arXiv:1106.5952v1 [hep-ex]



ATLAS-CONF-2011-073



- So far excellent agreement between data and predictions
- Dependence seen versus parton shower model, pt vs angular ordering, and even UE tune

Jet Tools

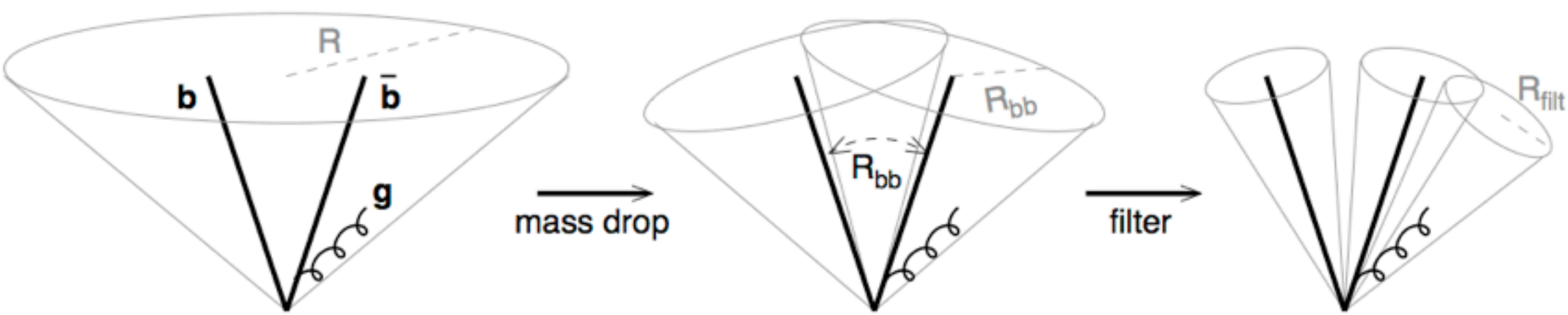
- "BDRS" Subjet/filter technique
  - Targeted tagger for identifying boosted Higgs (2-body decay)
  - Butterworth, Davison, Rubin, Salam: Phys.Rev.Lett.100:242001,2008
- JHU "top tagger"
  - Targeted tagger for identifying boosted top quarks (3-body decay)
  - Kaplan, Reherman, Schwartz,Tweedie: Phys.Rev.Lett. 101:142001,2008, arXiv:0806.0848
- University of Washington "jet pruning"
  - General-purpose tool to "clean up" jets
  - Ellis, Vermilion, Walsh: Phys. Rev. D80 (2009) 051501, arXiv: 0903.5081
- Jet "Superstructure"
  - Attempts to separate color singlets from color octets
  - Gallicchio and Schwartz, Phys.Rev.Lett.105:022001,2010, arXiv: 1001.5027v3

"BDRS" Subjet/Filter Details

- Butterworth et al (arXiv:0802.2470v2 [hep-ph])
- Use CA PFJets
- "Undo" last clustering step
- Examine "mass drop" and "subjet symmetry" (make subjet)

$$\mu = \frac{m_{j1}}{m_j} \qquad y = \frac{\min(p_{T1}, p_{T2}) \cdot (\Delta R_{12})^2}{m_{jet}^2}$$

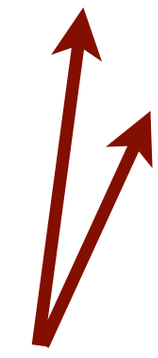
- If significant mass drop and symmetric subjets, reduce the jet scale and recluster (filter subjet)
  - This is DIFFERENT from QCD



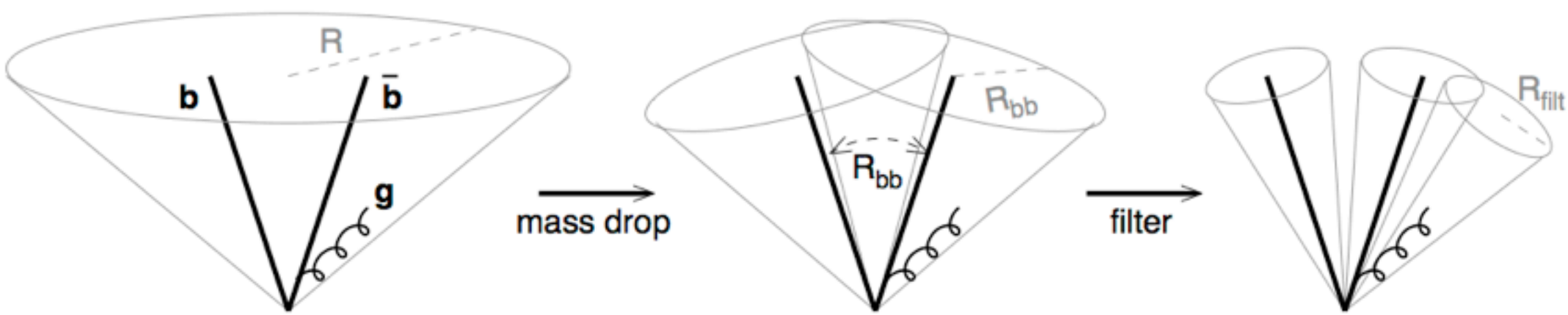
“BDRS” Subjet/Filter Details

- Butterworth et al (arXiv:0802.2470v2 [hep-ph])
- Use CA PFJets
- “Undo” last clustering step
- Examine “mass drop” and “subjet symmetry” (make subjet)

$$\mu = \frac{m_{j1}}{m_j} \qquad y = \frac{\min(p_{T1}, p_{T2}) \cdot (\Delta R_{12})^2}{m_{jet}^2}$$

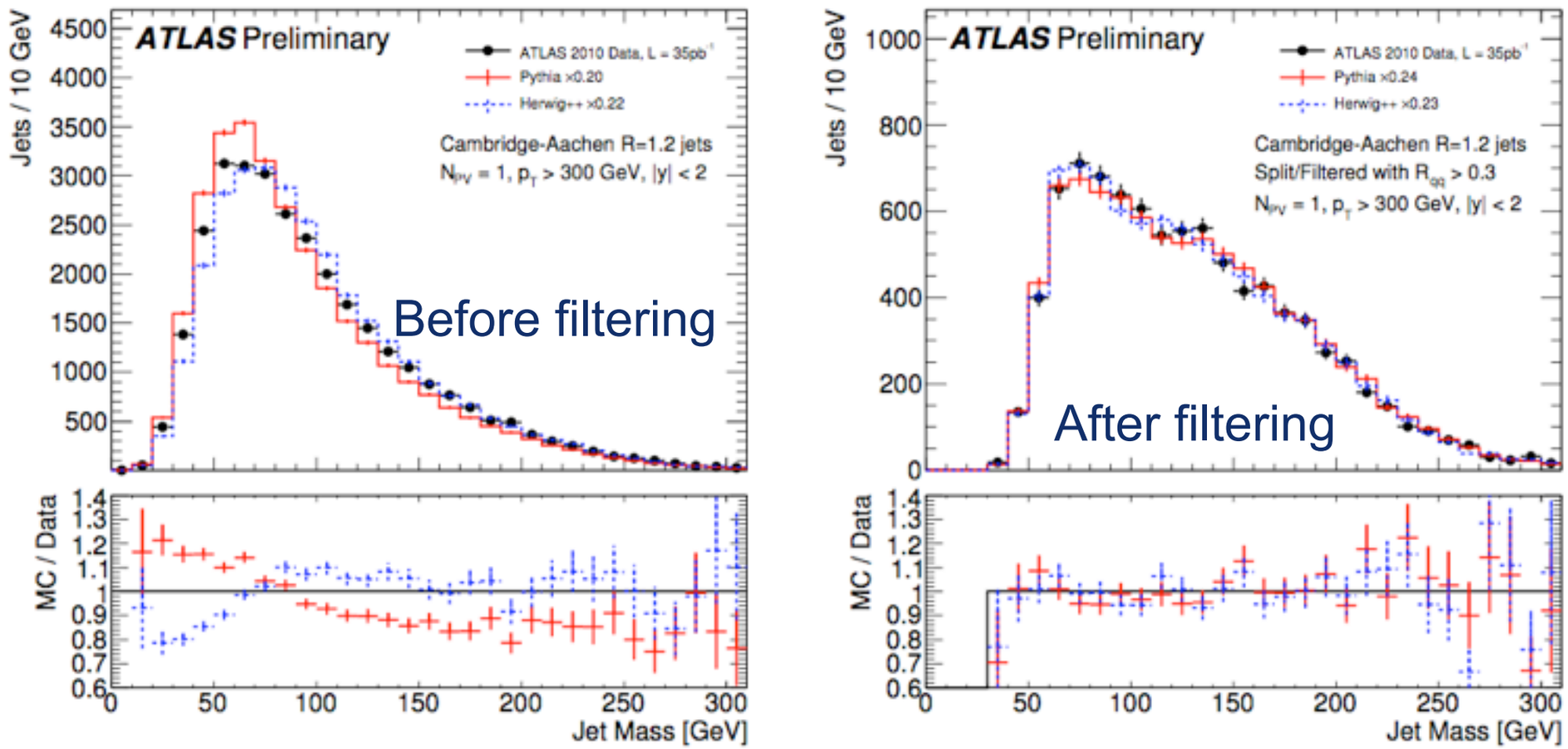


- If significant mass drop and symmetric subjets, reduce the jet scale and recluster (filter subjet)
  - This is DIFFERENT from QCD



"BDRS" Subjet/Filter Details

ATLAS-CONF-2011-073



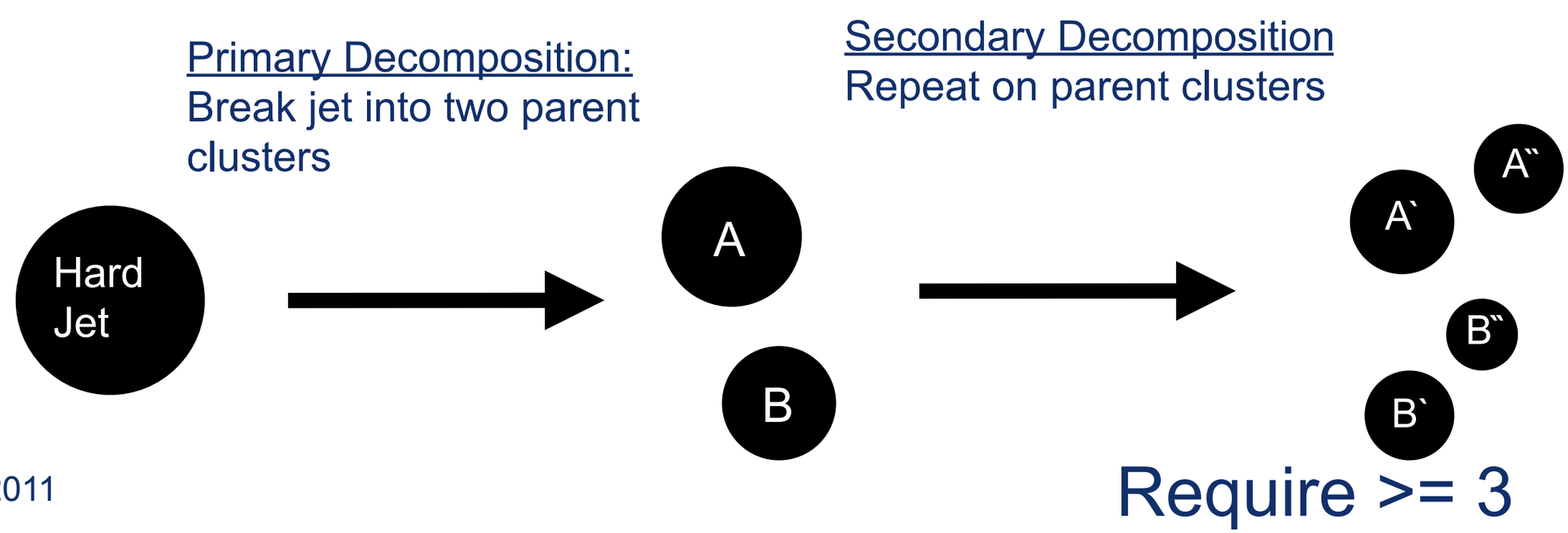
- Good agreement between data and MC
- Largest variations between PS and UE tunes

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×pThard jet
  - Adjacency criterion: ΔR(A, B) > 0.4 - 0.0004×pT
- Iterative process - throw out objects that fail momentum fraction cut and try to decluster again
- Then use :
  - Jet Mass ~ Top mass
  - Minimum mass pairing of subjets ~ W mass

Removes soft clusters

Removes wide angle clusters



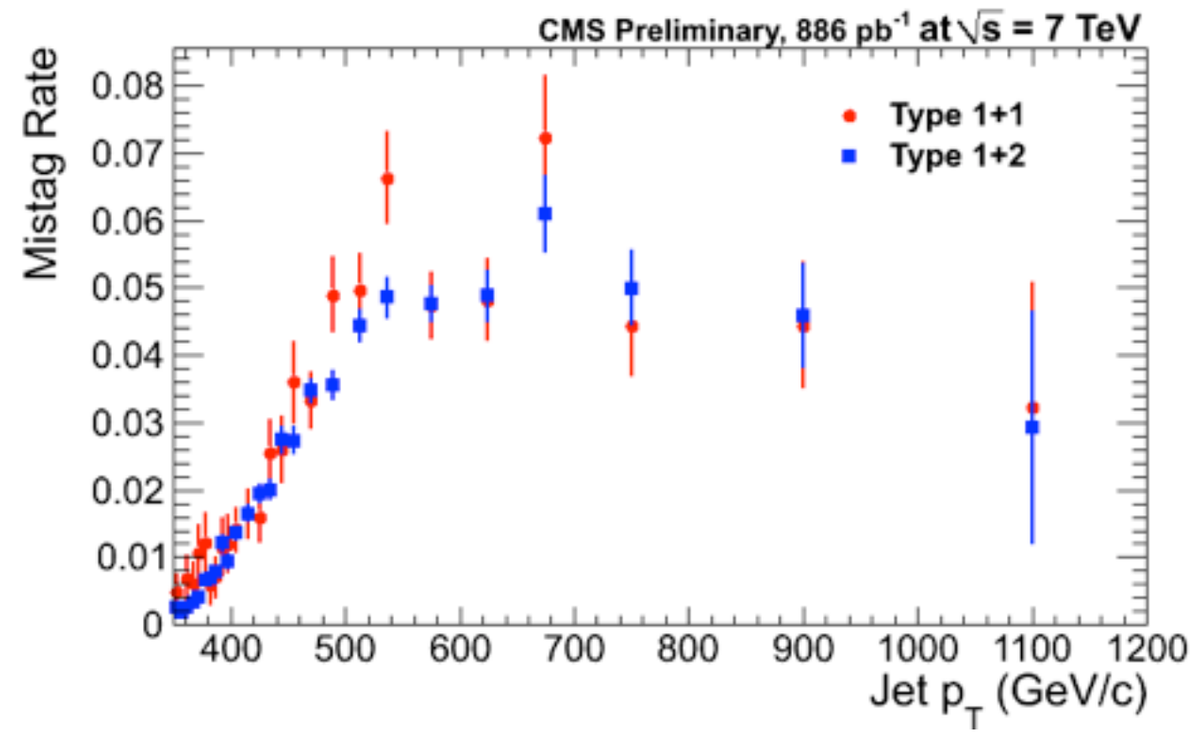
Top Tagging Mistag Rate From Data

CMS-PAS-EXO-11-006

- 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



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

* Type 1 + 1 = Dijet events, with mass requirement
* Type 1 + 2 = Trijet events

Jet Pruning Details

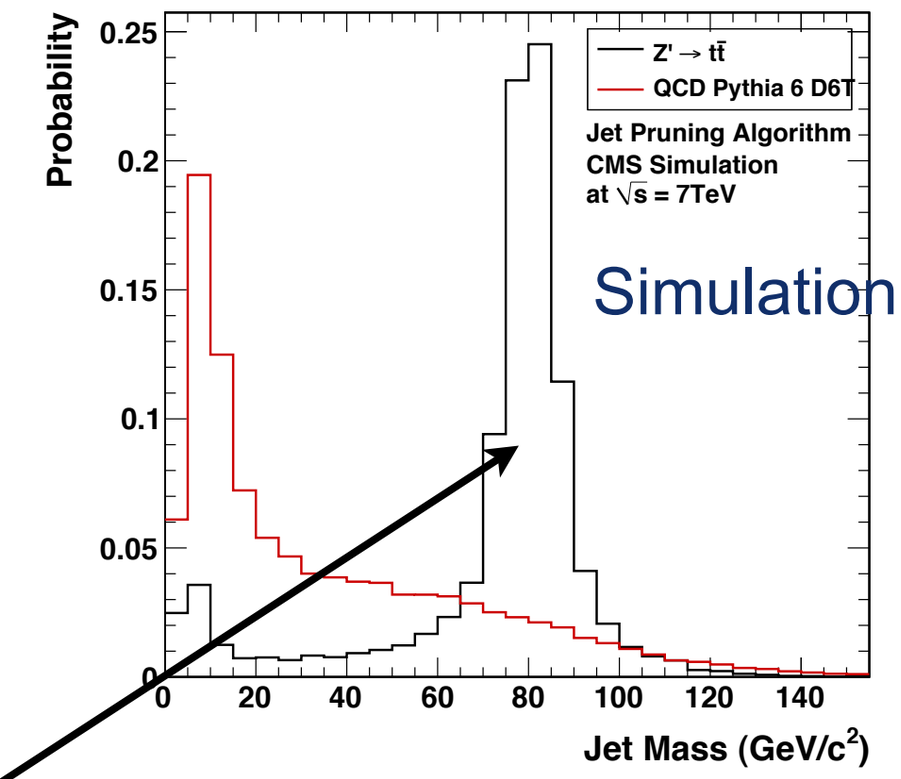
- Ellis et al. (arXiv:0903.5081)
- Improves mass resolution by removing soft, large angle particles from the jet
- Recluster each jet, requiring that each recombination satisfy the following:

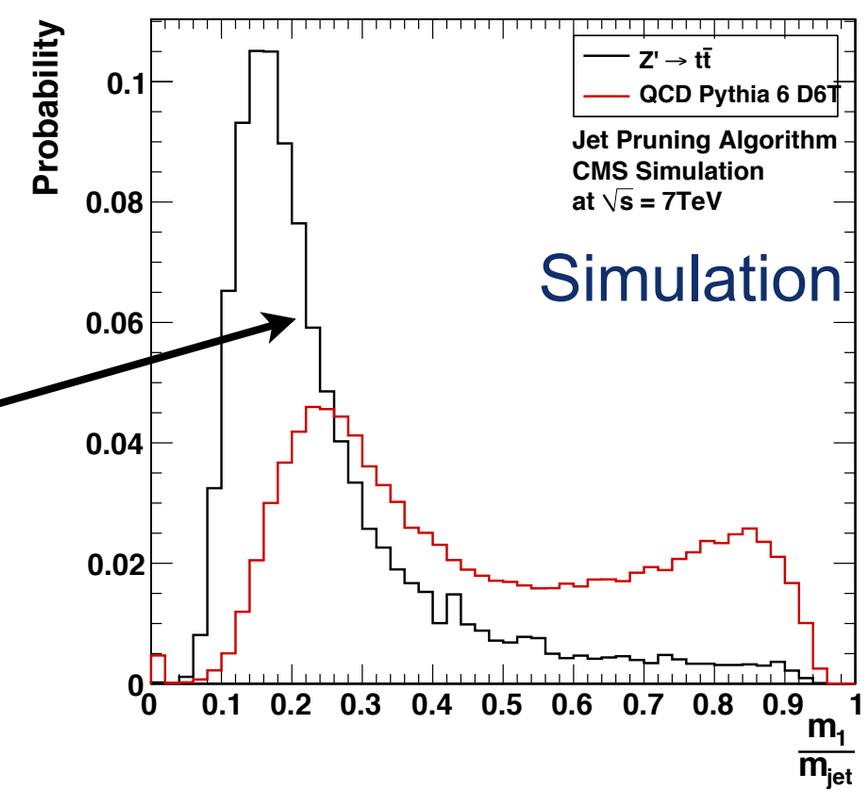
$$\frac{\min(p_{\mathrm{T}1}, p_{\mathrm{T}2})}{p_{\mathrm{T}p}} > 0.1$$

$$\Delta R_{12} < 0.5 \times \frac{m_{\mathrm{jet}}}{p_{\mathrm{T}}}$$

- For W tagging, require:
  - Jet mass in 60-100 GeV/c2
  - Mass drop (mu) < 0.4

$$\mu = \frac{m_{j1}}{m_j}$$

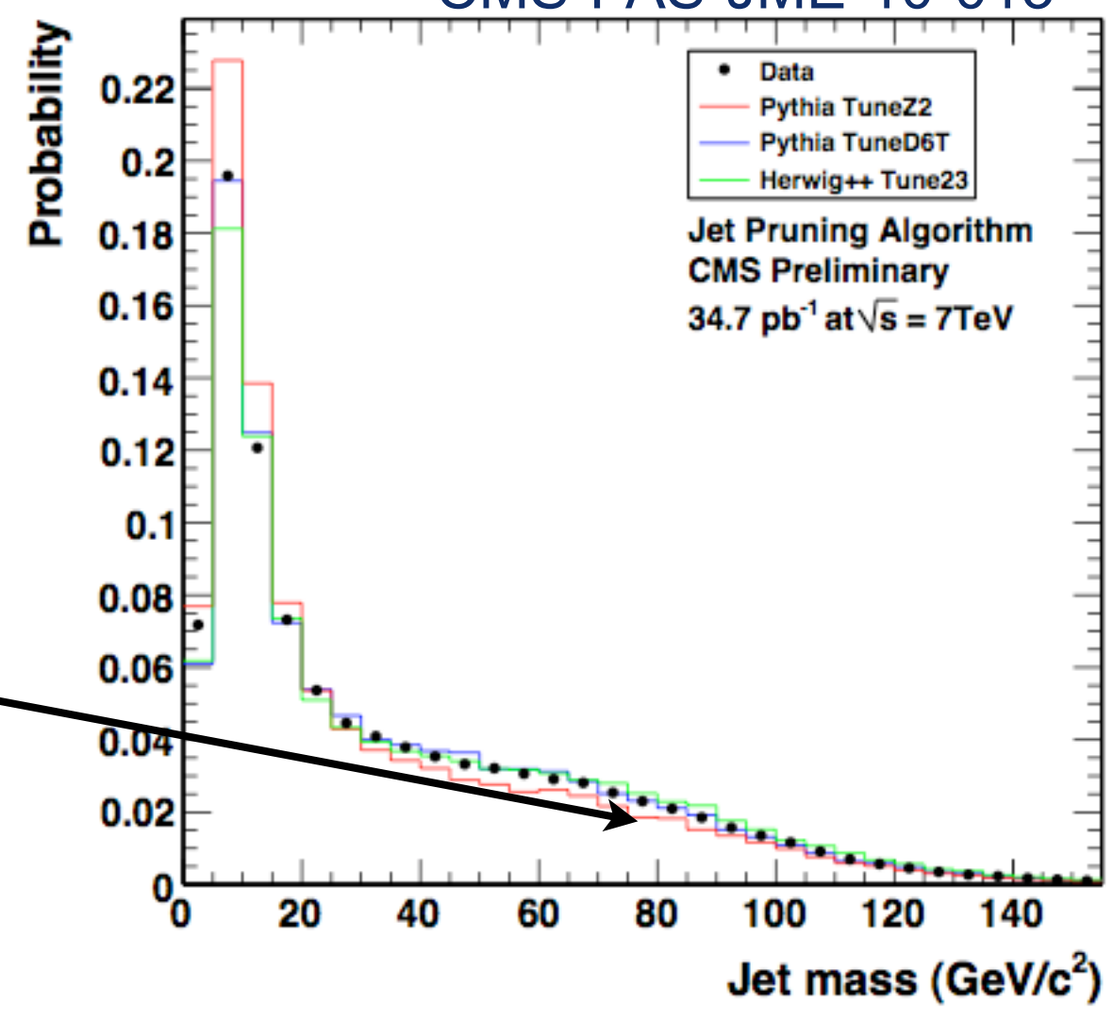




Pruned Jet Mass

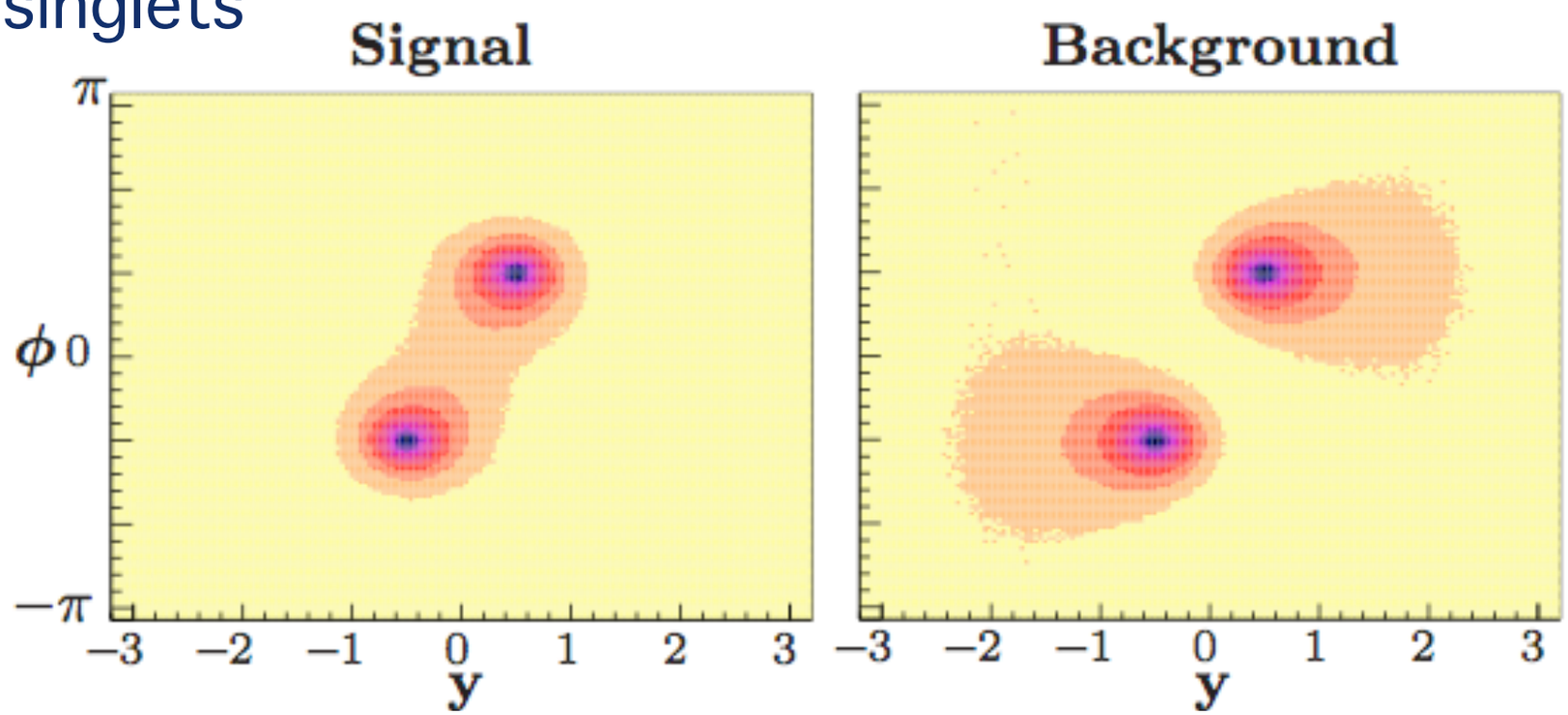
- Very good agreement between MC and data jet mass!
- Strong dependence on PS model and tune, but all very good in “signal region” for interesting particles (high jet mass)

CMS-PAS-JME-10-013



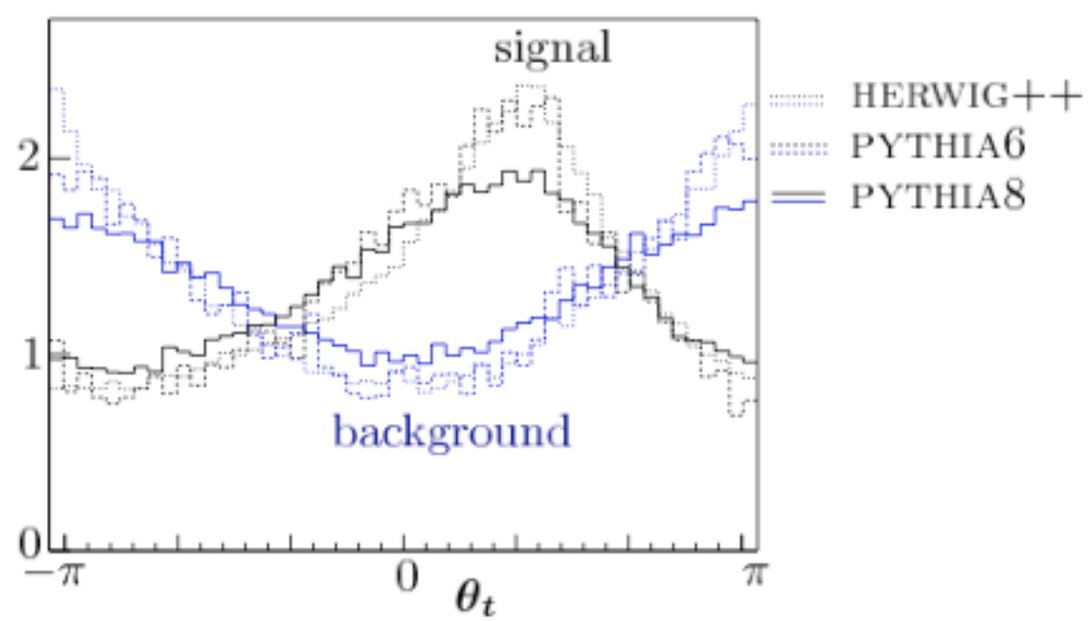
Jet Superstructure

- Gallicchio and Schwartz (arXiv:1001.5027v3)
- Take advantage of color connection between objects to distinguish color singlets



- Compute "pull":

$$\vec{t} = \sum_{i \in \text{jet}} \frac{p_T^i \, |r_i|}{p_T^{\text{jet}}} \vec{r}_i \, .$$

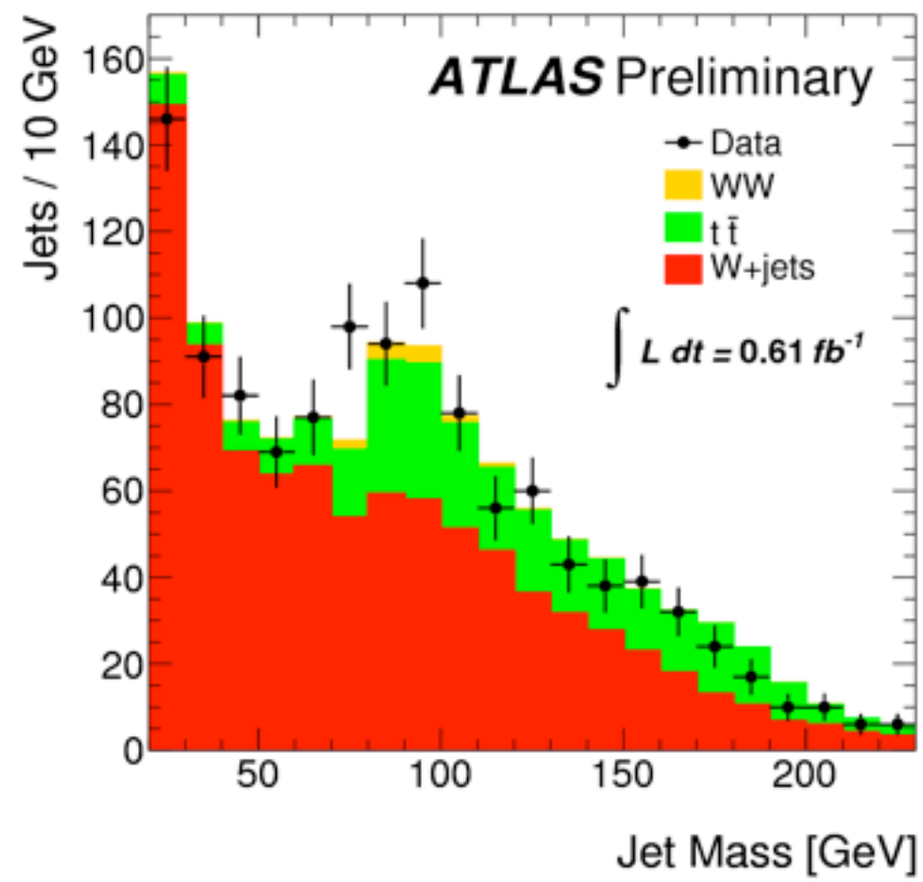


Characterization of Substructure Tools

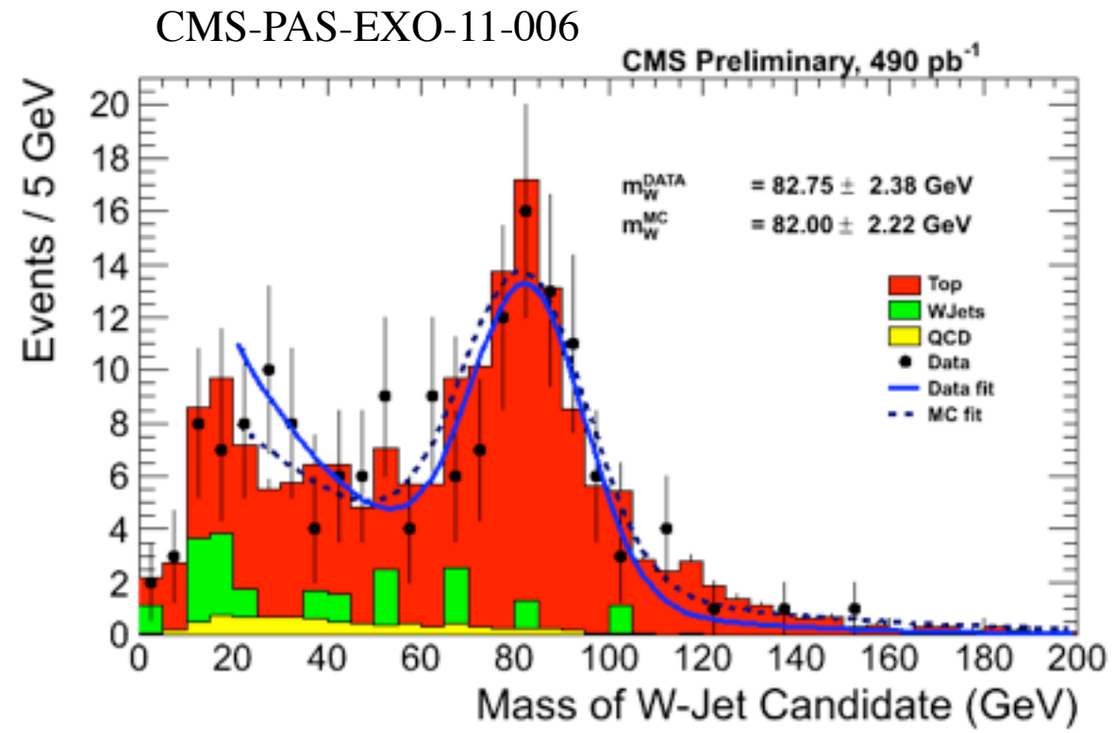
W Mass in Semileptonic Sample

- Standard candle : Find a W peak in jet masses!
  - Examine boosted semileptonic ttbar
  - Look at mass of W-jet candidates
  - Look at no b-tags (ATLAS) or with b-tags (CMS)
  - Both show a nice clear W peak from within jets
  - Also observe good data to MC agreement

ATLAS-CONF-2011-103



CMS-PAS-EXO-11-006



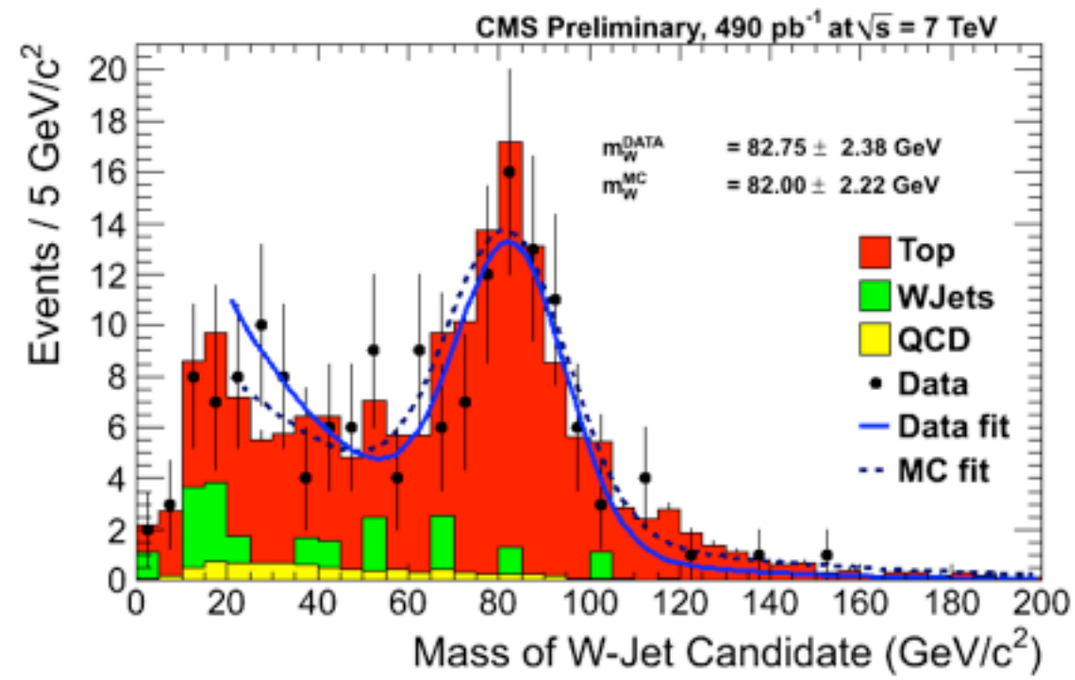
26 July 2011

W Tagging With Jet Pruning

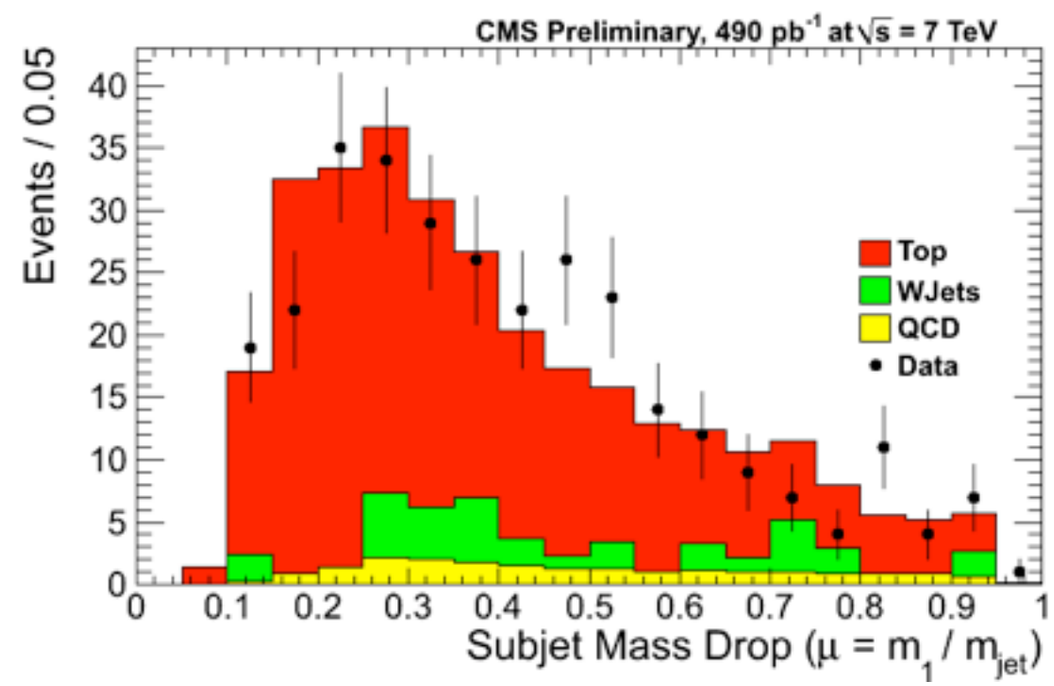
- Can compute ratio of efficiencies in data and MC
  - Completely statistically dominated, will improve with more data!

- Combined selection:
  - 60 < Mjet < 100
  - Mu < 0.4

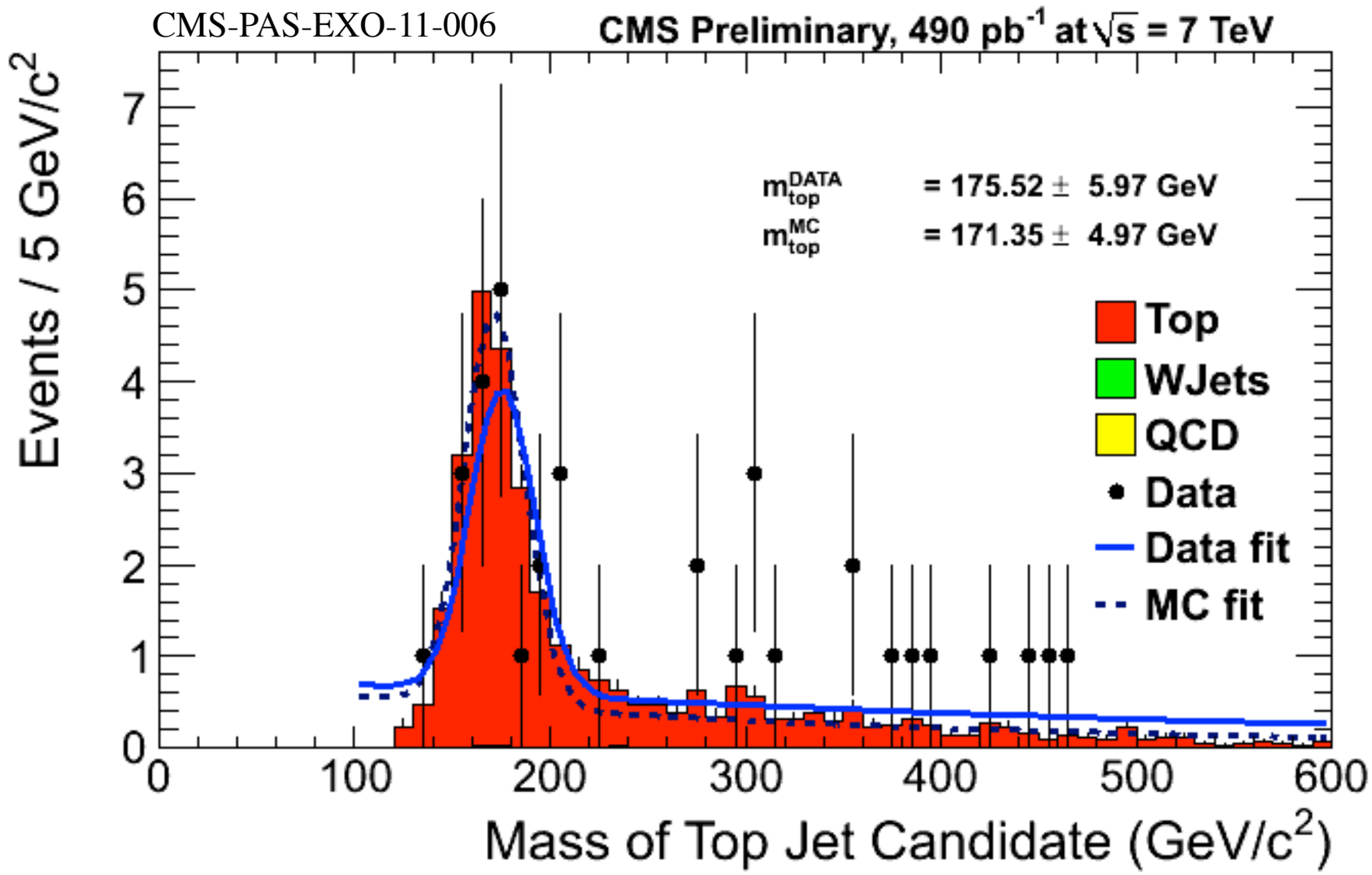
- Combined efficiencies:
  - Data: 28%
  - MC : 30%
- Evaluate "data-to-MC scale factor":
  - SF = 0.93 +- 0.13



CMS-PAS-EXO-11-006



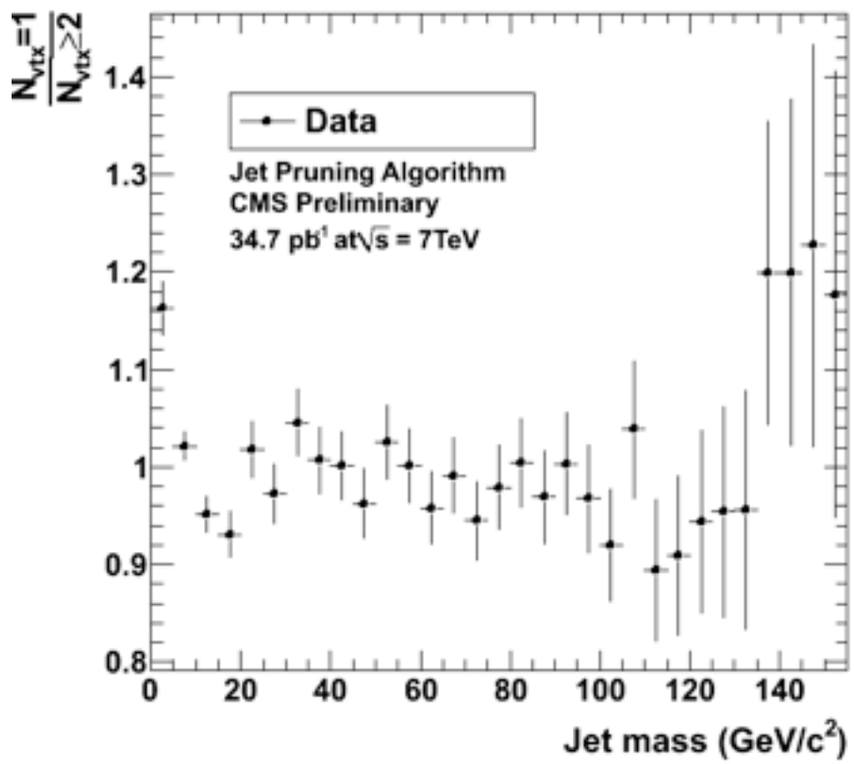
Top Mass in Semileptonic Sample



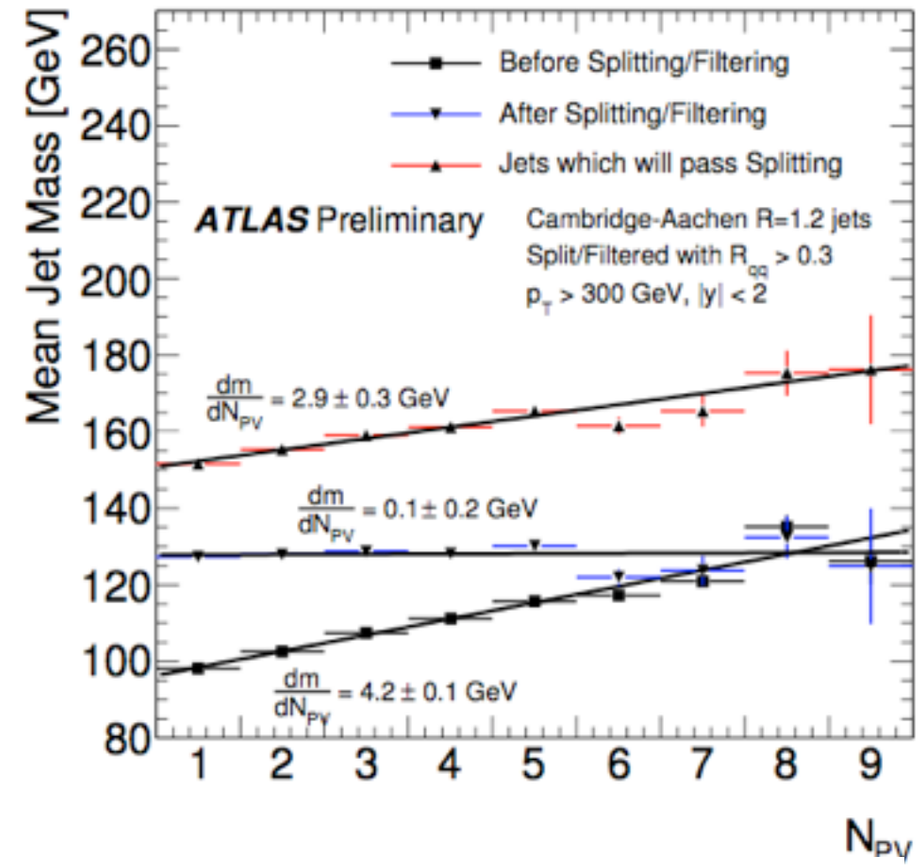
- Take events in the W window (60-100)
- Combine with closest jet, form top mass
- Very clear top signal, can even be used to measure top mass!

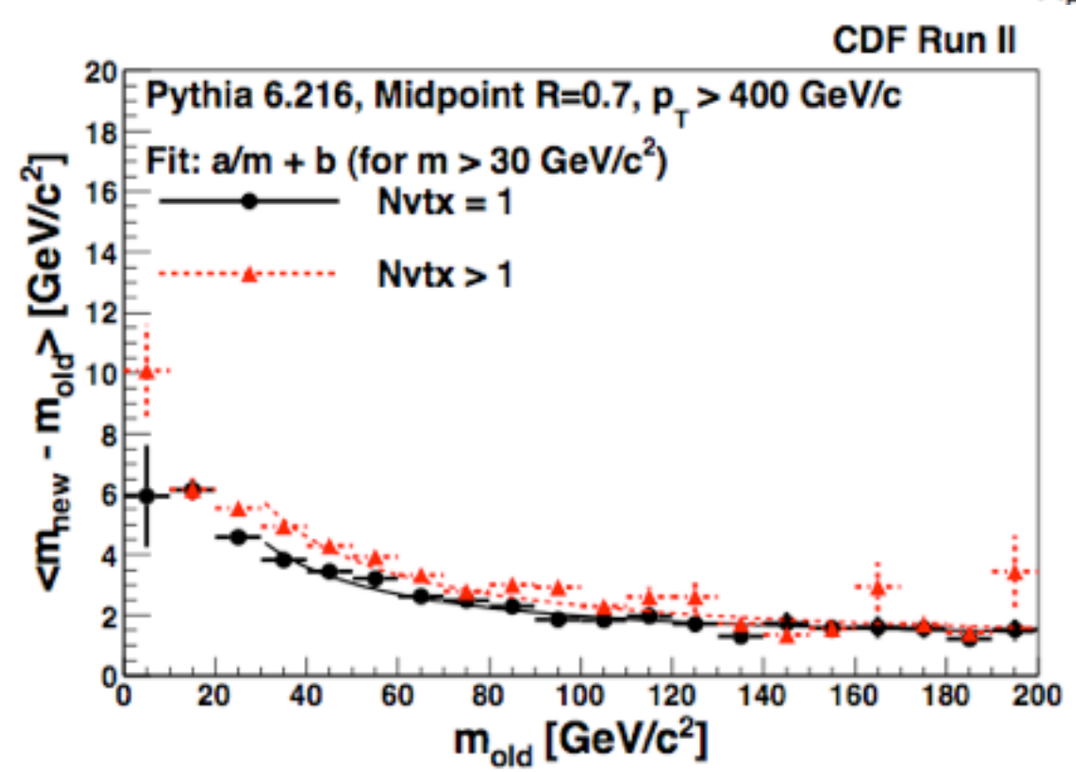
Pileup Dependence

CMS-PAS-JME-10-013



ATLAS-CONF-2011-073





- Some pileup dependence seen, but largely mitigated by substructure tools!

Searches with Substructure Techniques

Searches

- So far, most work done on characterizing algorithms in data
  - Compare QCD jets data/MC
  - Found the first signals of resonances within jets

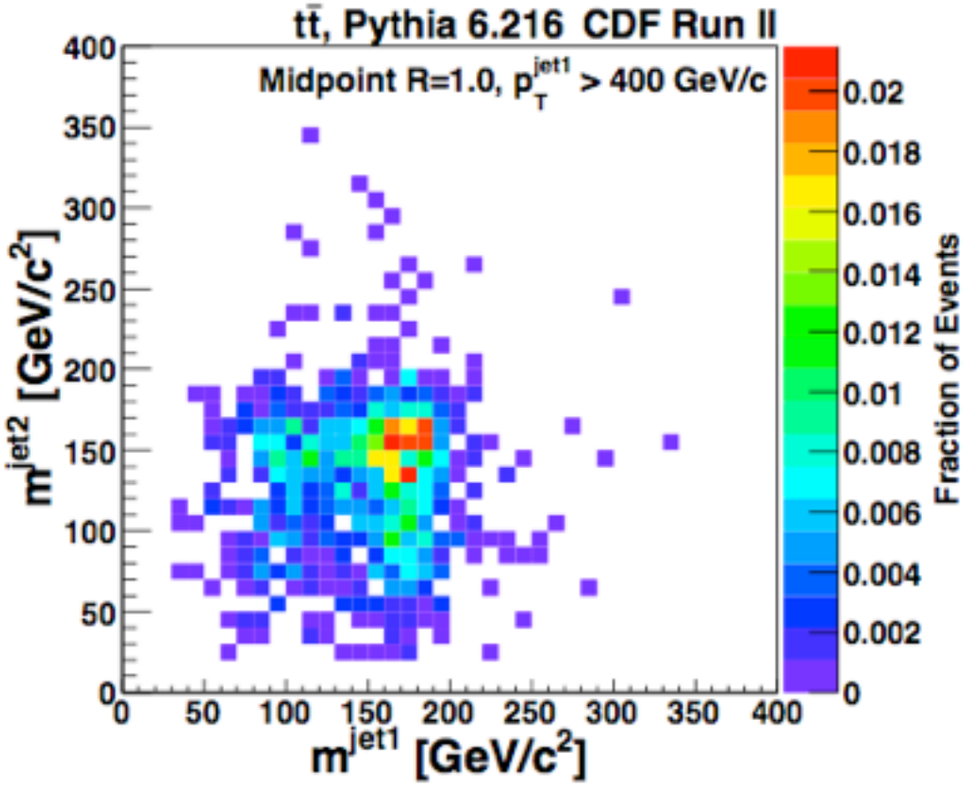
- However, already started using new ideas in searches to good effect

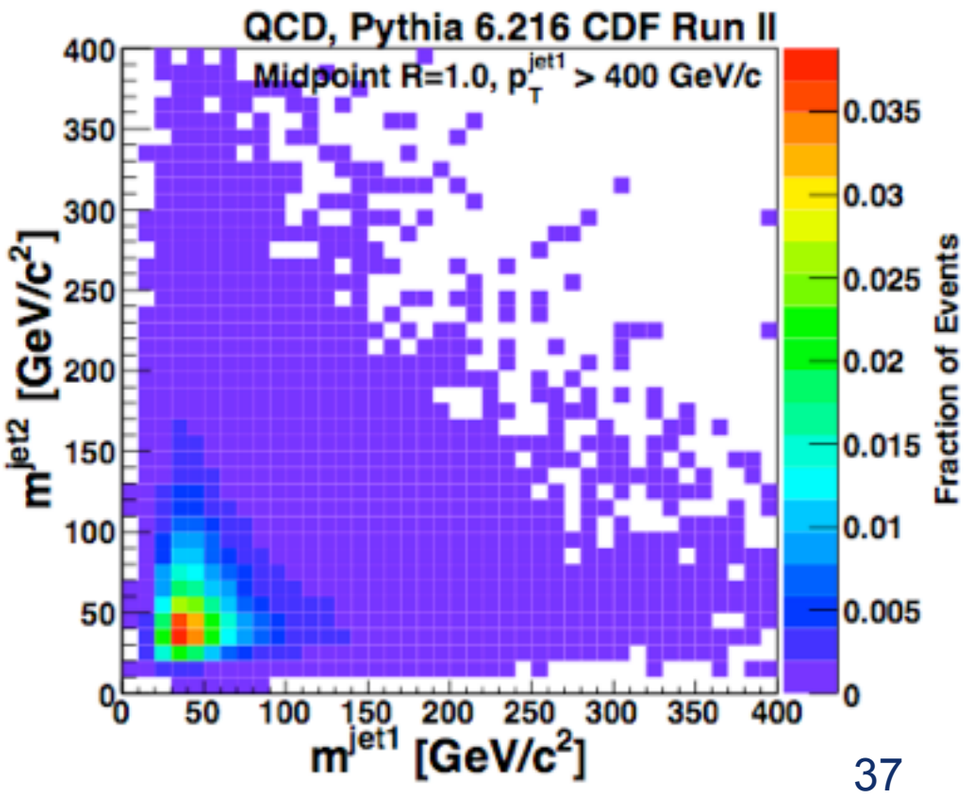
TTBar Search : CDF

CDF/ANAL/TOP/ PUBLIC/10234

- Use data-driven 2-dim sideband method
- Look for excess above QCD background

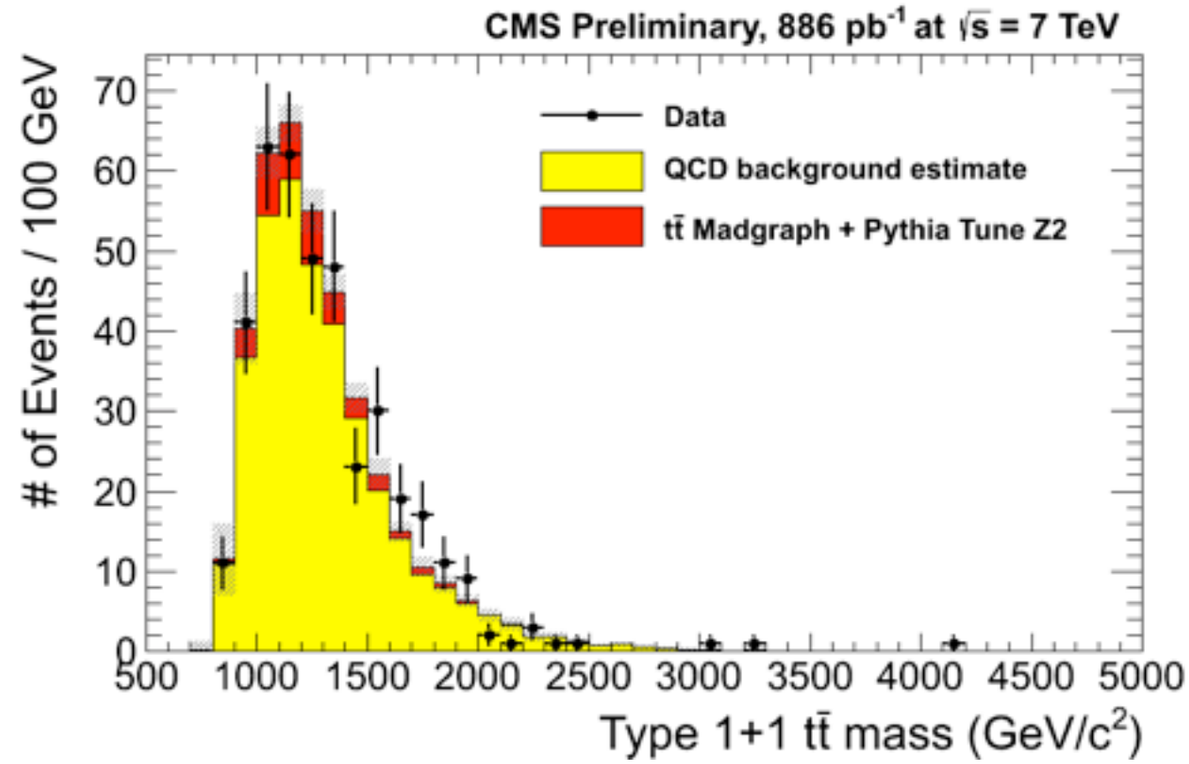
- 2 Channels:
  - All hadronic
    - Dijet
  - Semileptonic
    - MET+jet versus jet
- Observe 57, expect 46 +- 16
- No evidence for new physics, set a limit of 20 fb-1 for cross section of new objects with jet pt > 400 GeV

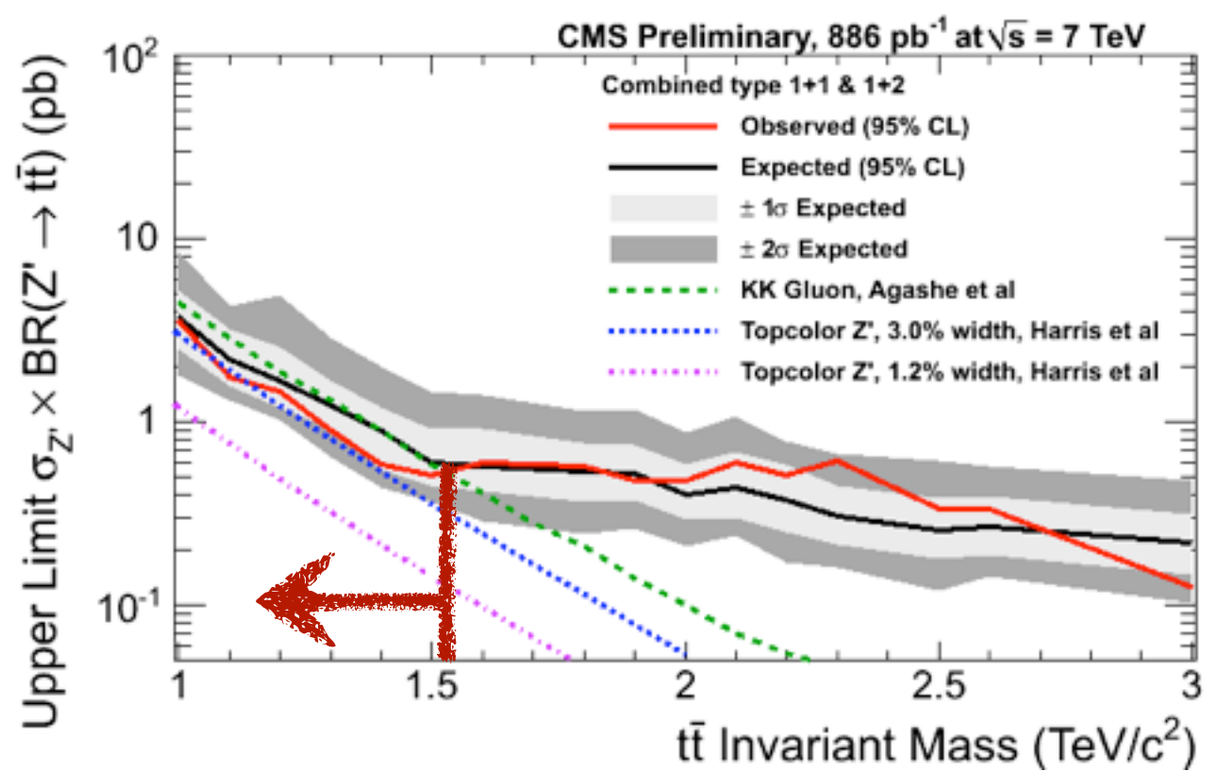




TTBar Search : CMS

- Mttbar Search
  - Yellow is data-driven QCD background
    - Weight pretag distribution with mistag rate
  - Red is TTbar
    - From MC
- Perform counting experiment in signal windows
- Exclude KK gluons with M in range 1.0-1.5 TeV

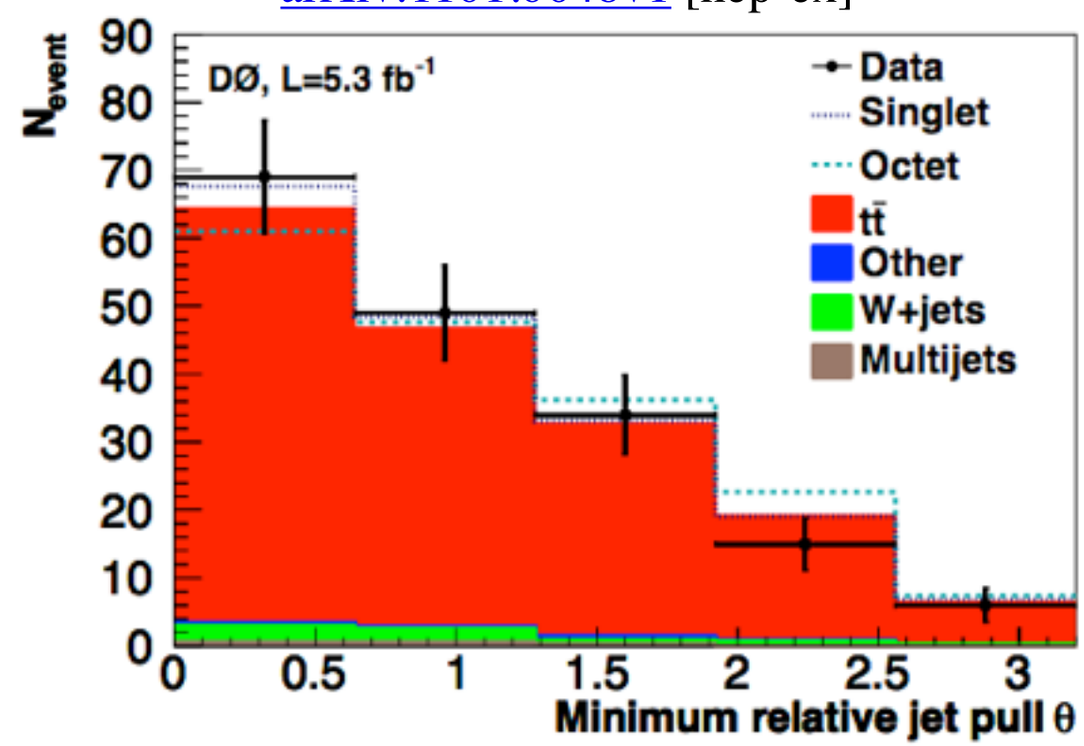




Measurement Using Jet Superstructure

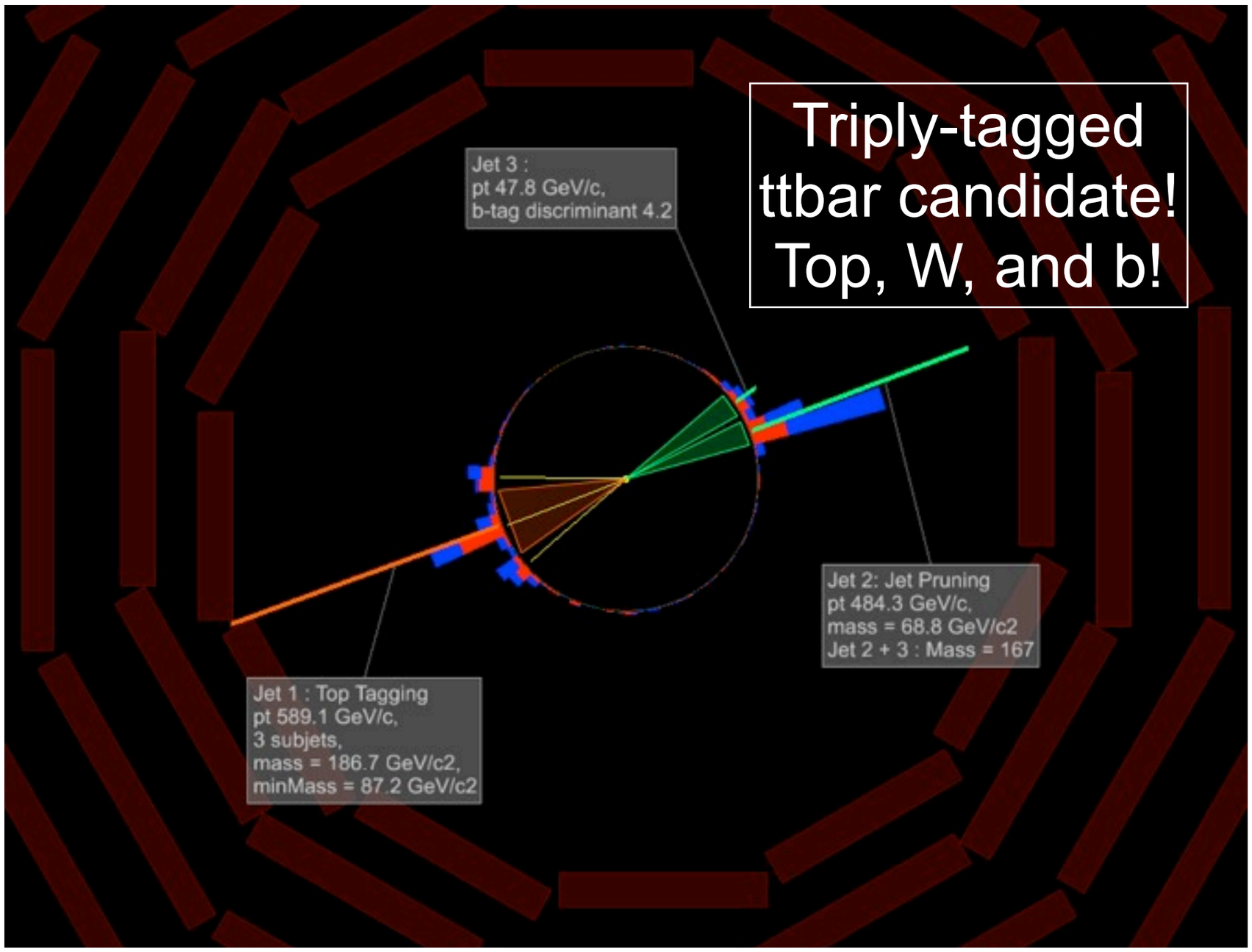
- Utilizing the jet superstructure variables, can examine color flow of hadronic W in ttbar events
- Can distinguish between color octet and color singlet models
- Find the data are consistent with the color singlet model
  - Fraction of uncolored W bosons is 0.56 +- 0.42
  - Large uncertainties, consistent with SM (1.0)

arXiv:1101.0648v1 [hep-ex]



Conclusions

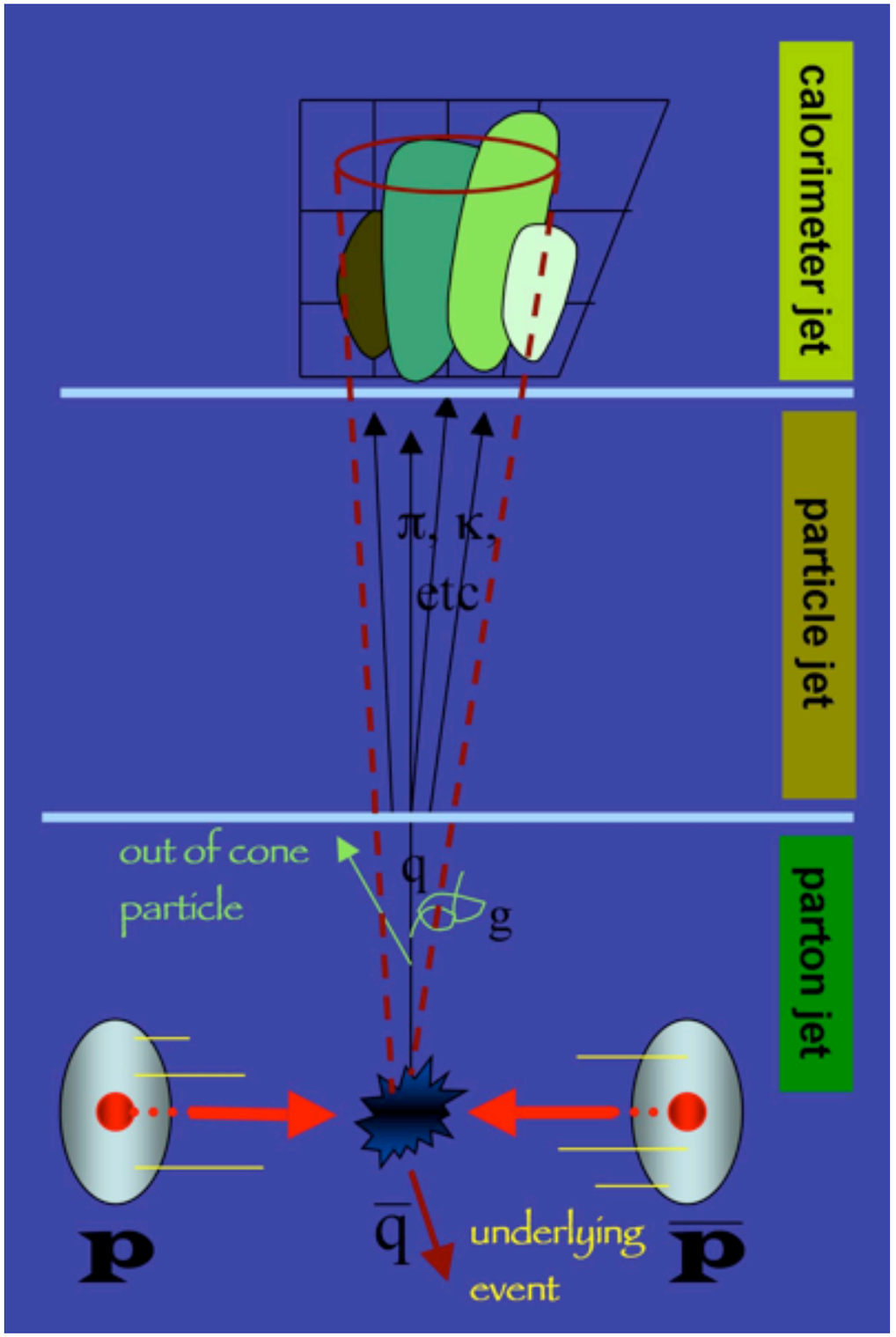
- Exciting times for jet substructure tools!
- Tevatron and LHC are both successfully deploying tools in new physics searches
- Already opening phase spaces to scrutiny for leading models of new physics



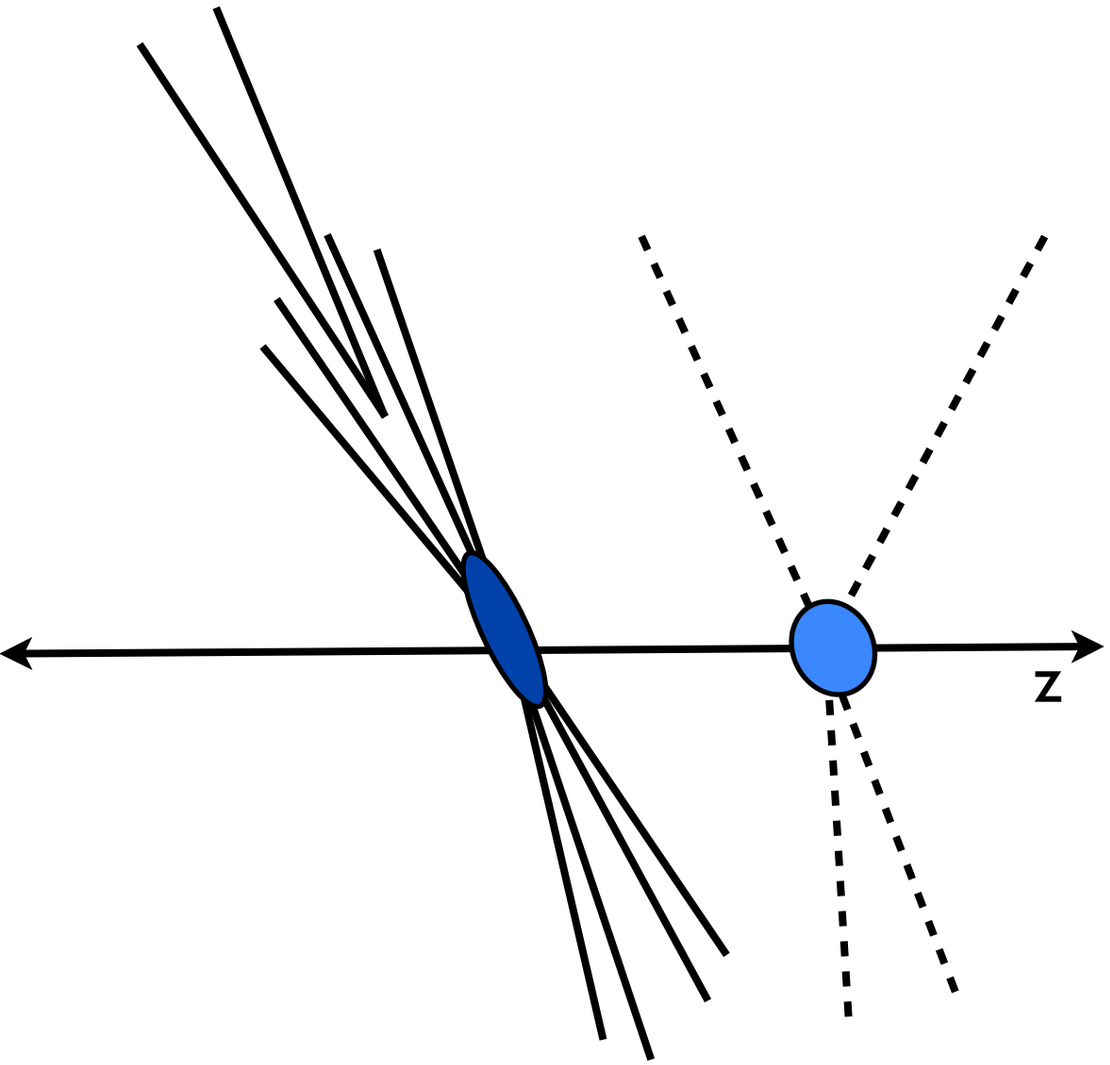
Backup Slides

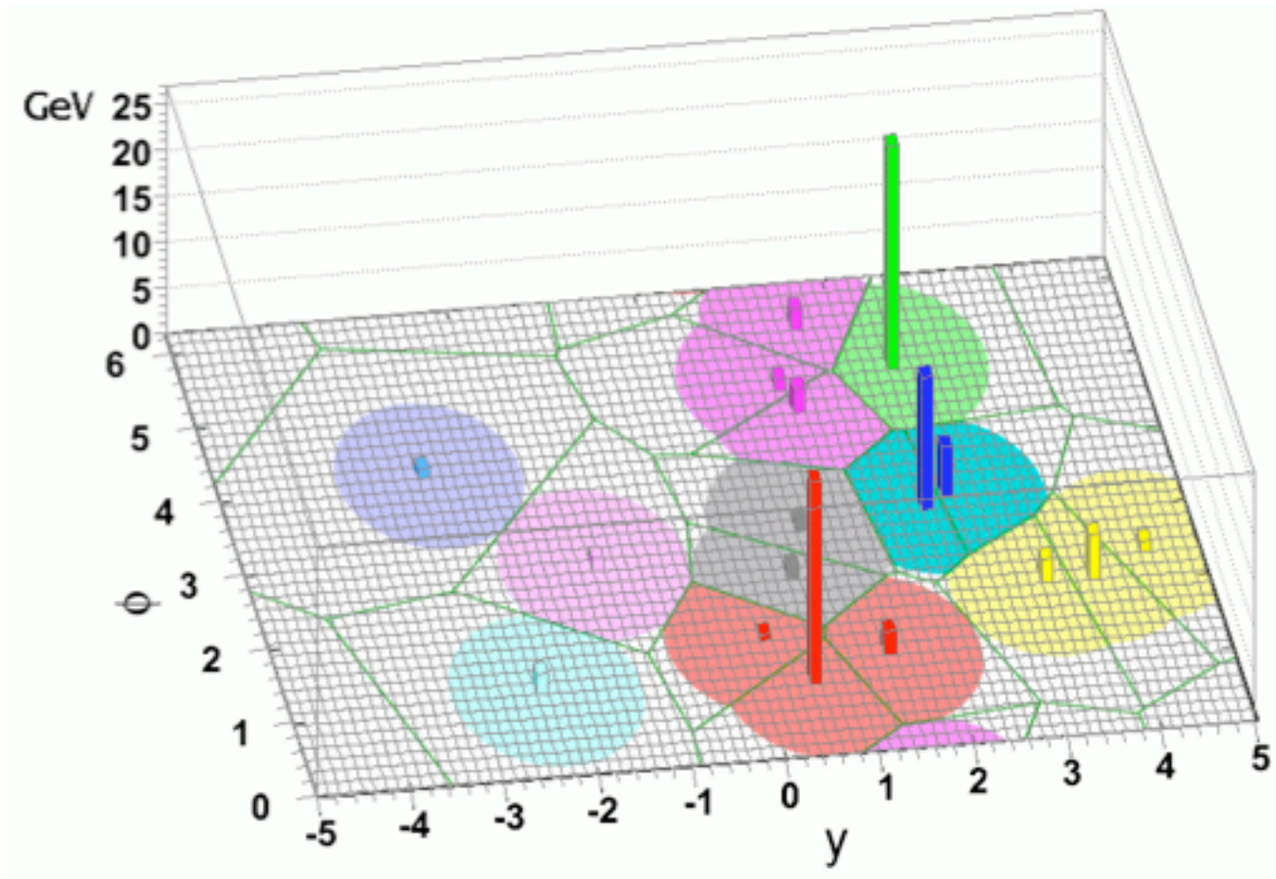
Jet Corrections

- Correct nonlinearities in detector
- Correct to constituent particles
- Correct back to parton level



Pileup



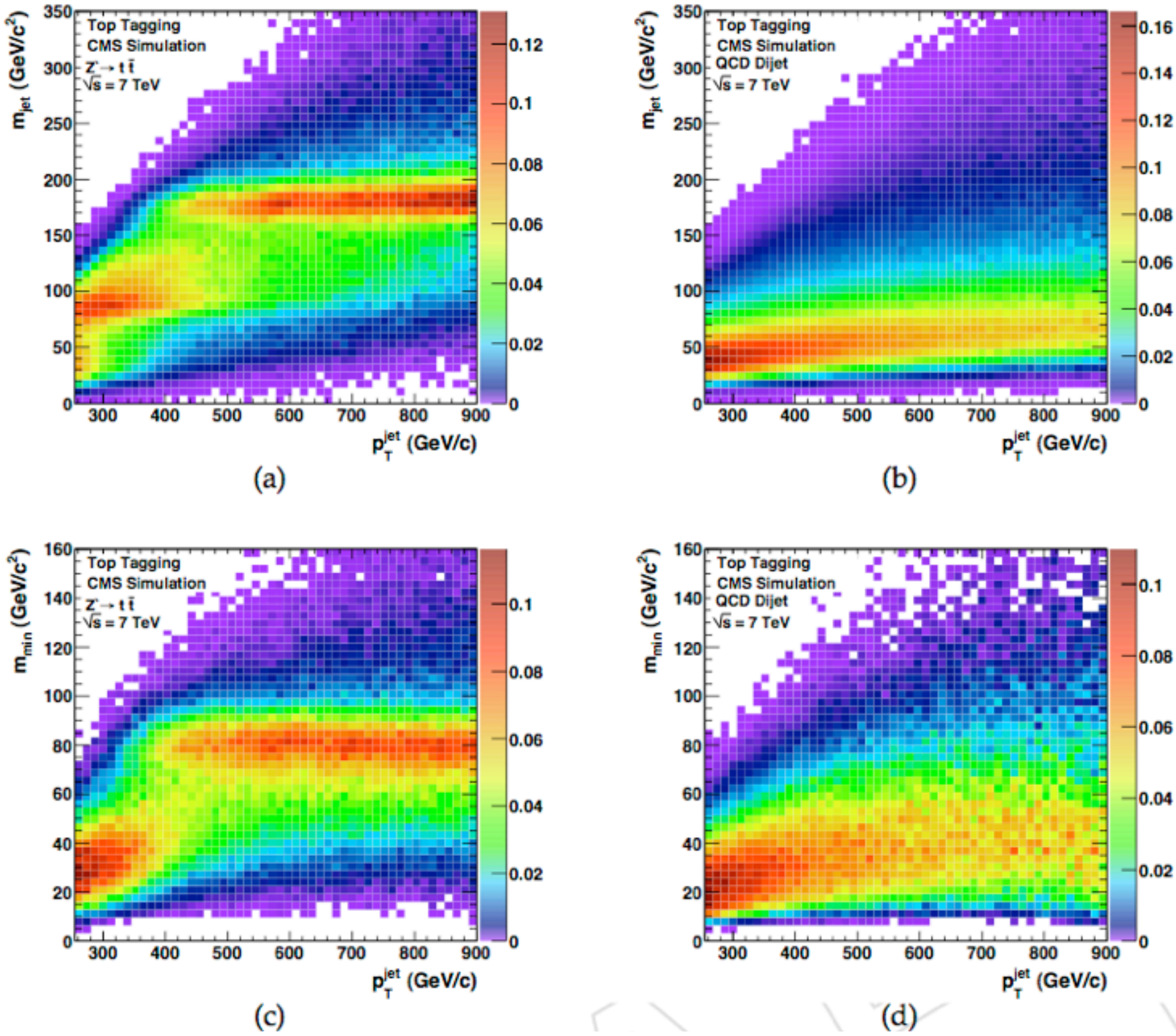


arXiv:0802.1188[hep-ph]
Cacciari, Salam, Soyez

- Explicit removal : charged hadron subtraction
- Average removal :
  - mean pt per unit area subtraction

Top Tagging Details

CMS-PAS-JME-10-013



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

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