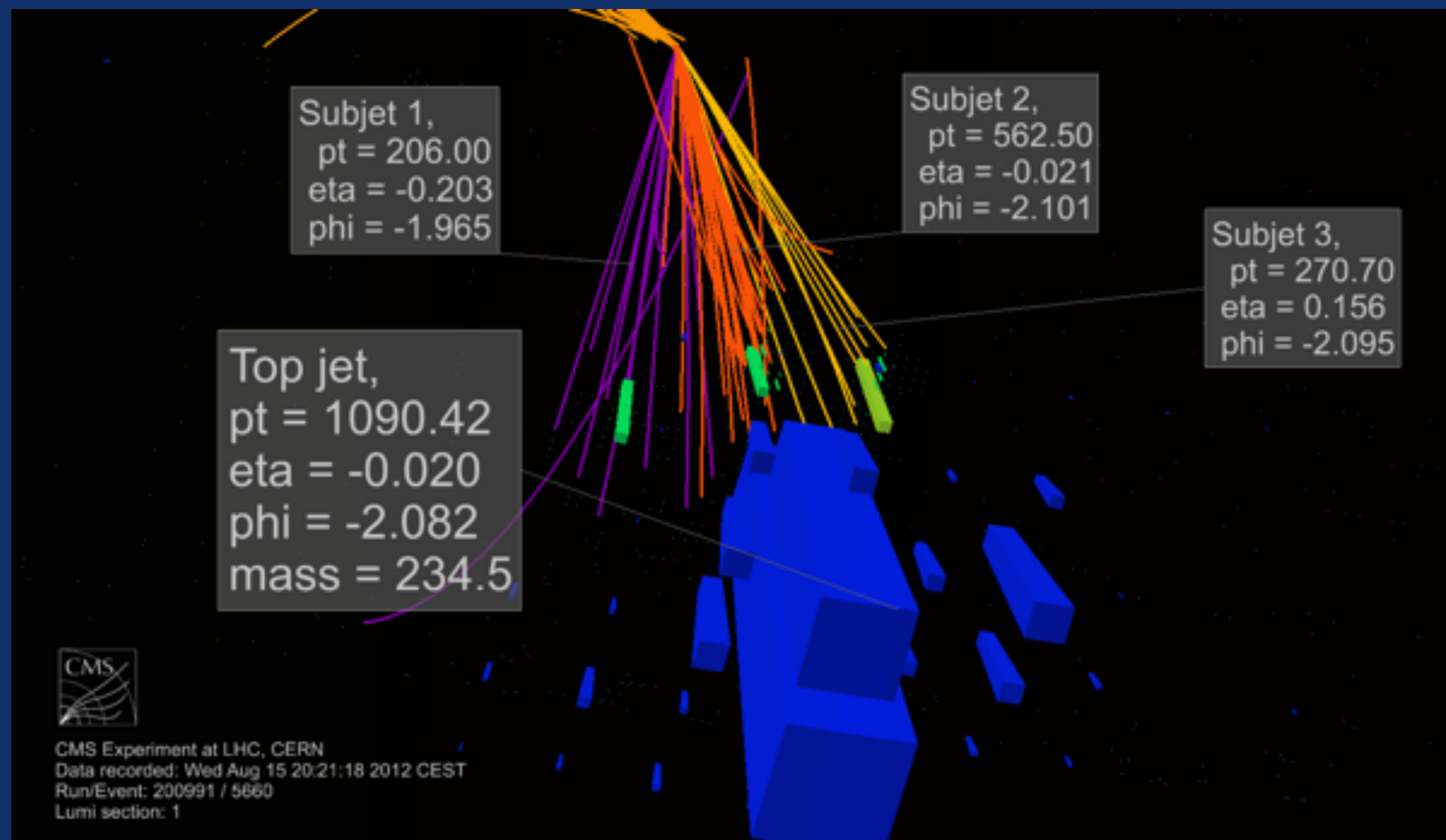
Boosted Jets and Jet Substructure



CTEQ Summer School

Salvatore Rappoccio
(State University of New York at Buffalo)

9-July 2015

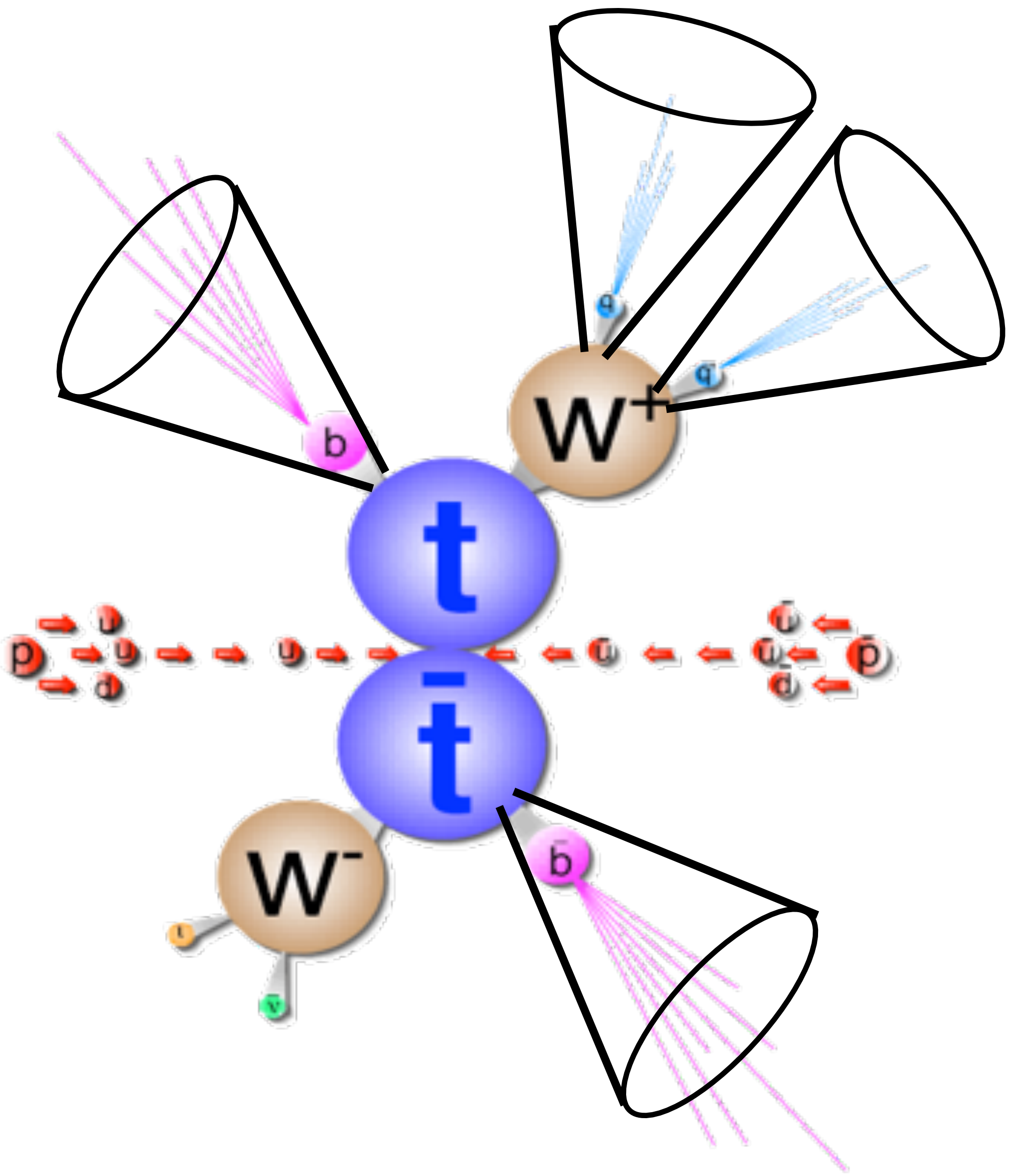
Outline

- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

(No pileup discussions today)

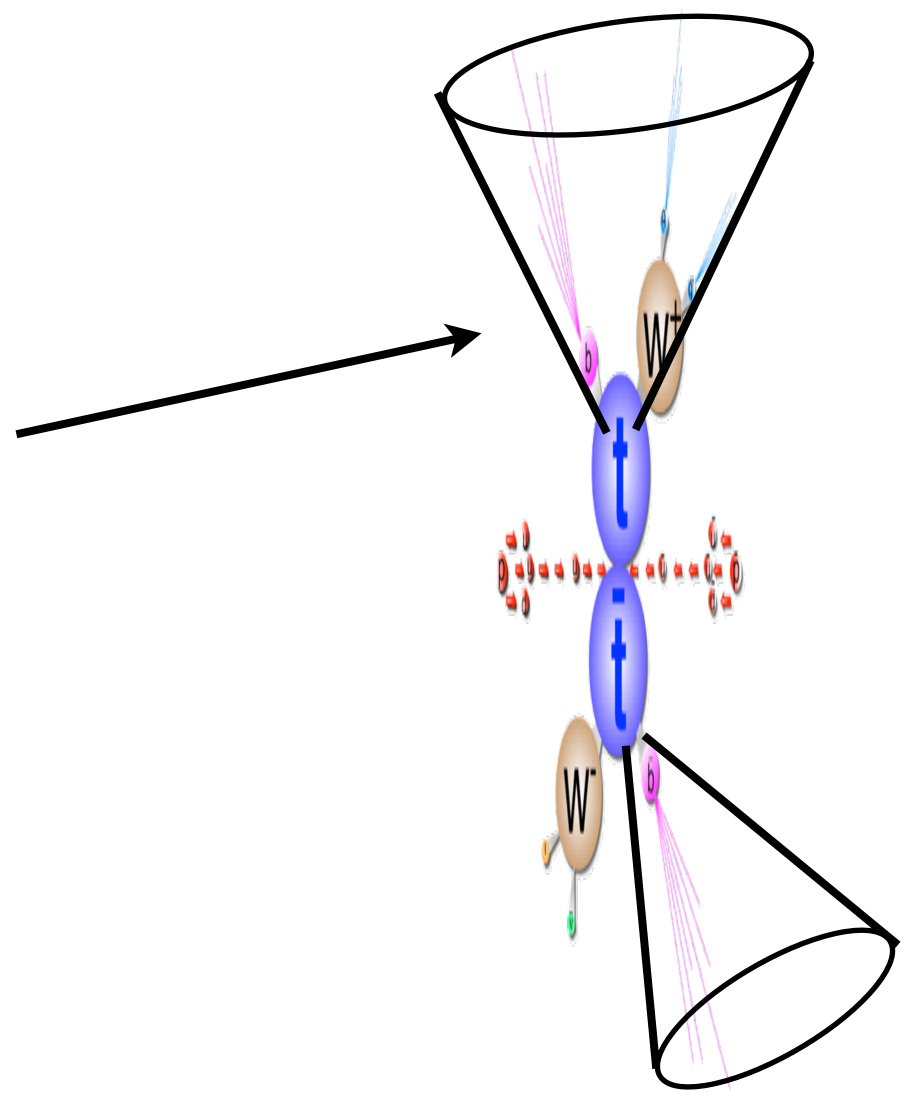
Motivation

- To study QCD partons : need jets!
- Need to associate jets to individual partons (quarks or gluons)
- Traditionally : 1-to-1 matching of jets to partons



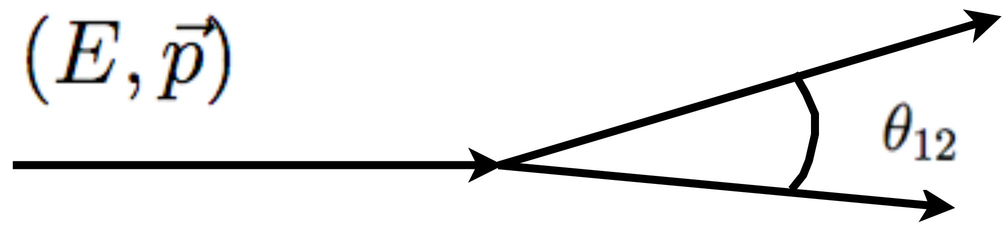
Motivation

- Problem! For boosted case cannot use this assumption!
- Have to consider cases where partons merge into a single jet



- Relativistic kinematics of boosted objects

$$p^\mu = (E, \vec{p})$$



$$p_1^\mu = (E_1, \vec{p_1})$$

$$p_2^\mu = (E_2, \vec{p_2})$$

$$\begin{aligned} m^2 &= (p_1^\mu + p_2^\mu)^2 \\ &= p_1^2 + p_2^2 + 2p_1^\mu p_{2,\mu} \end{aligned}$$

If m1 = m2 = 0 :

$$m^2 = 2(E_1 E_2 - \vec{p_1} \cdot \vec{p_2})$$

$$m^2 = 2E_1 E_2 (1 - \cos\theta_{12})$$

If E1=E2= E /2 :

$$m^2 = \frac{E^2}{2}(1 - \cos\theta_{12})$$

$$2\frac{m^2}{E^2} = (1 - \cos\theta_{12})$$

- Small angle approximation for theta :

$$2\frac{m^2}{E^2} = (1 - \cos\theta_{12})$$

$$2\frac{m^2}{E^2} = (1 - (1 - \frac{\theta_{12}^2}{2}) = \frac{\theta_{12}^2}{2}$$

$$4\frac{m^2}{E^2} = \theta_{12}^2$$

$$\theta_{12} = 2\frac{m}{E} = \frac{2}{\gamma}$$

Motivation

- Limiting cases:

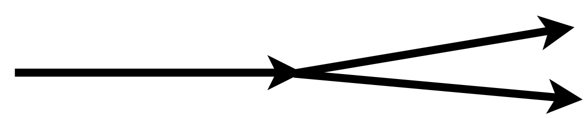
Small angle

$$\theta_{12} = 2\frac{m}{E} = \frac{2}{\gamma}$$

Full expression

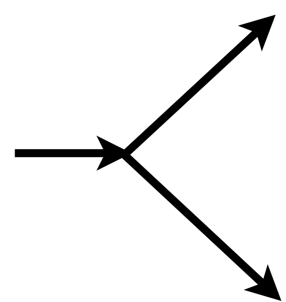
$$2\frac{m^2}{E^2} = (1 - \cos\theta_{12})$$

$$\gamma \rightarrow \infty$$



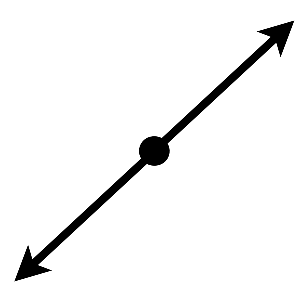
highly Lorentz boosted

$$\gamma \rightarrow 2$$



moderately Lorentz boosted

$$\gamma \rightarrow 1$$



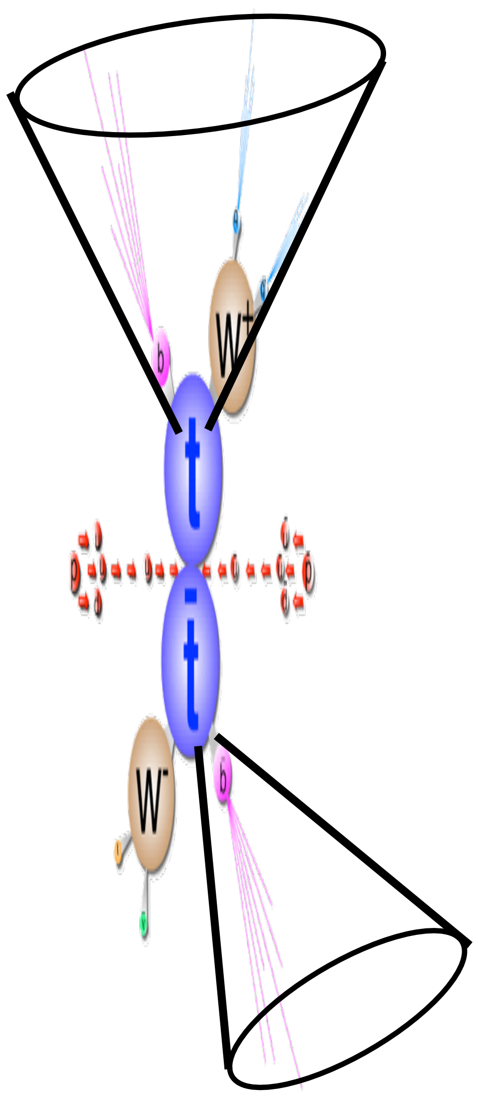
at rest : back to back
(use full expression again)

Motivation

- Now can quantify when 1-to-1 parton-jet assignment breaks down!

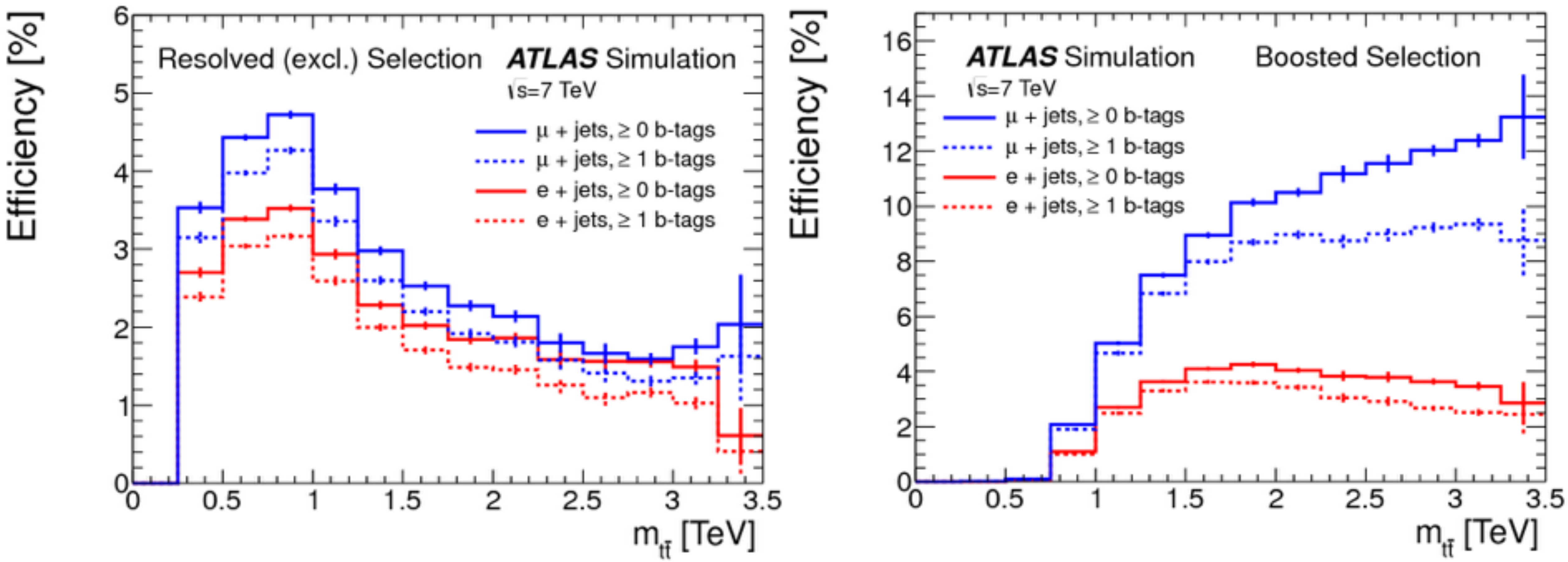
$$\theta_{12} = 2\frac{m}{E} = \frac{2}{\gamma}$$

- If theta required to be > 0.8, you throw away events with gamma > 2.5



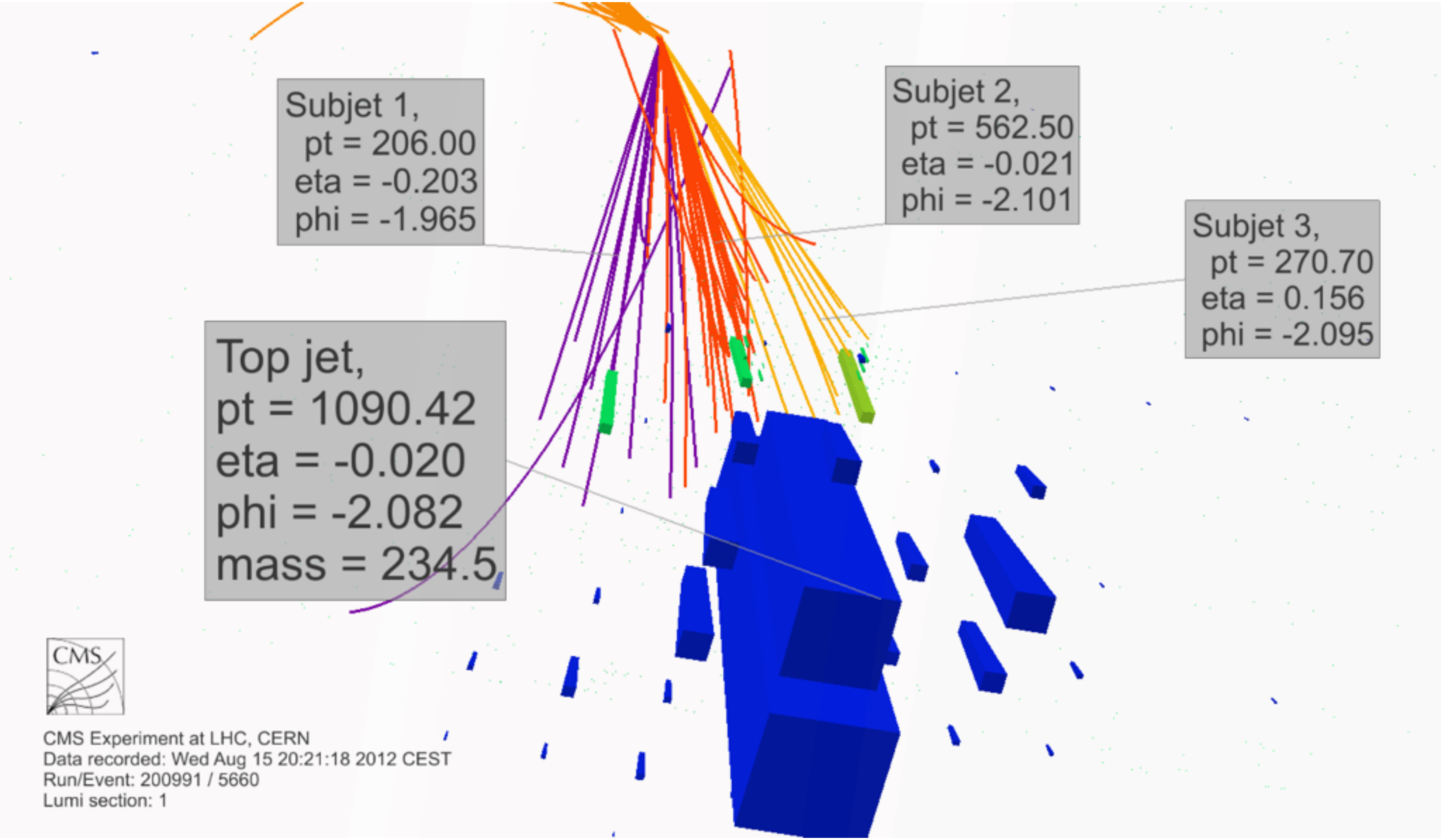
Inadvertently place a maximum E cut!

- Can see this in analyses



“Resolved” selections turn off, “boosted” selections turn on

Motivation



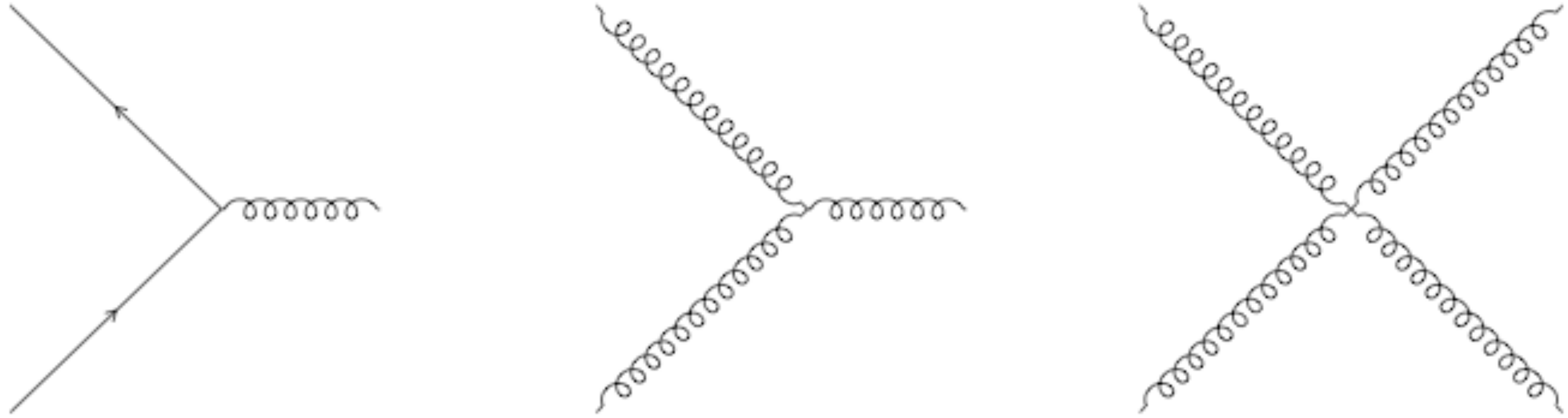
Subjet 1,
pt = 206.00
eta = -0.203
phi = -1.965
Subjet 2,
pt = 562.50
eta = -0.021
phi = -2.101
Subjet 3,
pt = 270.70
eta = 0.156
phi = -2.095
Top jet,
pt = 1090.42
eta = -0.020
phi = -2.082
mass = 234.5
CMS
CMS Experiment at LHC, CERN
Data recorded: Wed Aug 15 20:21:18 2012 CEST
Run/Event: 200991 / 5660
Lumi section: 1

Outline

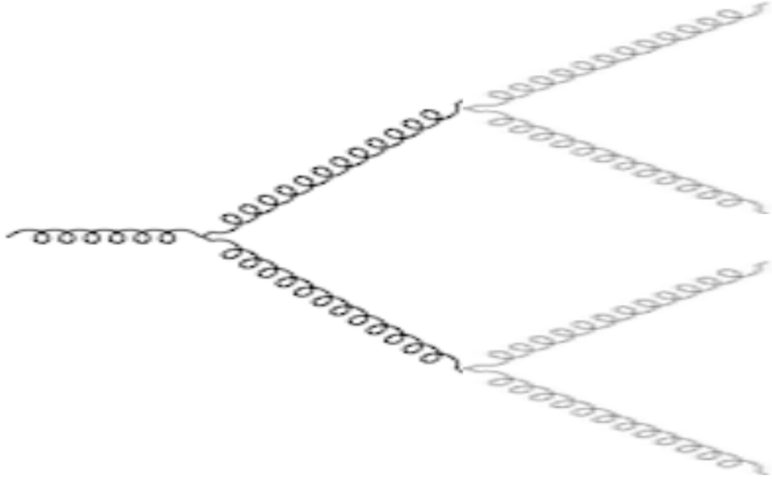
- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

A little QCD

- Let's take it from the top:
  - QCD has these vertices:



  - An evolution of a QCD final state should therefore mostly have the first two, but in reverse (1->2)



Sequential Clustering Algorithms

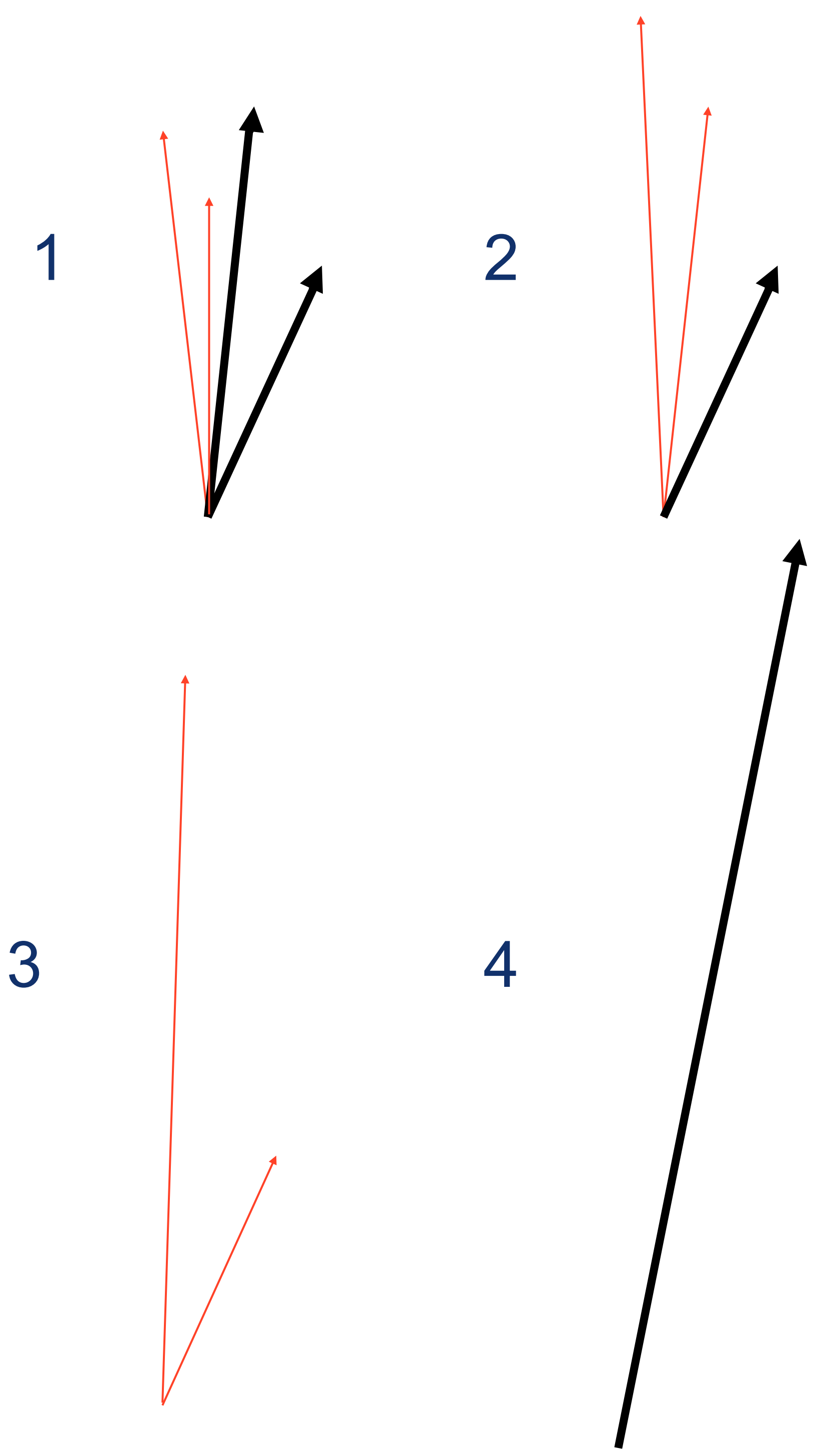
- Pairwise examination of input 4-vectors
- Calculate

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

- Also find the "beam distance"

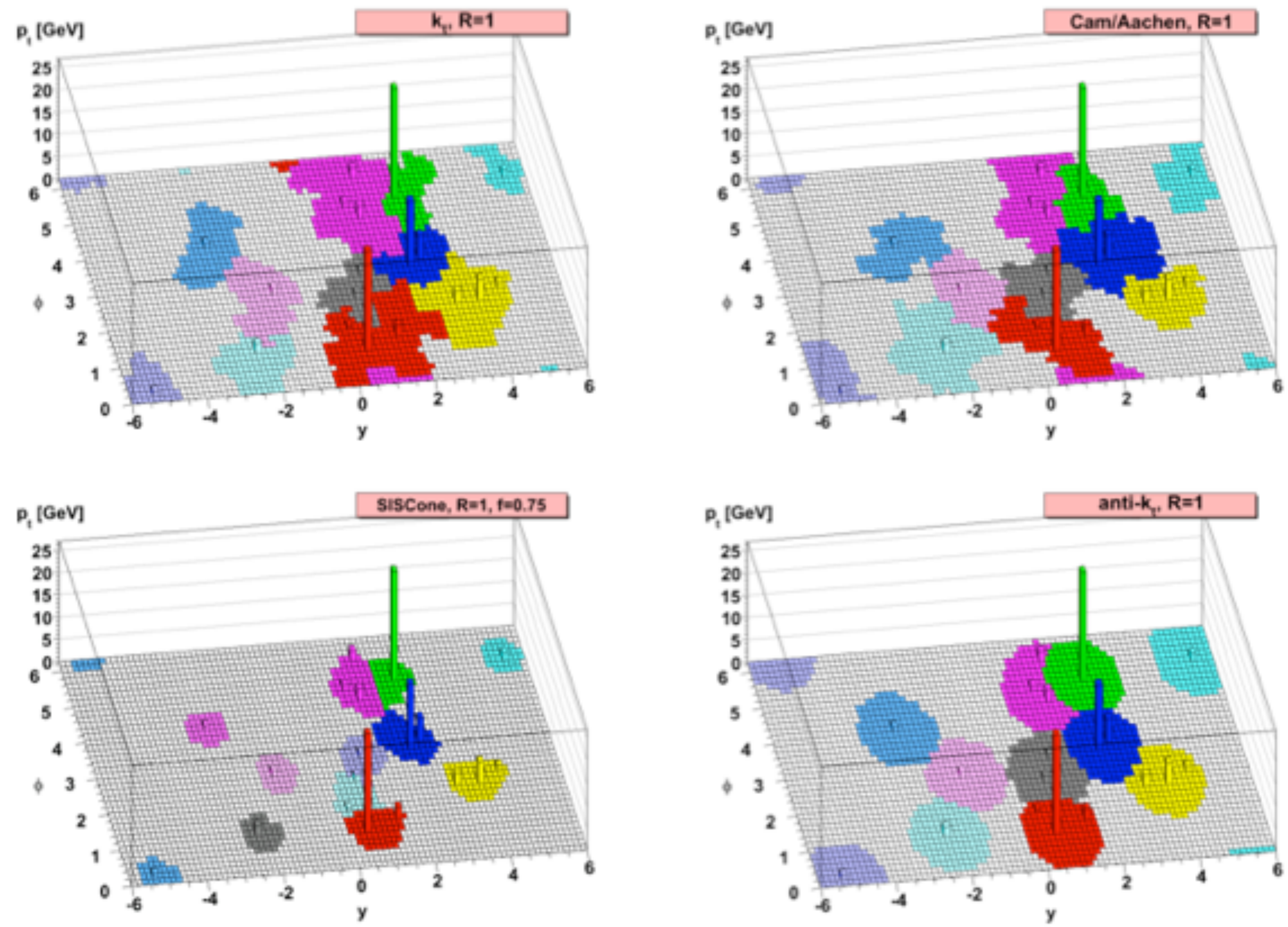
$$d_{iB} = k_{T,i}^n$$

- 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

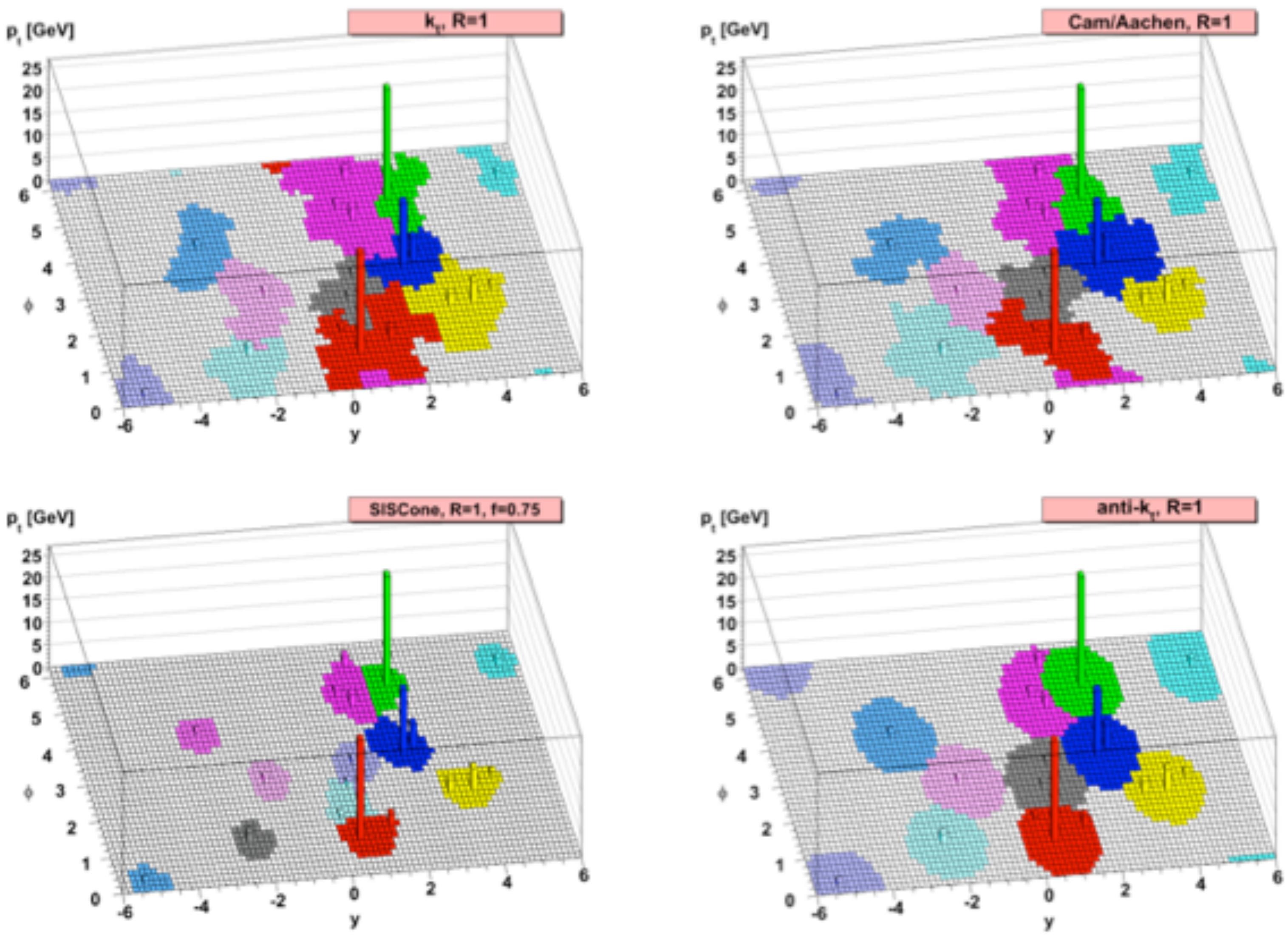
- Different types
  - N = 2: "kT"
    - "QCD in reverse"
  - N = 0 : "Cambridge-Aachen" (CA)
    - Distance only, irregular, very useful for substructure!
  - N = -2: "anti-kT"
    - "Idealized" cone algorithm



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

Sequential Clustering Algorithms

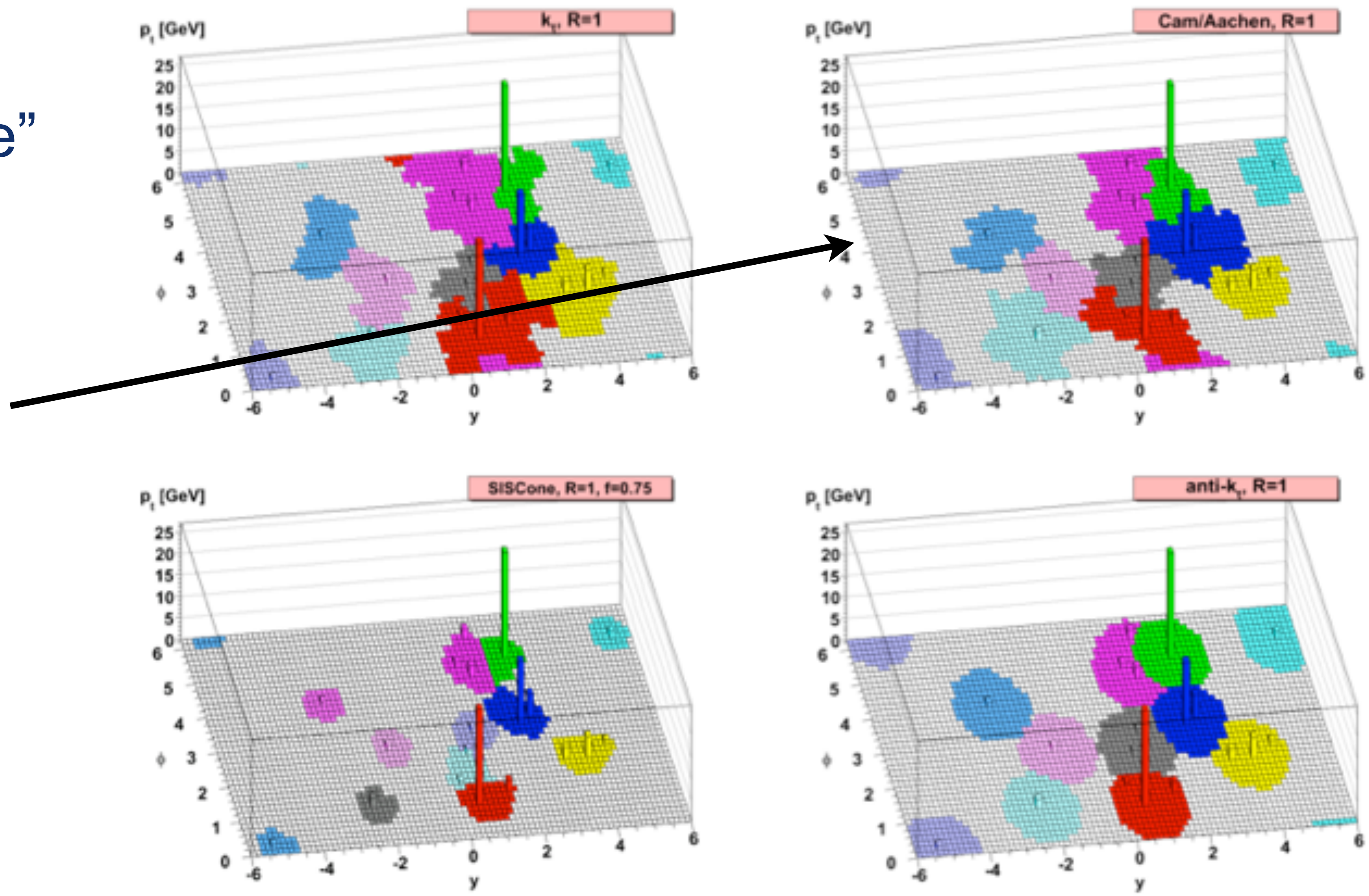
- Different types
  - N = 2: “kT”
    - “QCD in reverse”
  - N = 0 : “Cambridge-Aachen” (CA)
    - Distance only, irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



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

Sequential Clustering Algorithms

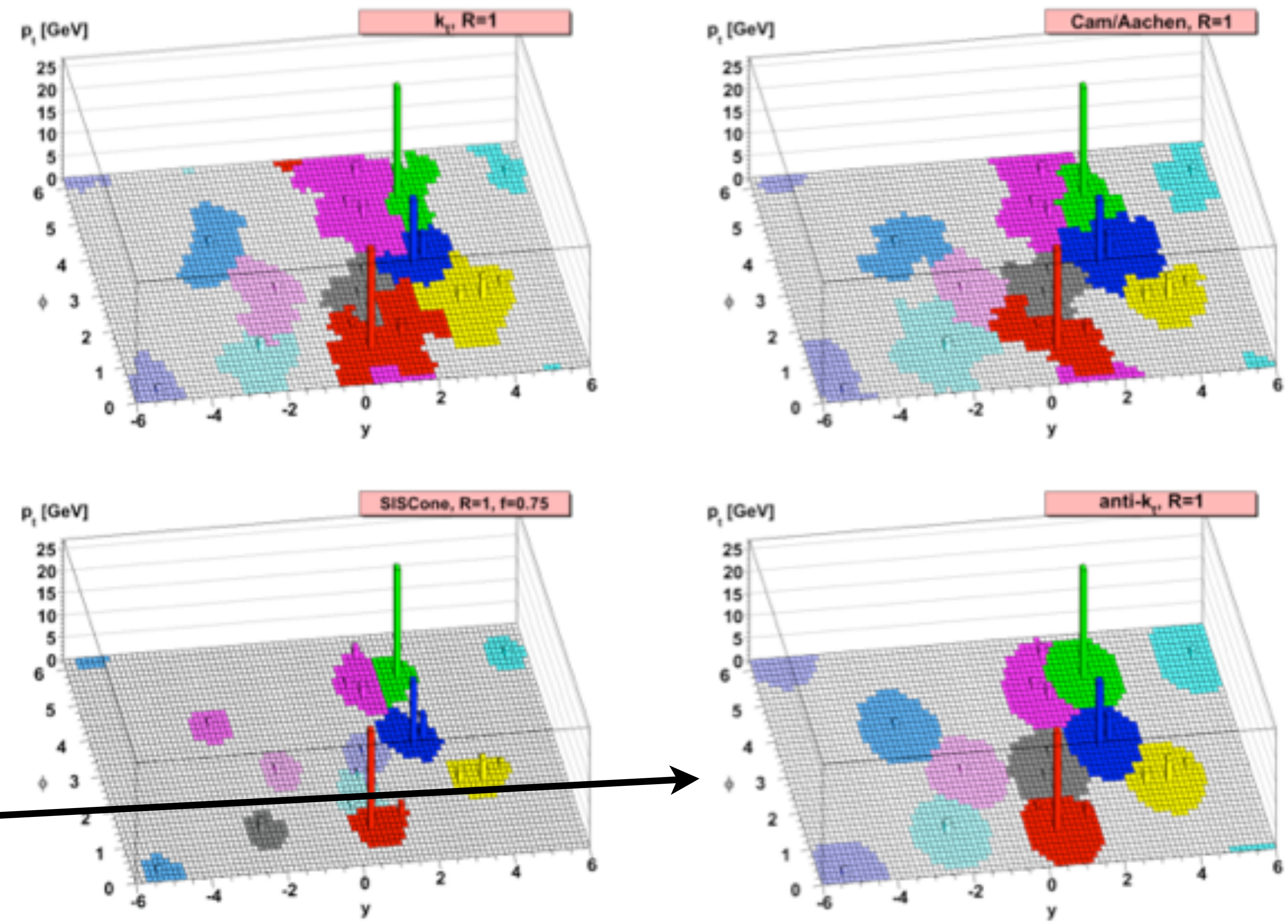
- Different types
  - N = 2: “kT”
    - “QCD in reverse”
  - N = 0 : “Cambridge-Aachen” (CA)
    - Distance only, irregular, very useful for substructure!
  - N = -2: “anti-kT”
    - “Idealized” cone algorithm



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

Sequential Clustering Algorithms

- Different types
  - N = 2: "kT"
    - "QCD in reverse"
  - N = 0 : "Cambridge-Aachen" (CA)
    - Distance only, irregular, very useful for substructure!
  - N = -2: "anti-kT"
    - "Idealized" cone algorithm



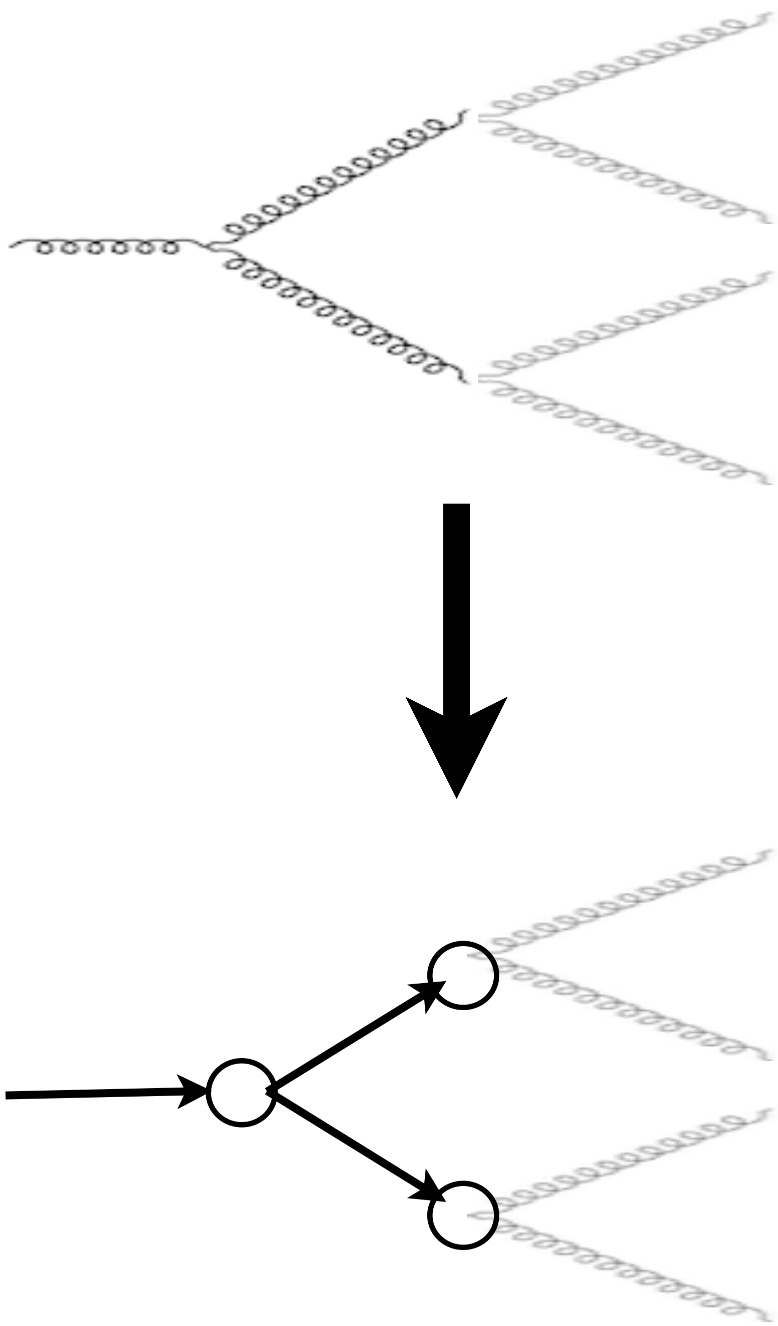
arXiv:0802.1189v2 [hep-ph]
Cacciari, Salam, Soyez

Outline

- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

Jet Substructure

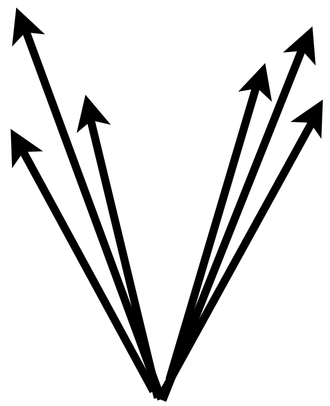
- Let's go back to the QCD jet... what happens if there's a particle with mass > 0 at a vertex?
  - Different kinematics
  - Invariant mass of constituents is > 0
  - The clustering sequence has specific features that relate to mass scales



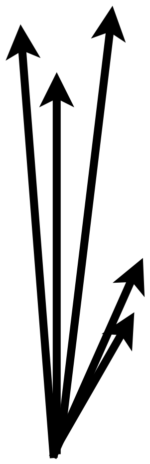
Jet Substructure

$$\theta_{12} = 2\frac{m}{E} = \frac{2}{\gamma}$$

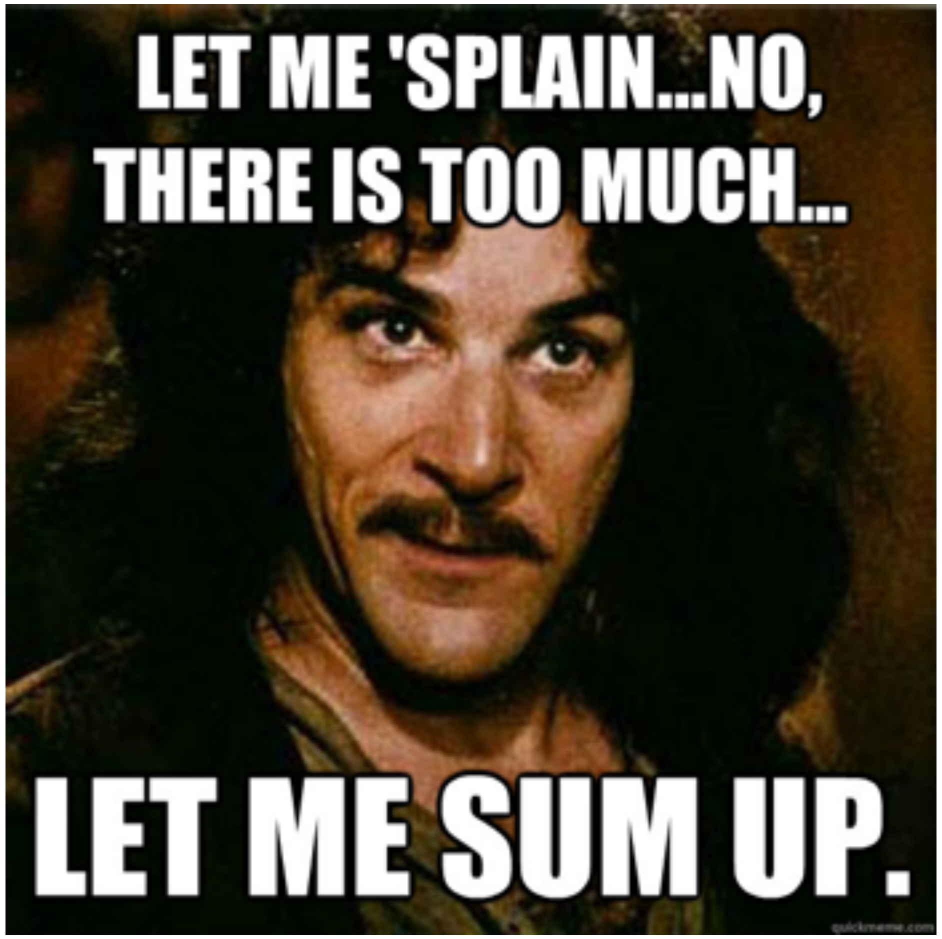
- Massive particles:
  - Wider-angle splittings
  - Symmetric splittings



- QCD Jets : m ~ 0
  - Many low-angle splittings
  - Asymmetric splittings



History of Jet Substructure and Boosted Jet Tools



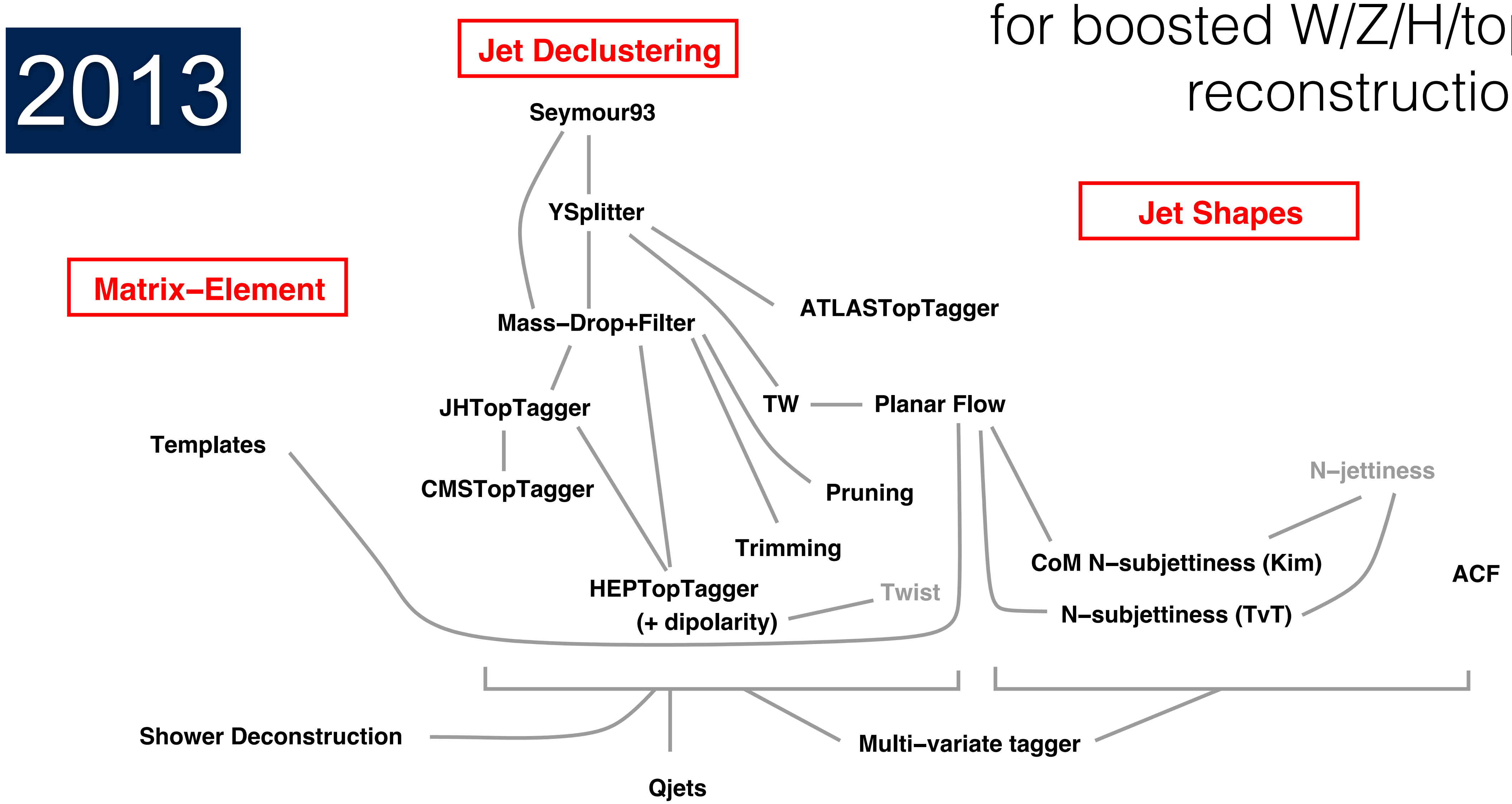
- M. H. Seymour, Phys. C62 (1994) 127–138
- J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, Phys. Rev. Lett. 100 (2008) 242001
- S. D. Ellis, C. K. Vermilion and J. R. Walsh, Phys. Rev. D 81 (2010) 094023
- S. D. Ellis, C. K. Vermilion and J. R. Walsh, Phys. Rev. D 80 (2009) 051501
- D. Krohn, J. Thaler and L. -T. Wang, JHEP 1002 (2010) 084
- Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, JHEP 9708 (1997) 001
- M. Wobisch and T. Wengler, In *Hamburg 1998/1999, Monte Carlo generators for HERA physics* 270-279
- M. Dasgupta, A. Fregoso, S. Marzani and A. Powling, arXiv:1307.0013
- M. Dasgupta, L. Magnea and G. P. Salam, JHEP 0802 (2008) 055
- D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys.Rev.Lett. 101 (2008) 142001
- ATLAS Collaboration, ATL-PHYS-CONF-2008-008. ATL-COM-PHYS-2008-001
- CMS Collaboration, CMS-PAS-JME-09-001, 2009
- J. Thaler and K. Van Tilburg, JHEP 1103 (2011) 015
- T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, JHEP 1010 (2010) 078
- Backovic, Gabizon, Juknevich, Perez, Soreq, JHEP 1404 (2014) 176
- D. E. Soper and M. Spannowsky, Phys.Rev. D84 (2011) 074002
- M. Jankowiak and A. J. Larkoski, HEP 1106 (2011) 057
- S. D. Ellis, A. Hornig, T. S. Roy, D. Krohn, and M. D. Schwartz, Phys.Rev.Lett. 108 (2012) 182003
- M. Backovic, J. Juknevich, and G. Perez, JHEP 1307 (2013) 114
- A. J. Larkoski, G. P. Salam, and J. Thaler, JHEP 1306 (2013) 108
- J. Gallicchio and M. D. Schwartz, Phys.Rev.Lett. 107 (2011) 172001
- S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, JHEP 1011 (2010) 101
- M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, JHEP 1210 (2012) 126
- D. Bertolini, T. Chan, and J. Thaler, arXiv:1310.7584
- M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, JHEP 1309 (2013) 029,
- M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, Eur. Phys. J. C, 73 11 (2013) 2623
- D. Krohn, M. Low, M. D. Schwartz, and L.-T. Wang, arXiv:1309.4777
- A. Larkoski, S. Marzani, G. Soyez, J. Thaler,JHEP 1405 (2014) 146
- S.D. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tonnesmann, Prog.Part.Nucl.Phys. 60 (2008) 484-551

- This list is by no means exhaustive
- If you can read this, you have passed your eye exam. Congratulations.

History of Jet Substructure and Boosted Jet Tools

Very active research field

Some of the tools developed for boosted W/Z/H/top reconstruction



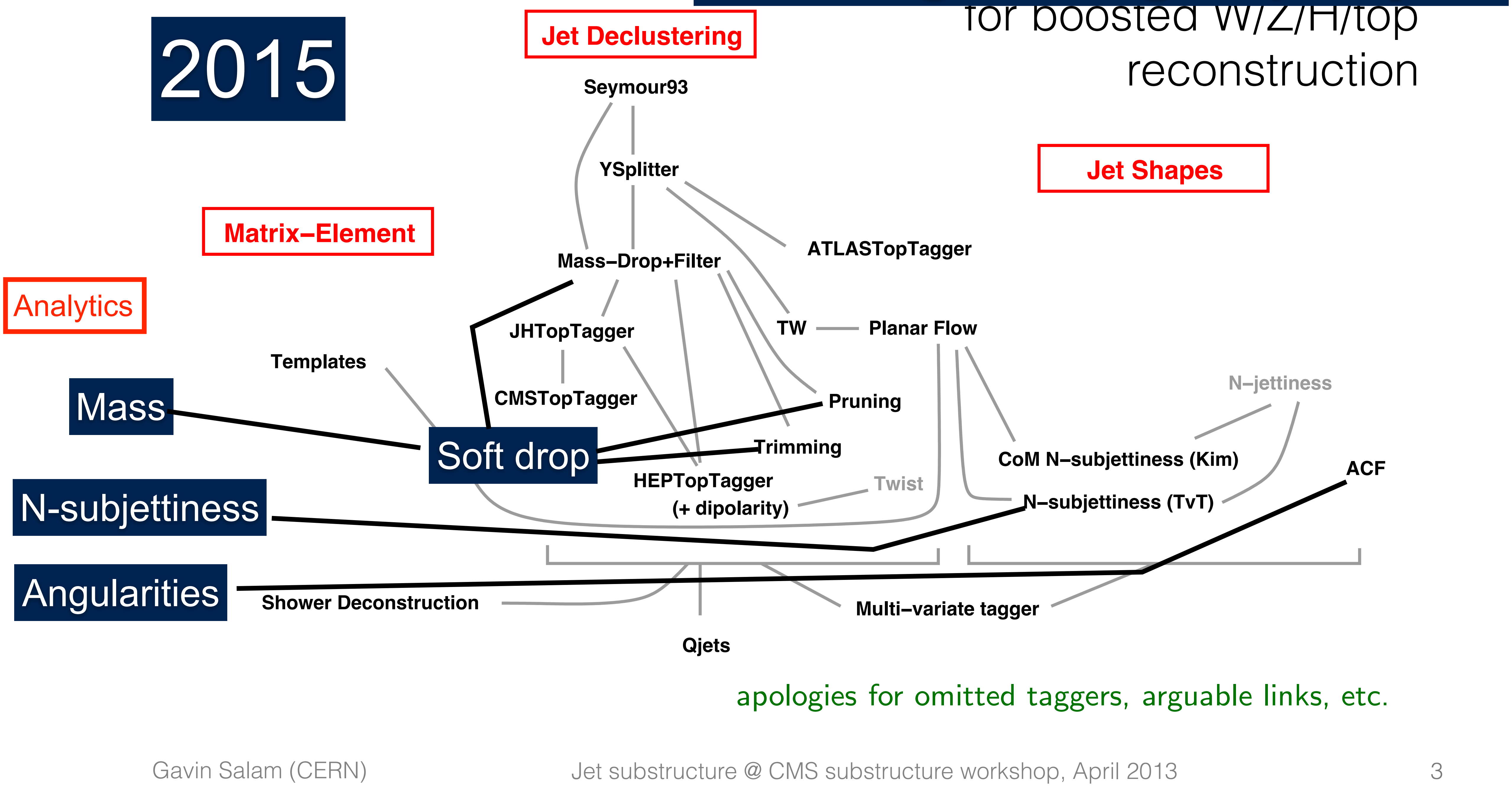
apologies for omitted taggers, arguable links, etc.

Gavin Salam (CERN) Jet substructure @ CMS substructure workshop, April 2013 3

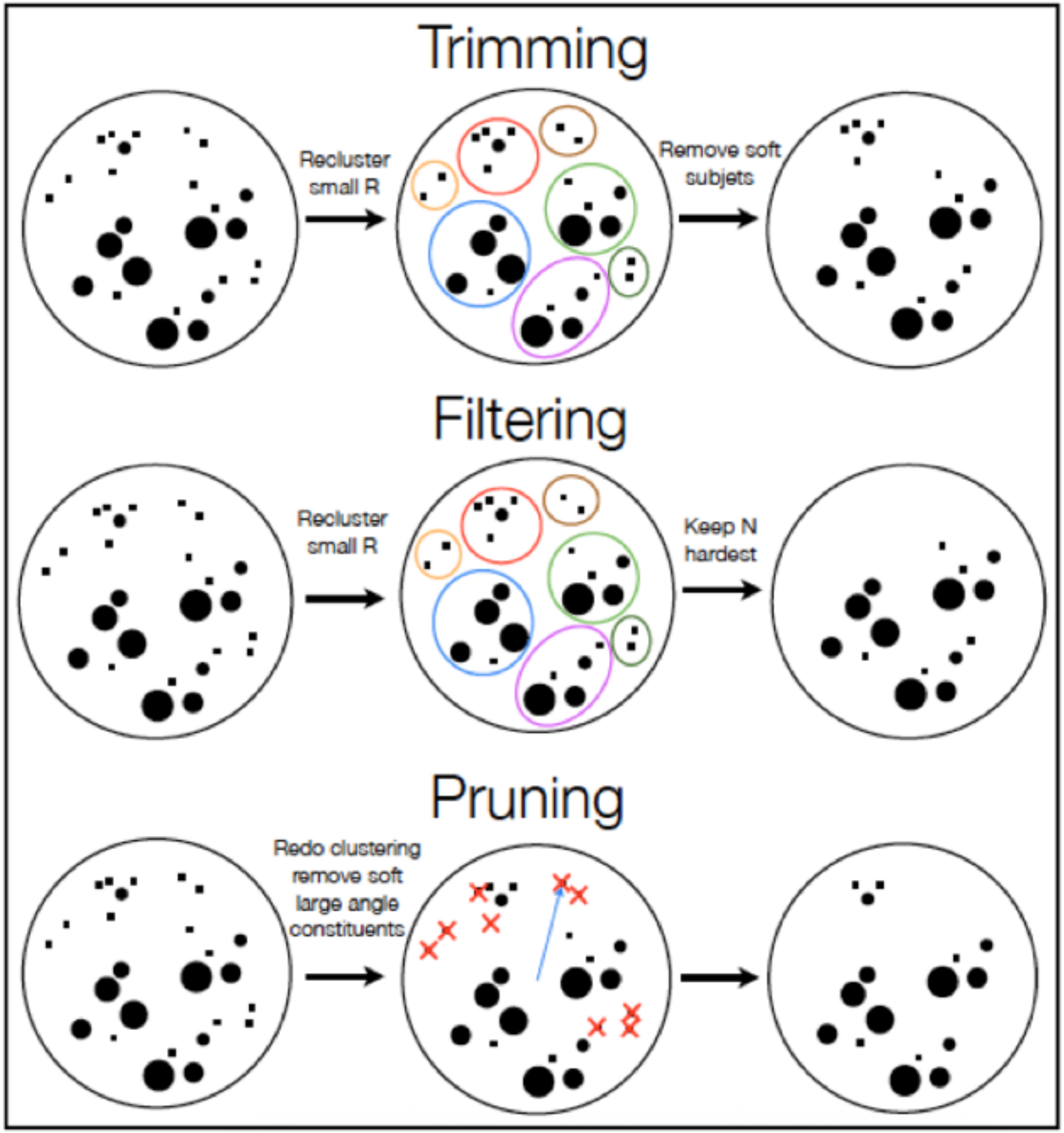
History of Jet Substructure and Boosted Jet Tools

Very active research field

And growing rapidly!



Jet Grooming



Outline

- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

Jet Analytics

- Major advances recently in first-principles analytic calculations of jet properties
  - Overhauled understanding of what these techniques are actually doing
  - Allowed a formation of “theoretically sound” techniques
  - Informing decisions for the experiments!

Jet Analytics

- First need to understand jet mass

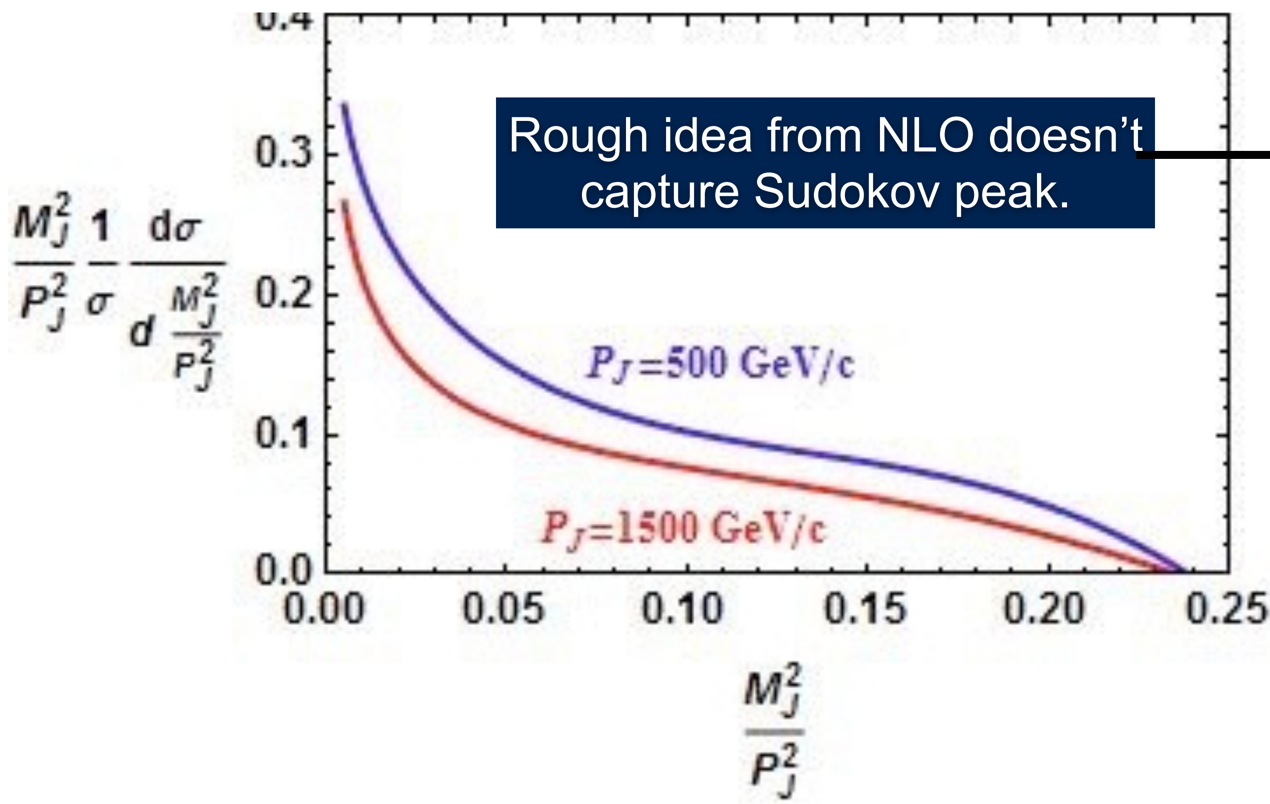
At NLO :

Log. divergence at low mass

Scales ~linearly with momentum

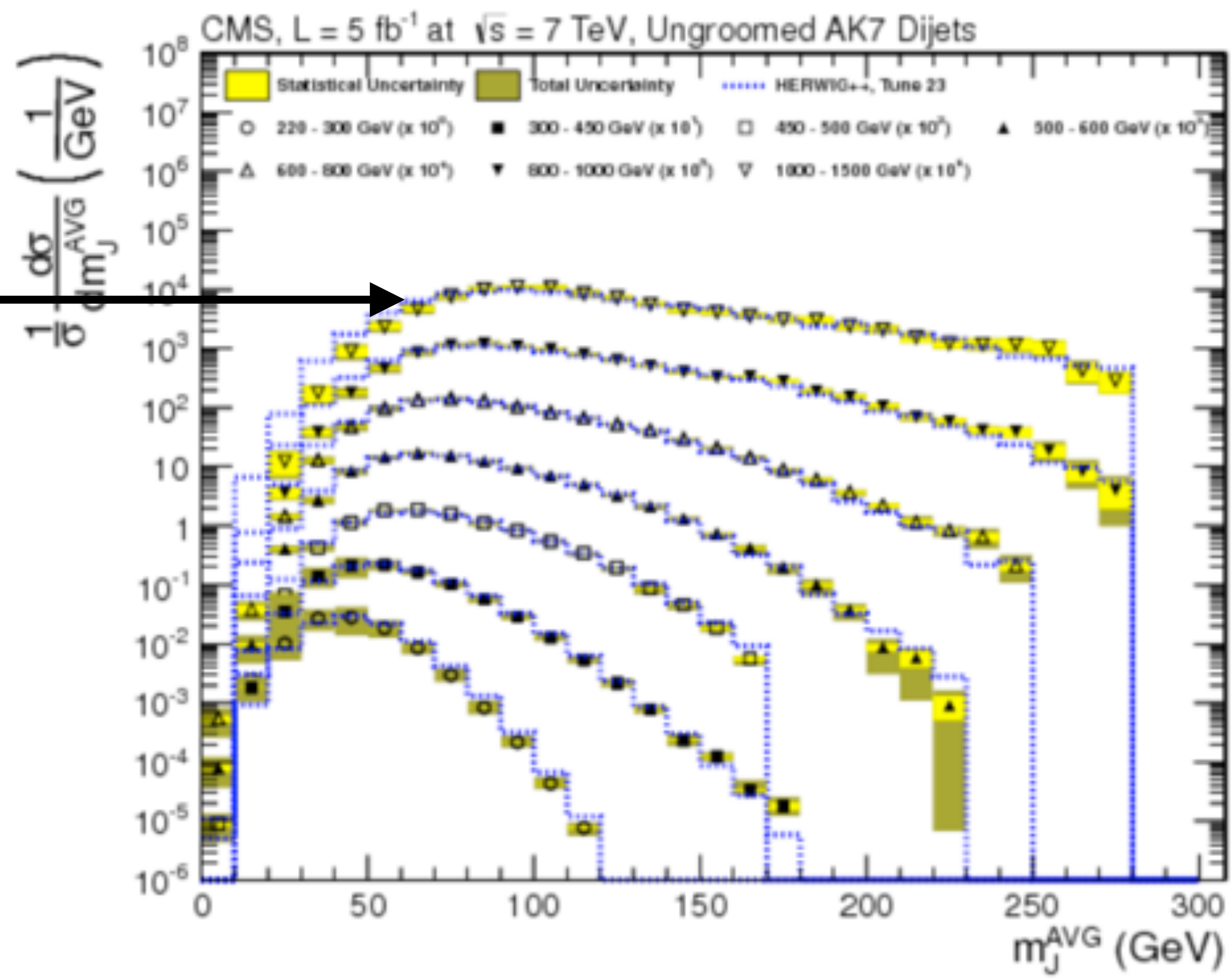
Finite-size effects from cutoff

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



Rough idea from NLO doesn't capture Sudokov peak.

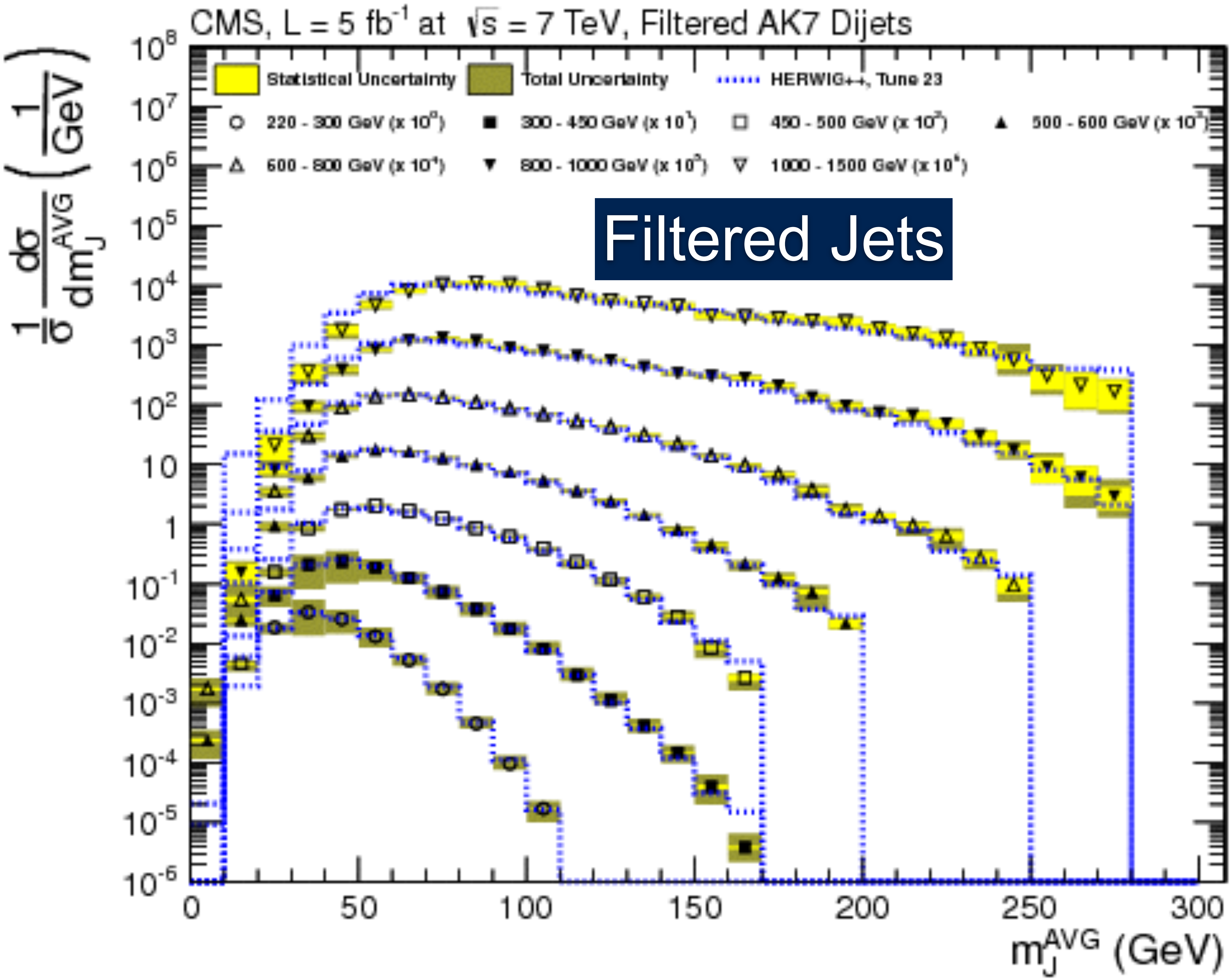
Good prediction of jet data from MC



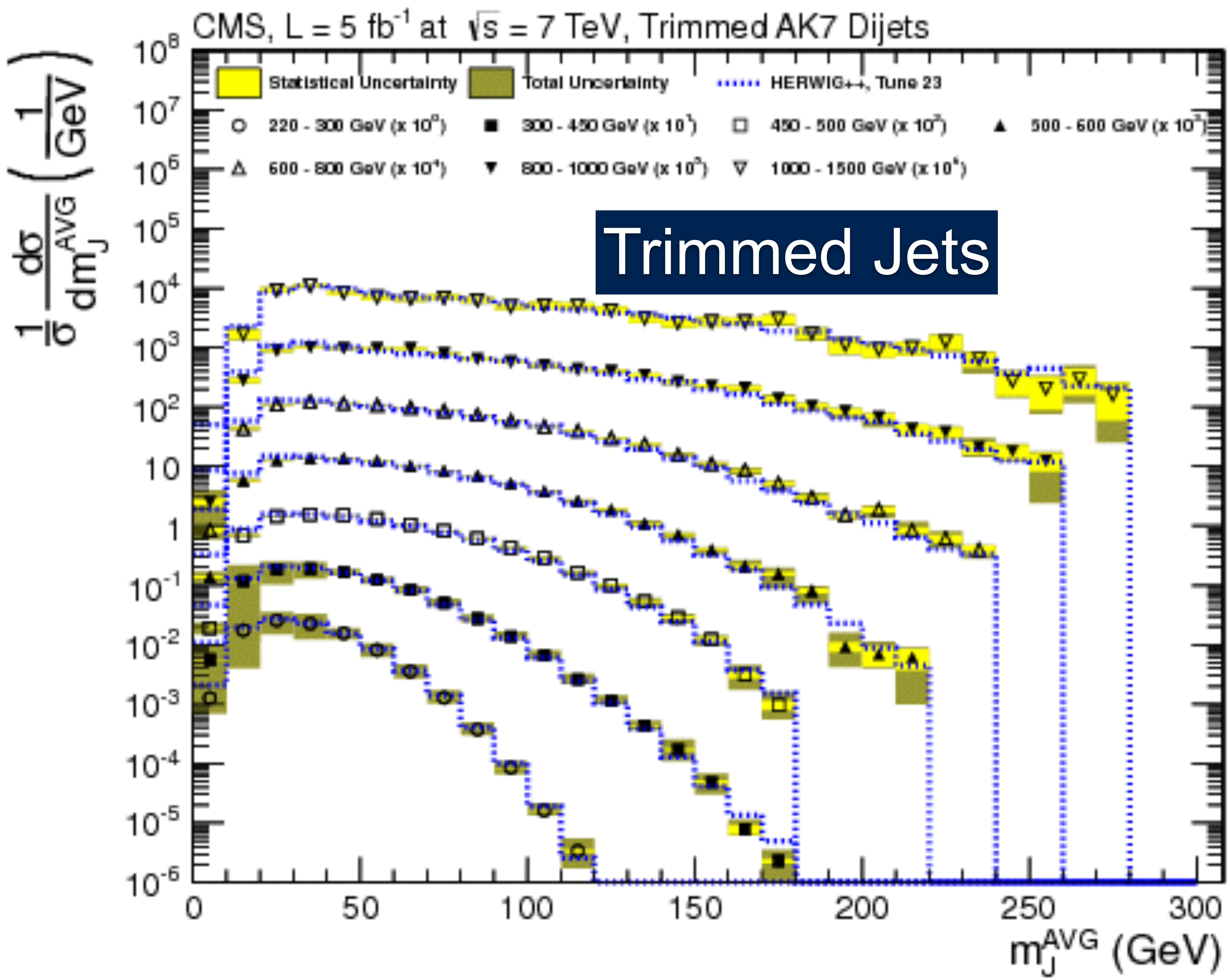
S.D. Ellis, J. Huston, K. Hatakeyama, P. Loch, M. Tonnesmann, Prog.Part.Nucl.Phys. 60 (2008) 484-551

JHEP 1305 (2013) 090

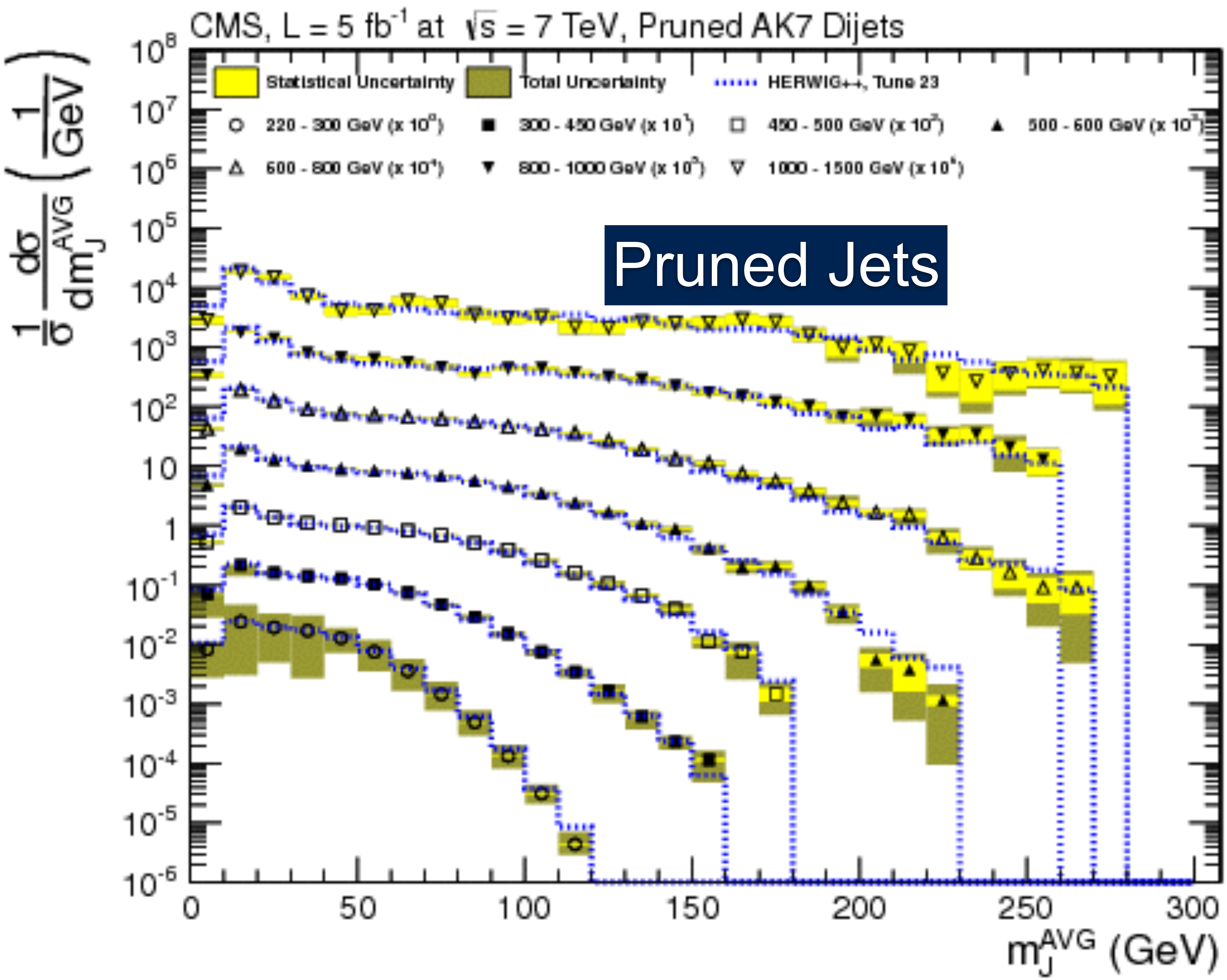
Jet Analytics



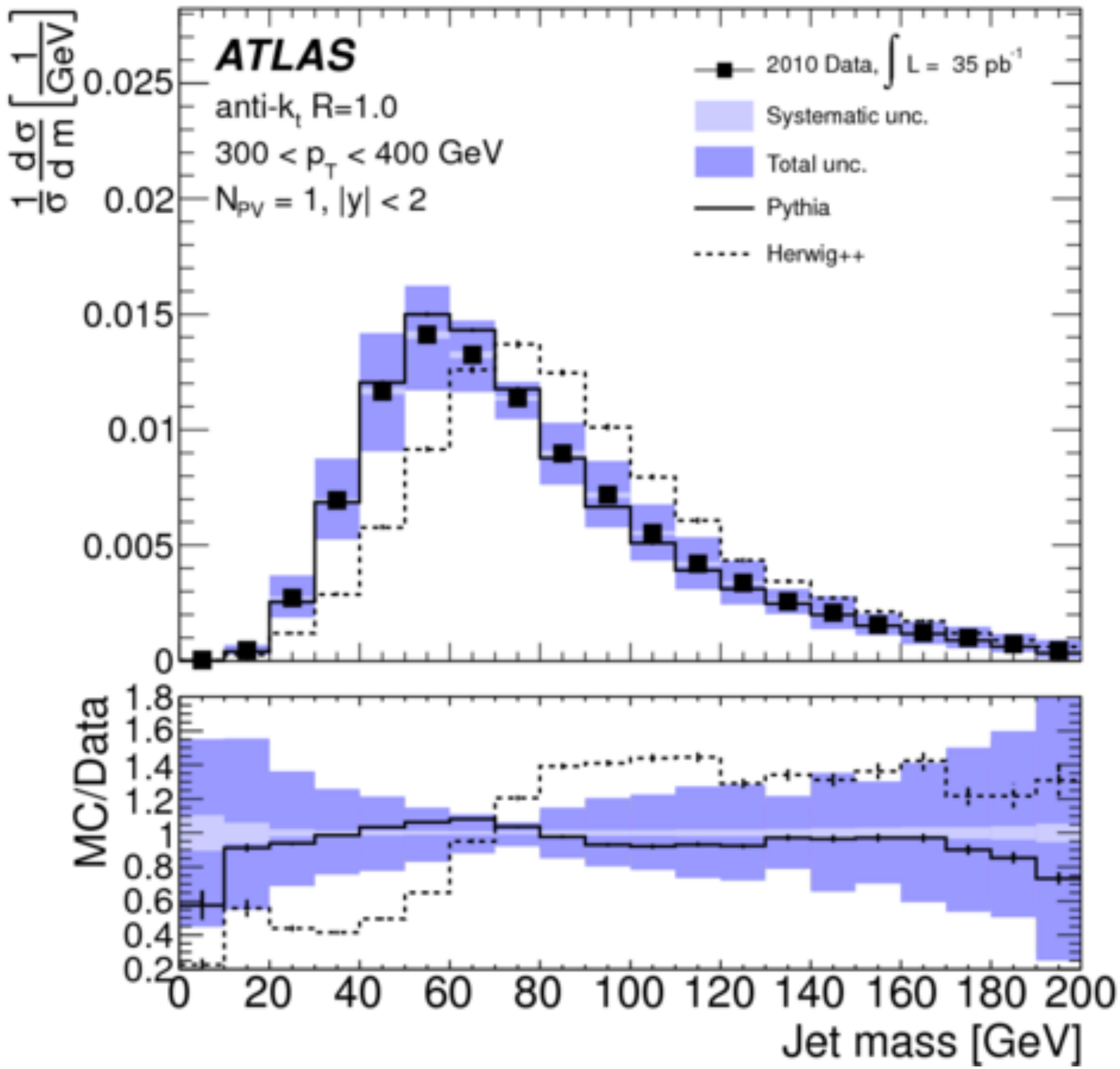
Jet Analytics



Jet Analytics



Jet Analytics



Jet Analytics

- Take-home message : QCD MC is basically getting the right answer (some better than others)
- Why is that?

Jet Analytics

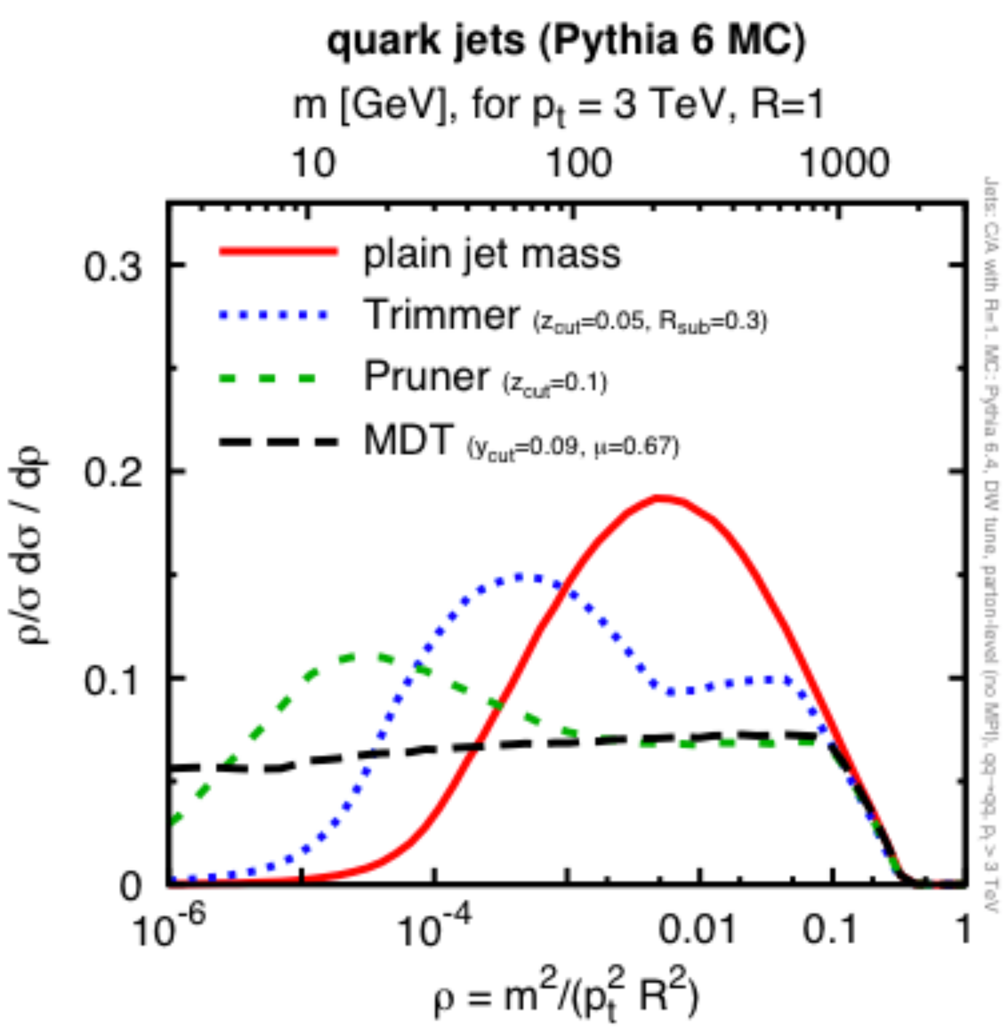
- First need to understand jet mass

At "NLL" :

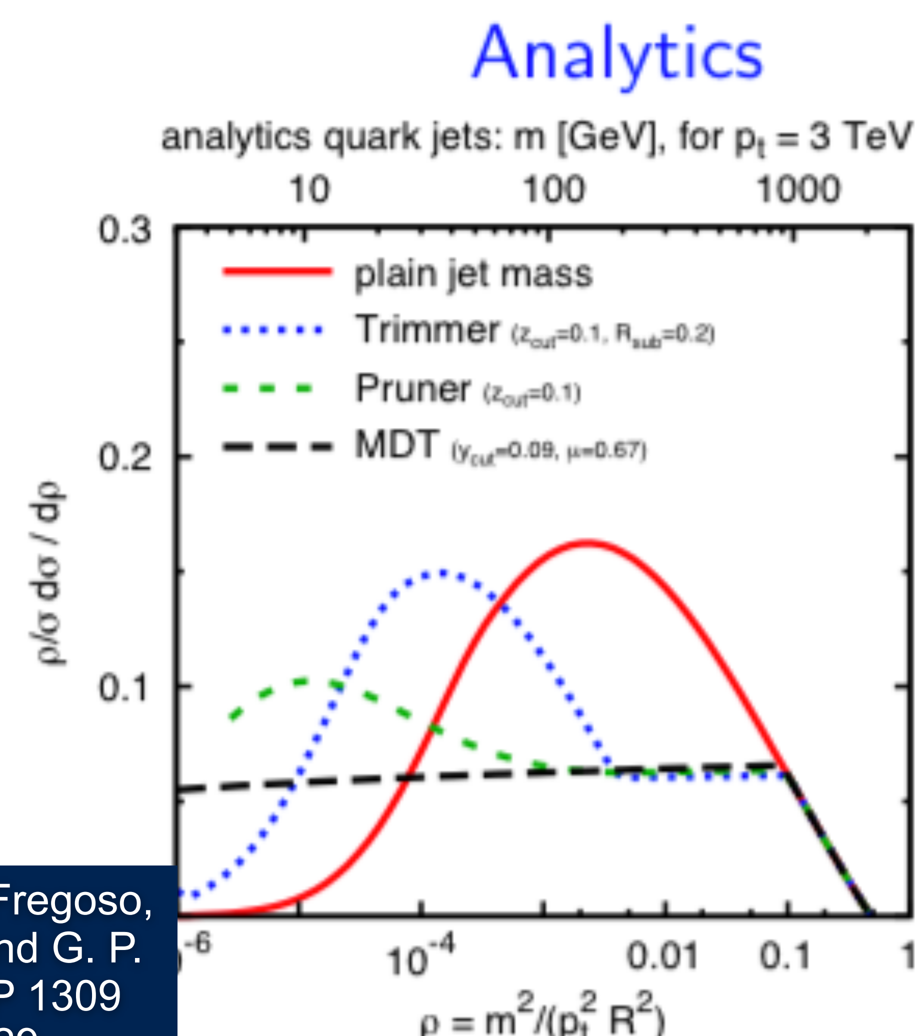
$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} \simeq \frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{\rho} - \frac{3}{4}\right) e^{-\frac{\alpha_s C_F}{2\pi}\left(\ln^2\frac{1}{\rho} - \frac{3}{2}\ln\frac{1}{\rho} + \mathcal{O}(1)\right)}$$

$$\rho \equiv \frac{m^2}{p_t^2 R^2}$$

Slide from G. Soyez



Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, JHEP 1309 (2013) 029,



Jet Analytics

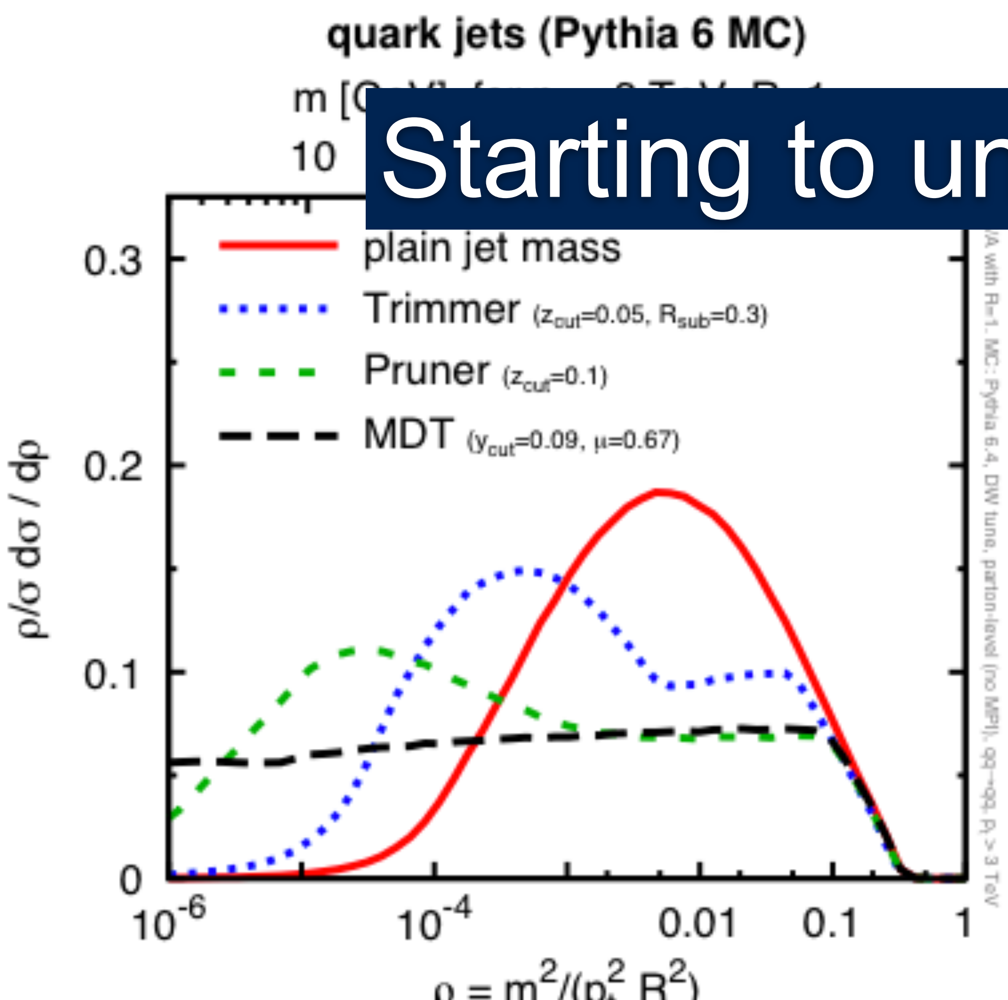
- First need to understand jet mass

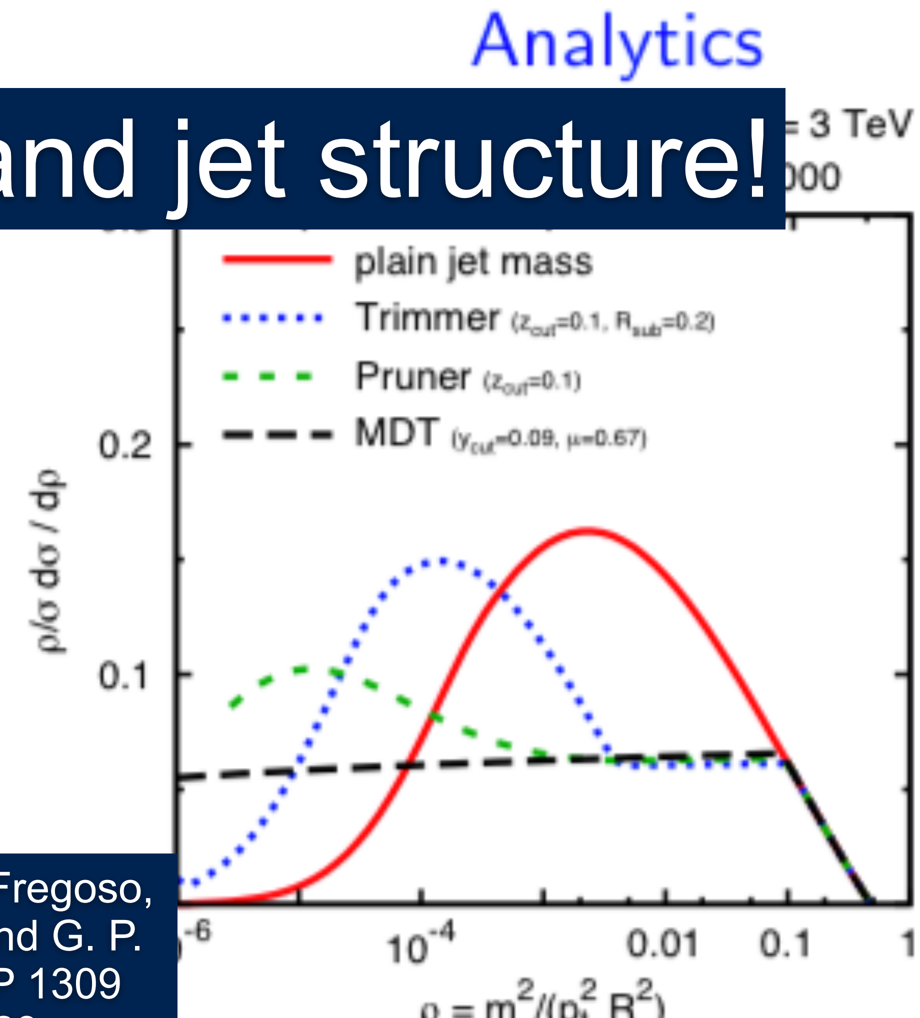
At "NLL" :

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} \simeq \frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{\rho} - \frac{3}{4}\right) e^{-\frac{\alpha_s C_F}{2\pi}\left(\ln^2\frac{1}{\rho} - \frac{3}{2}\ln\frac{1}{\rho} + \mathcal{O}(1)\right)}$$

$$\rho \equiv \frac{m^2}{p_t^2 R^2}$$

Slide from G. Soyez



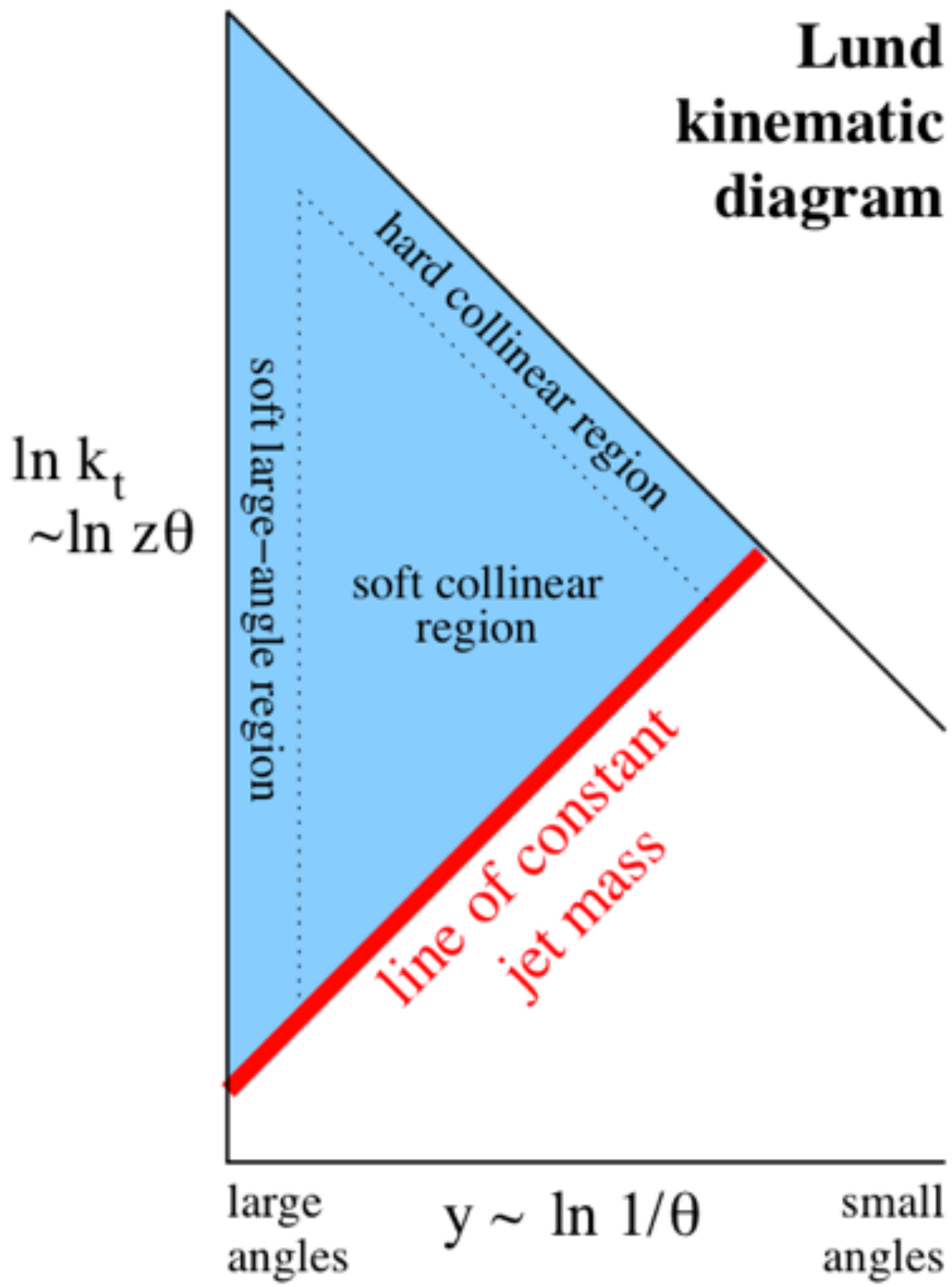


Starting to understand jet structure!

Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, JHEP 1309 (2013) 029,

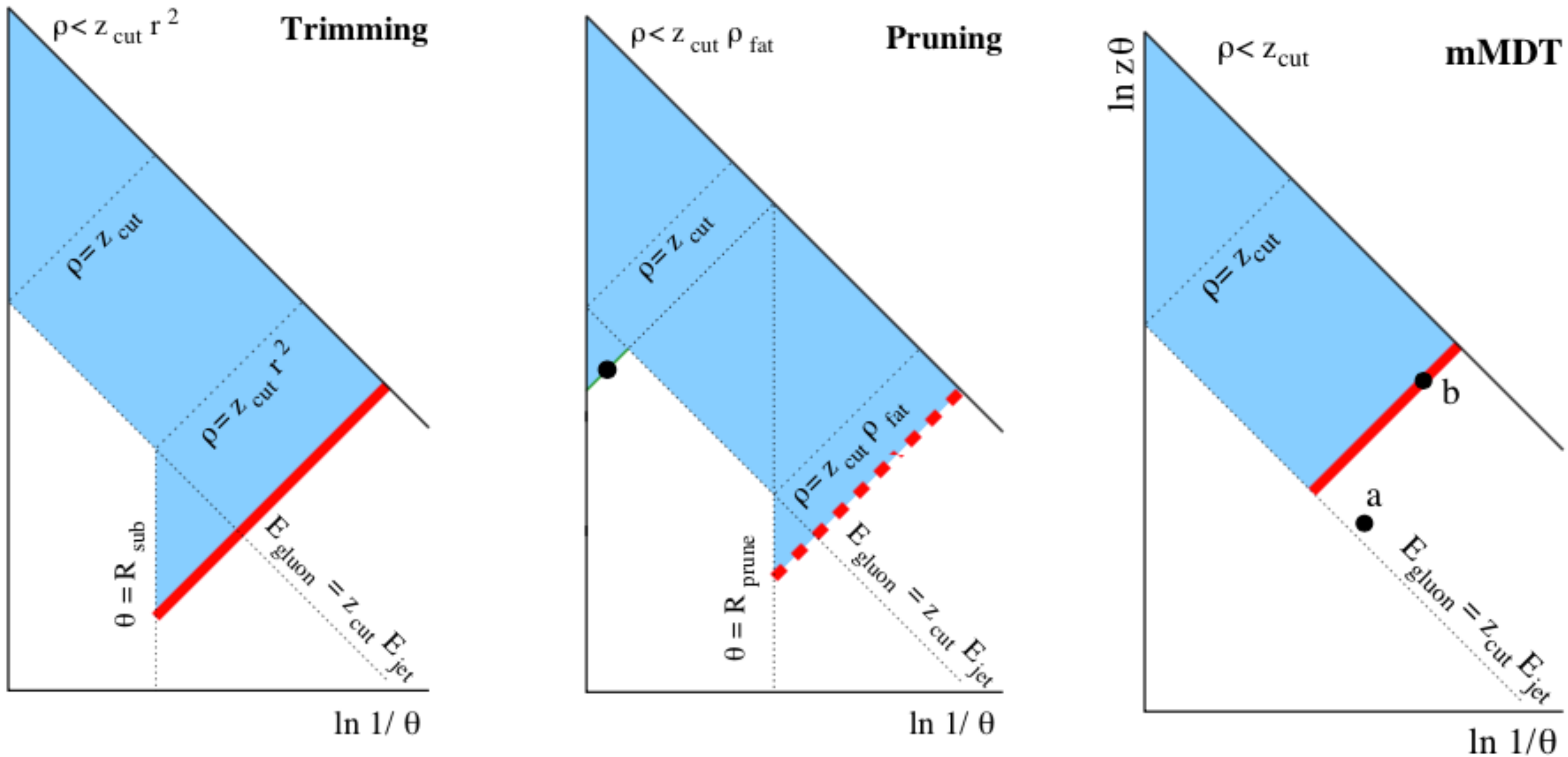
Jet Grooming Analytics

- What are groomers doing?



M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, JHEP 1309 (2013) 029,

- What are groomers doing?

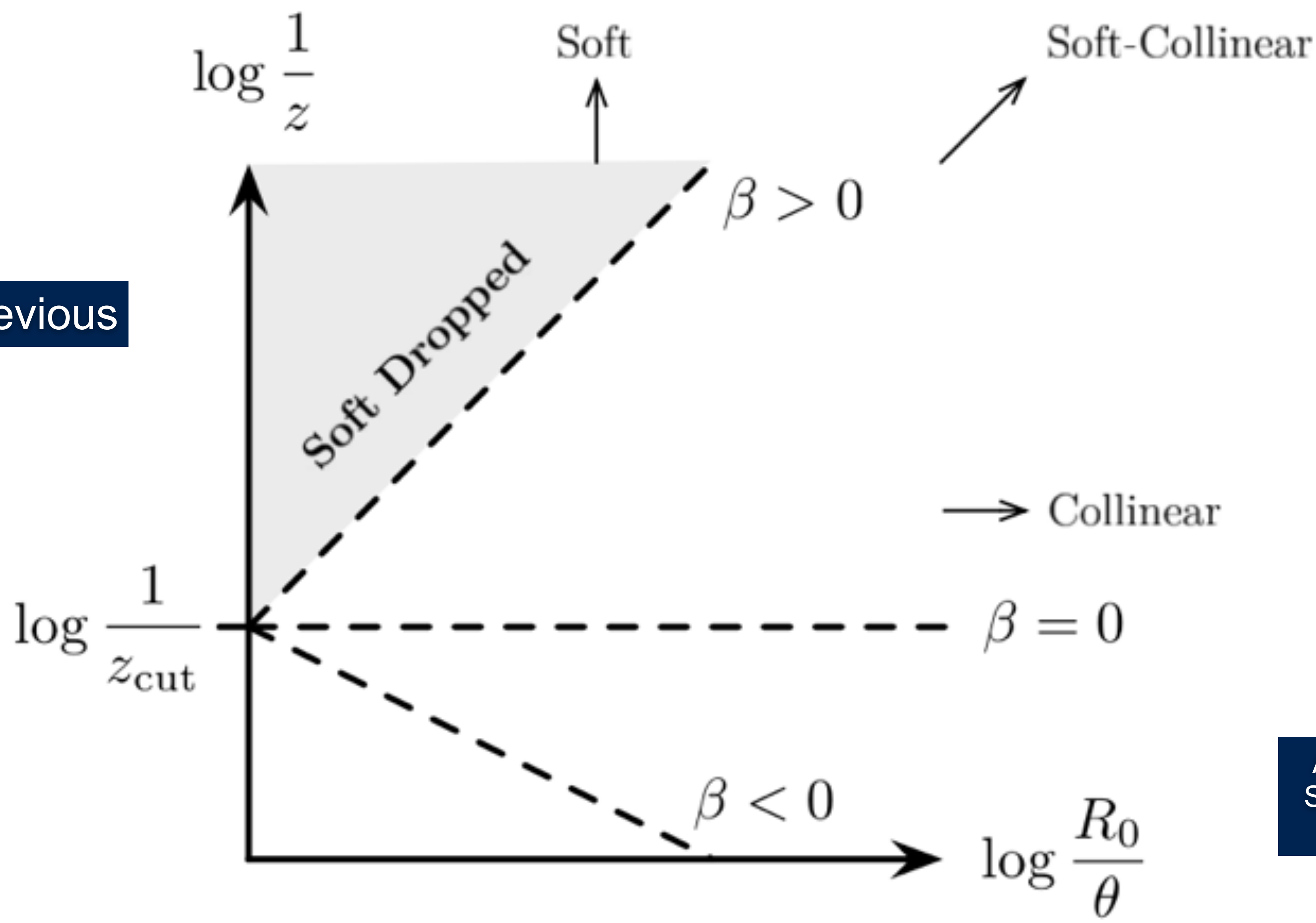


M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, JHEP 1309 (2013) 029,

Jet Grooming Analytics

- Understanding gained from jet analytics even gives new and better ways to groom and tag!

Note : y-axis now 1/previous



A. Larkoski, S. Marzani, G. Soyez, J. Thaler,JHEP 1405 (2014) 146

Soft drop : “simple” behavior in this plane, with tunable parameter for many algorithms!

Jet Grooming Analytics

- Soft drop :
  - Undo last stage of C/A clustering, label subjets j1,j2
  - If :

$$\frac{min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta}$$

  then j is soft dropped
  else redefine j to be the harder, and iterate

  - Recovers (modified) mass drop BDRS tagger for beta=0
    - This case always removes soft radiation entirely (hence the name)

Substructure Variables

- A plethora of variables to choose from :
  - N-subjettiness

$$\tau_N^{(\beta)} = \sum_i p_{Ti} \min\left\{R_{1,i}^{\beta}, R_{2,i}^{\beta}, \ldots, R_{N,i}^{\beta}\right\}$$

  - Energy correlation function

$$\mathrm{ECF}(N,\beta) = \sum_{i_1<i_2<\ldots<i_N\in J} \left(\prod_{a=1}^{N} E_{i_a}\right)\left(\prod_{b=1}^{N-1}\prod_{c=b+1}^{N}\theta_{i_b i_c}\right)^{\beta}$$

  - Mass drop (mass of heaviest subjet over mass of jet)
  - Subjet momentum balance (or subjet asymmetry)

$$\sqrt{y} \equiv \min(p_{\mathrm{T}\,j_1}, p_{\mathrm{T}\,j_2}) \frac{\Delta R_{(j_1,j_2)}}{m_0}$$

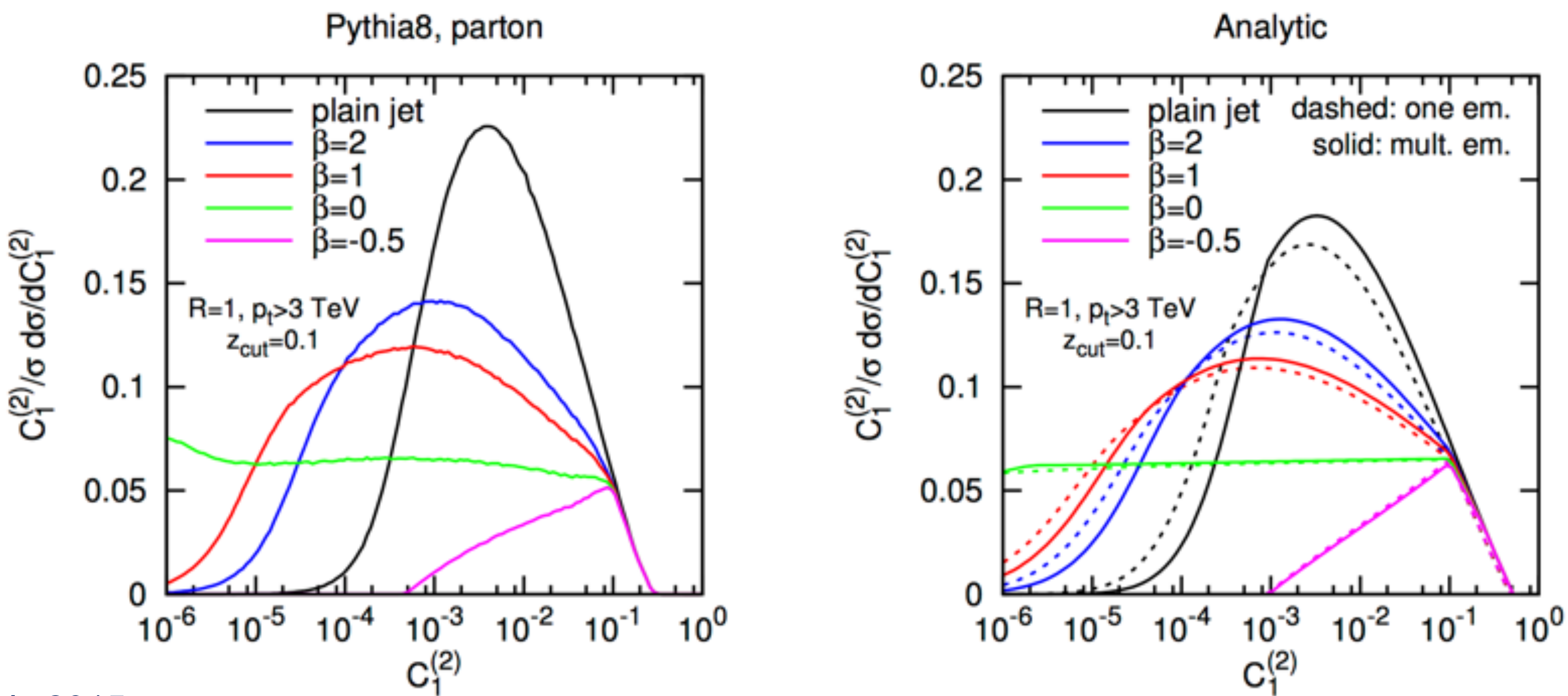
Substructure Variables

- Can also look into n-subjettiness, energy correlation functions

$$C_1^{(\alpha)} = \frac{\text{ECF}(2,\alpha)\,\text{ECF}(0,\alpha)}{\text{ECF}(1,\alpha)^2},$$

A. Larkoski, S. Marzani, G. Soyez, J. Thaler,JHEP 1405 (2014) 146

$$\begin{aligned}
\text{ECF}(0,\alpha) &= 1, \\
\text{ECF}(1,\alpha) &= \sum_{i\in \text{jet}} p_{Ti}, \\
\text{ECF}(2,\alpha) &= \sum_{i<j\,\in \text{jet}} p_{Ti}\,p_{Tj} \left(\frac{\Delta R_{ij}}{R_0}\right)^{\alpha}.
\end{aligned}$$

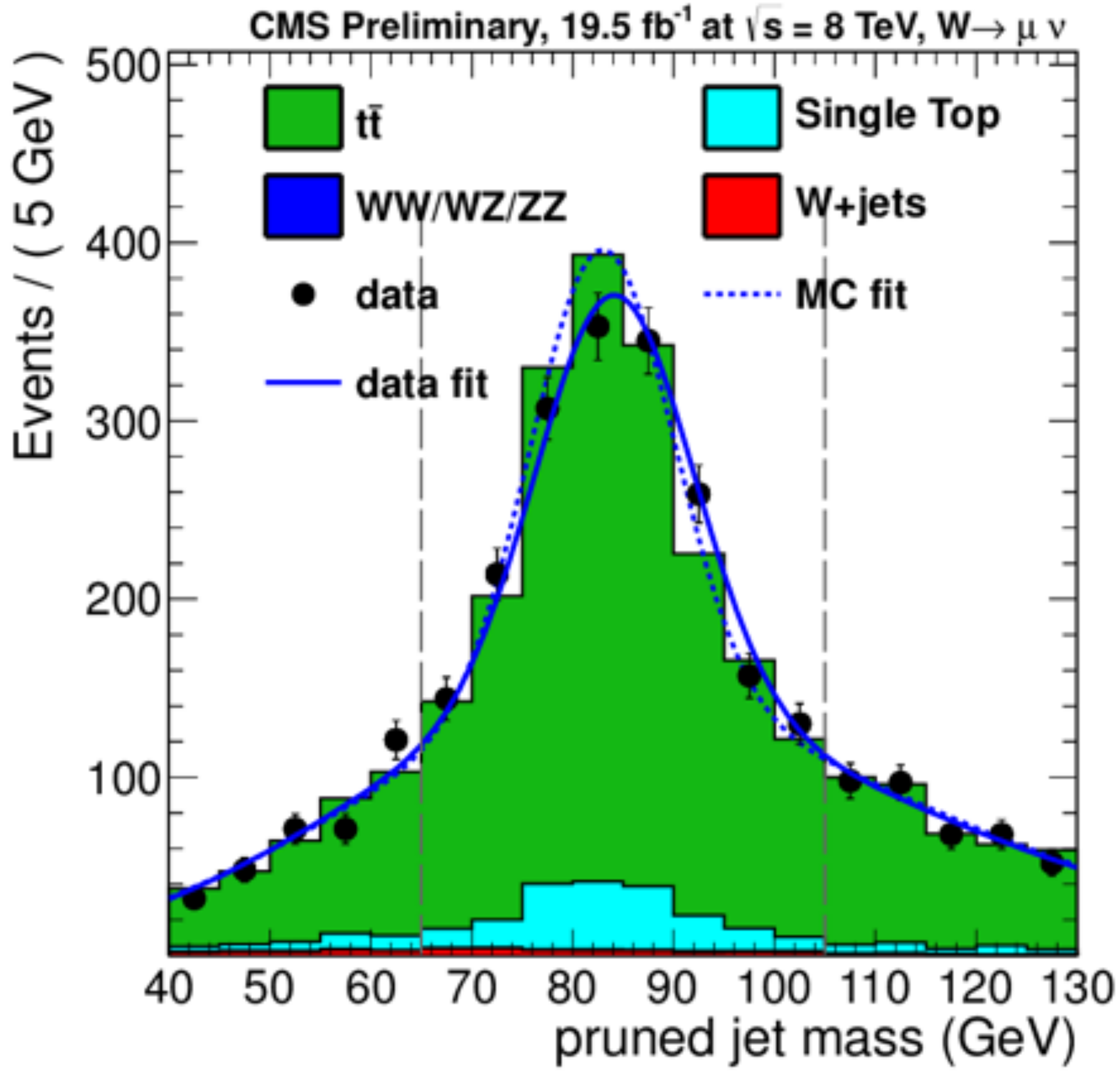


Outline

- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

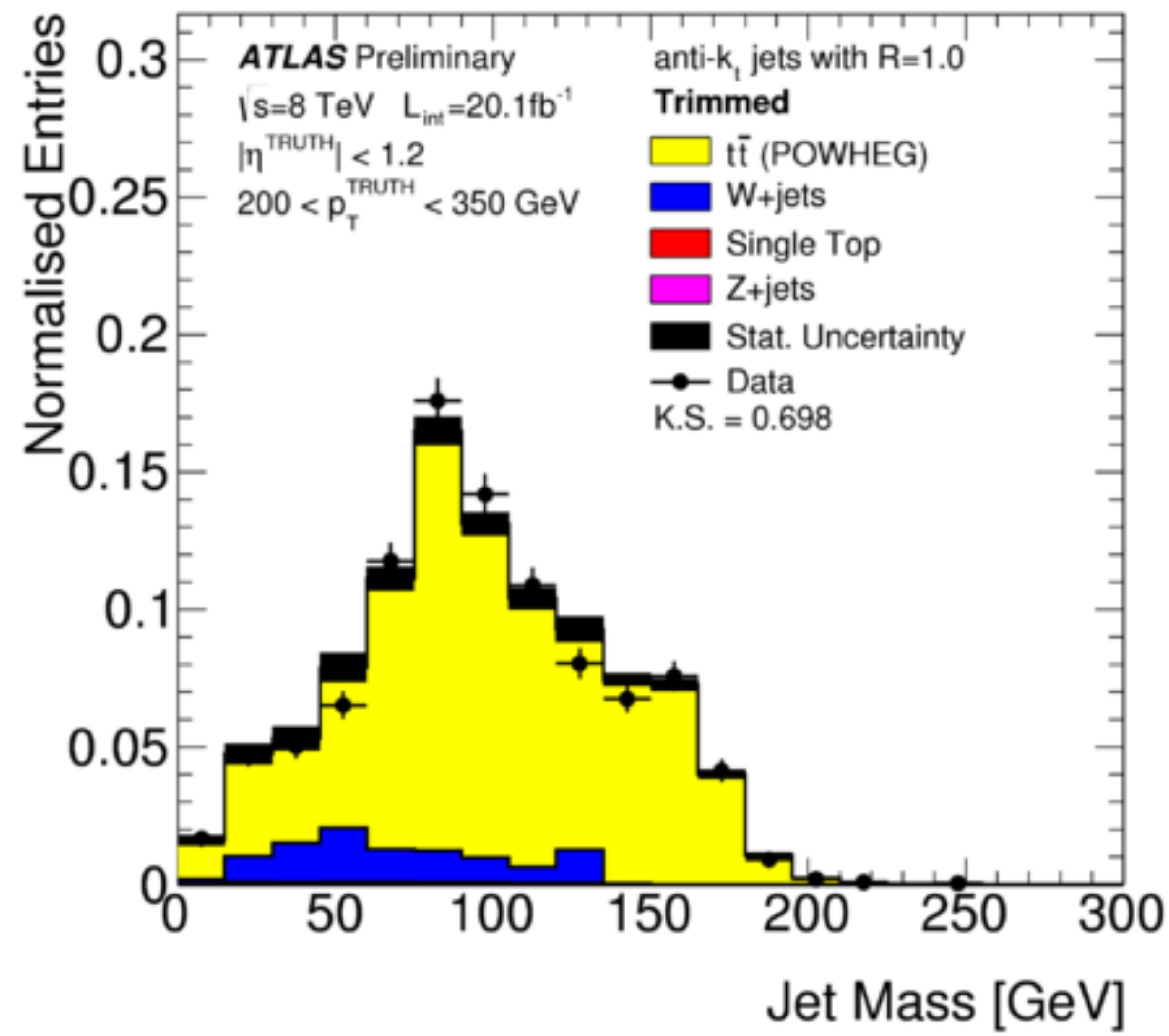
W/Z/H Tagging

- CMS : pruned jet mass window and 2-jettiness



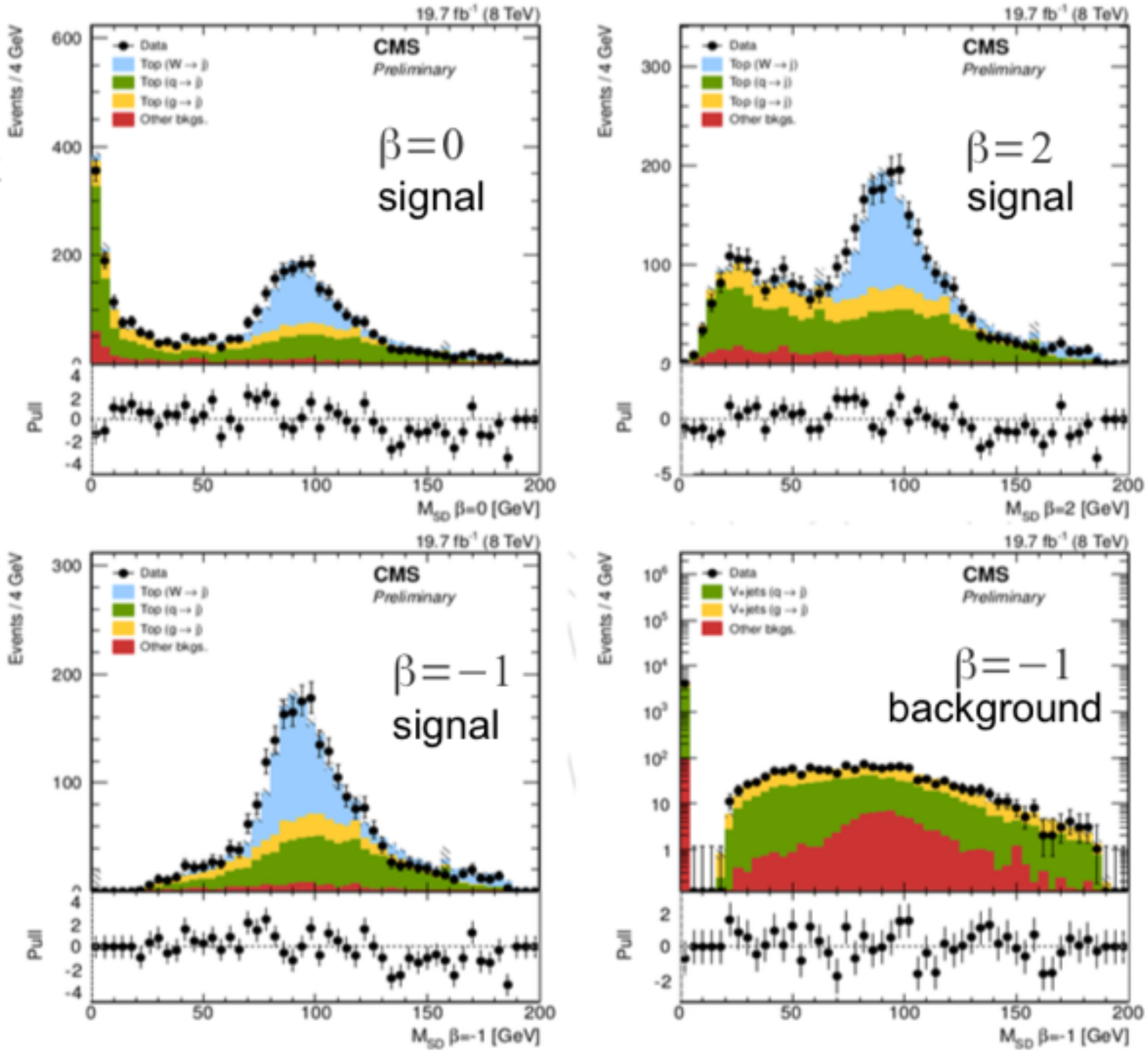
CMS-JME-13-006

- ATLAS : filtered jet mass window and asymmetry cut



ATL-PHYS-PUB-2014-004

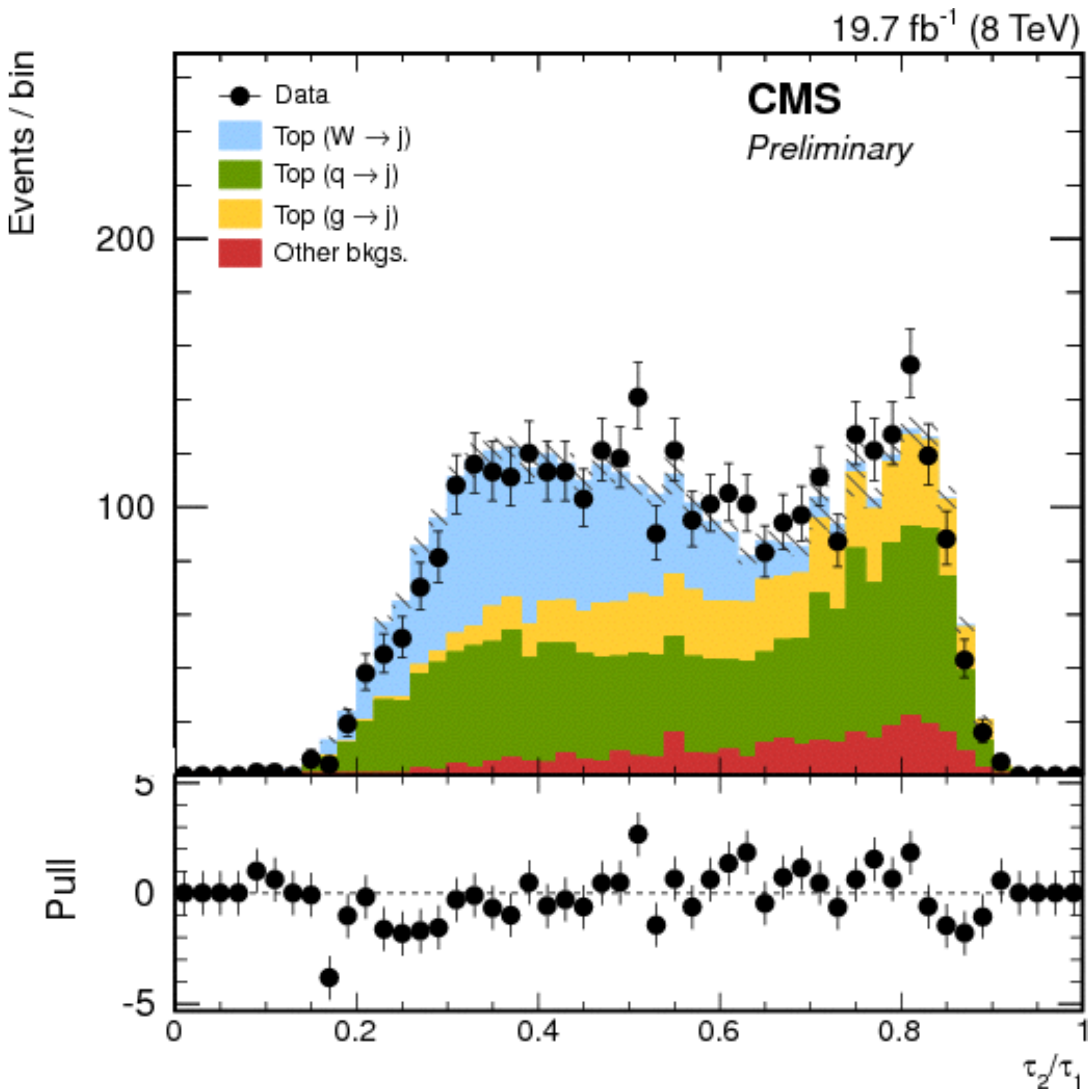
- CMS moving to soft drop instead of pruning

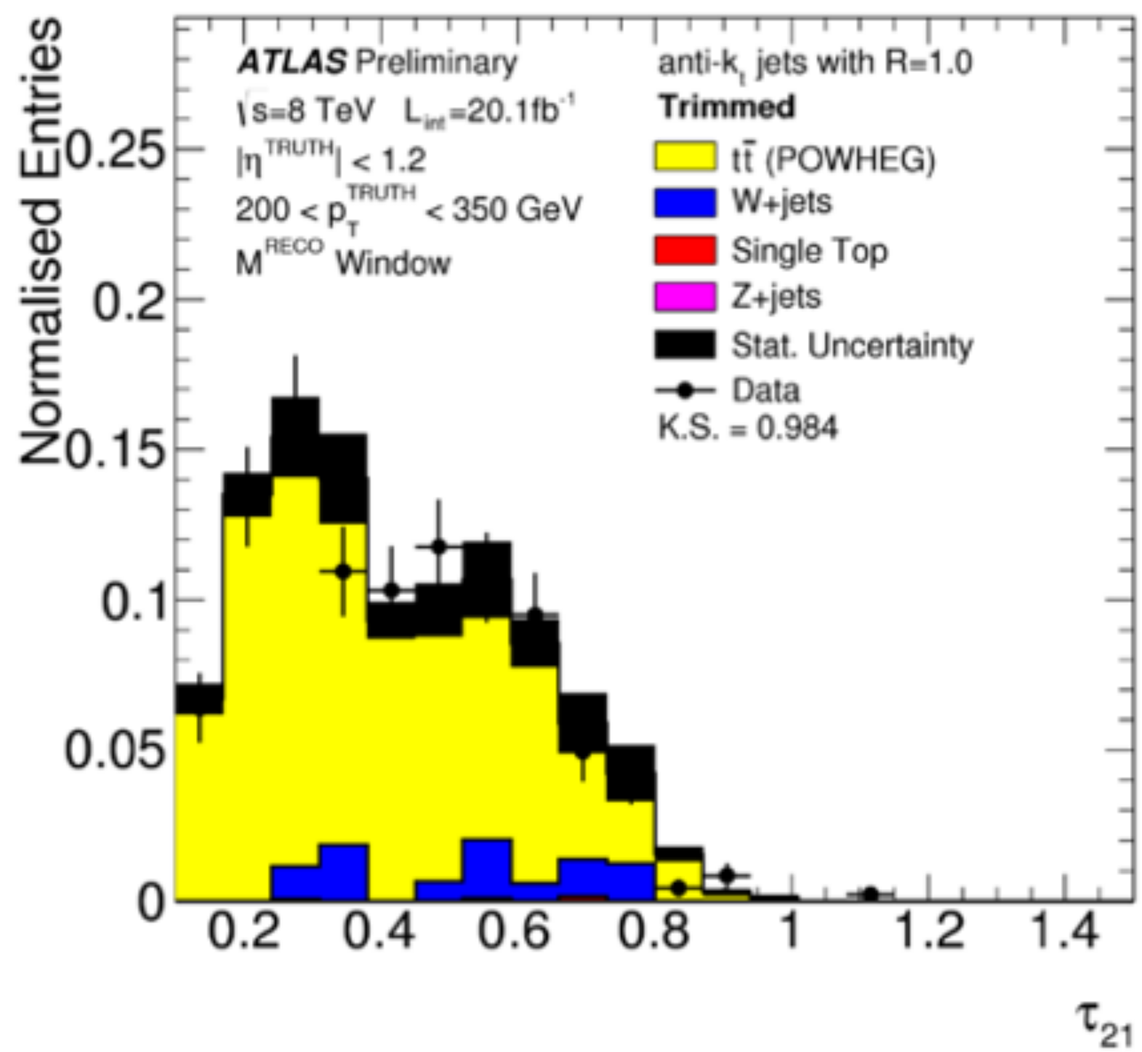


CMS-JME-14-002

W/Z/H Tagging

N-subjettiness





CMS-JME-14-002

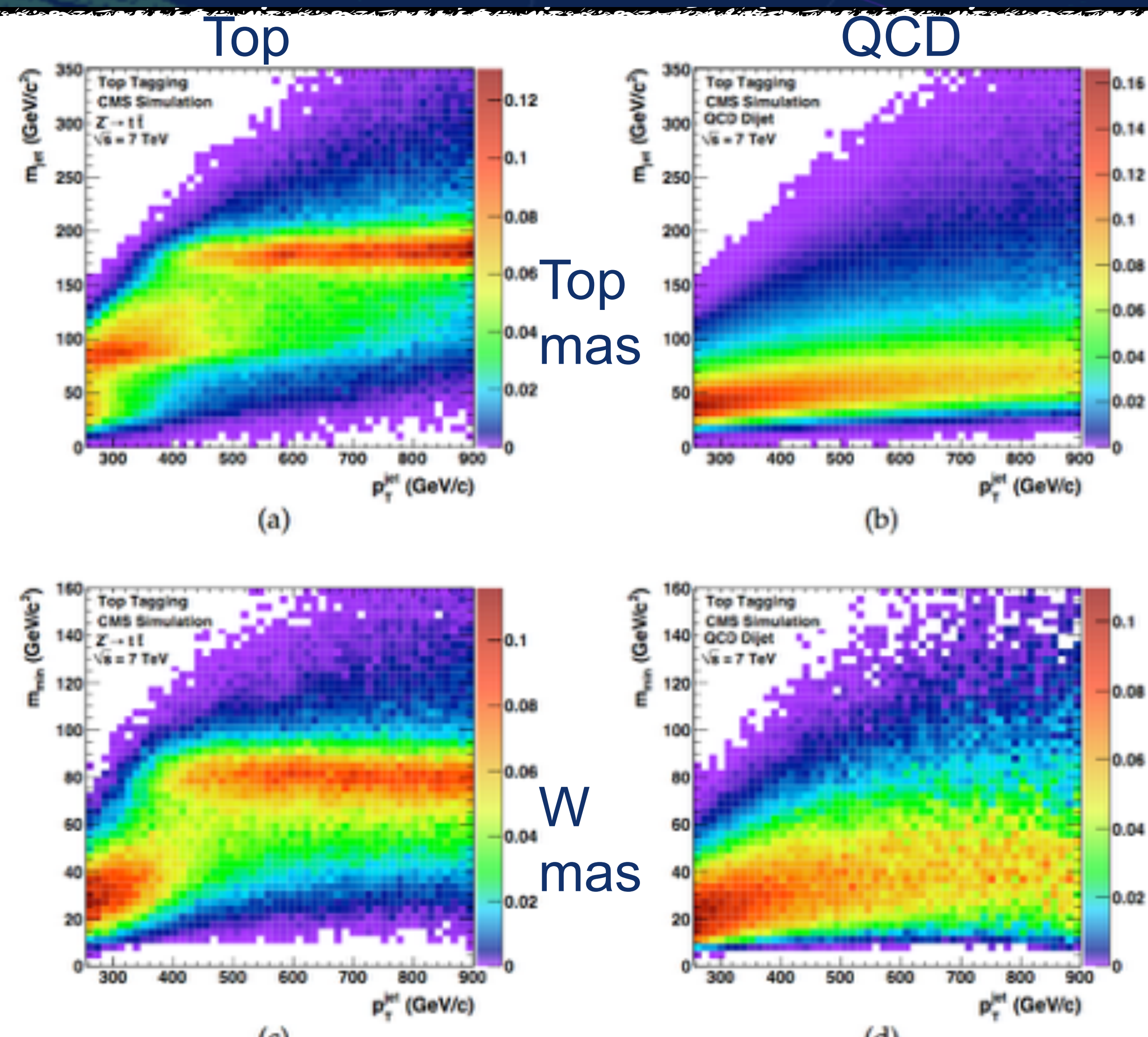
ATL-PHYS-PUB-2014-004

Outline

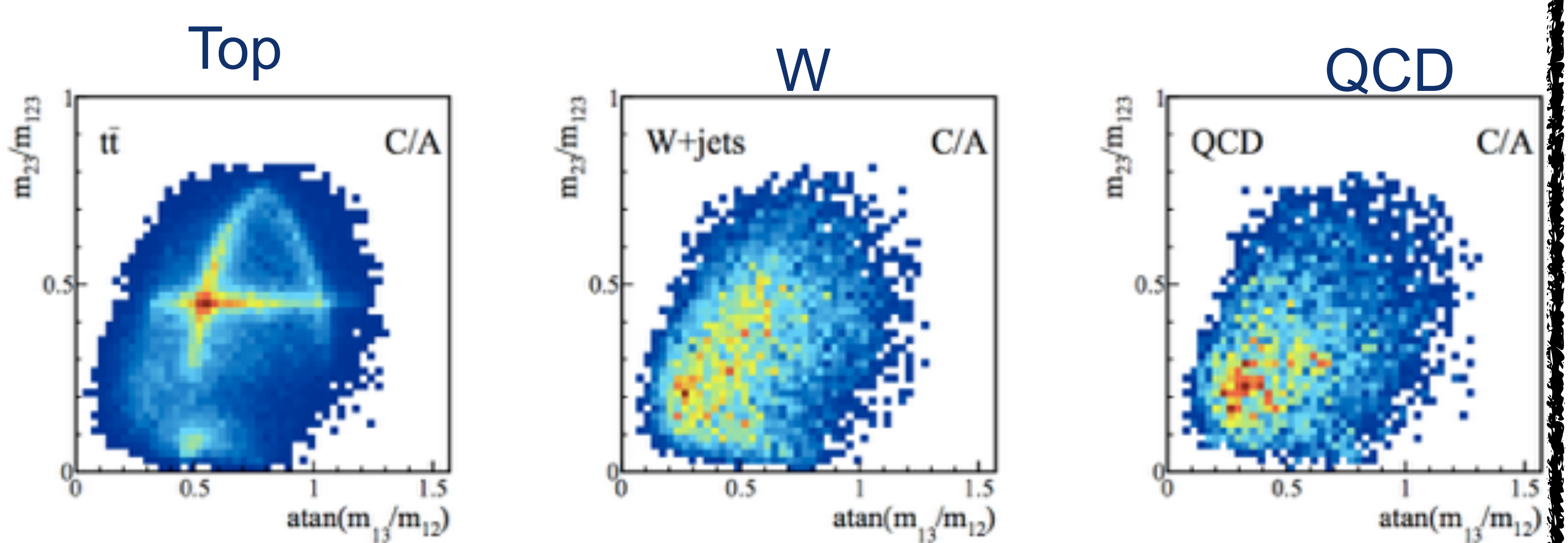
- Motivation
- Jet Algorithms
- Substructure
- Analytics
- W/Z/H taggers
- Top quark taggers

Top Tagging

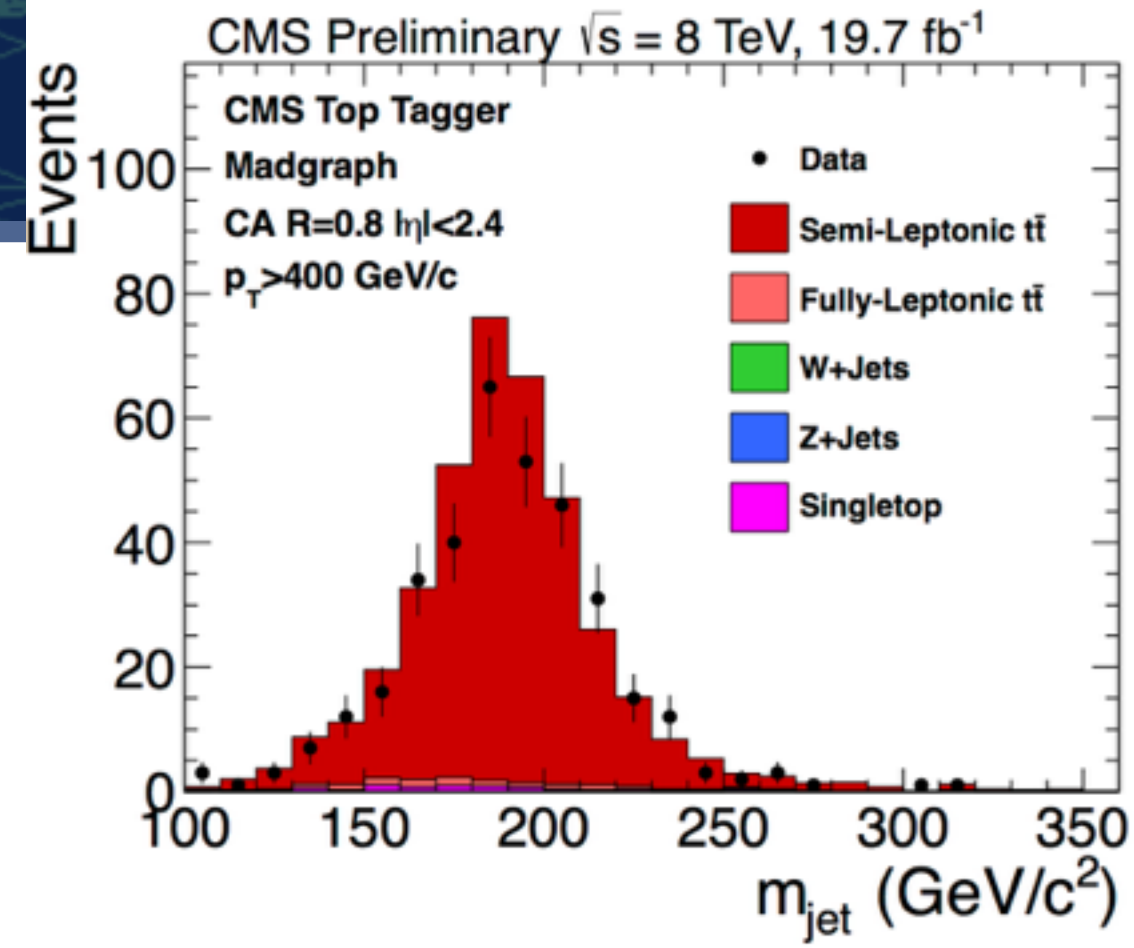
- JHU / CMS top tagger
  - Phys.Rev.Lett. 101 (2008) 142001
  - Break up cluster sequence to get three or four subjets
  - Impose top and W mass
    - Top mass ~ jet mass
    - W mass ~ min pairwise mass

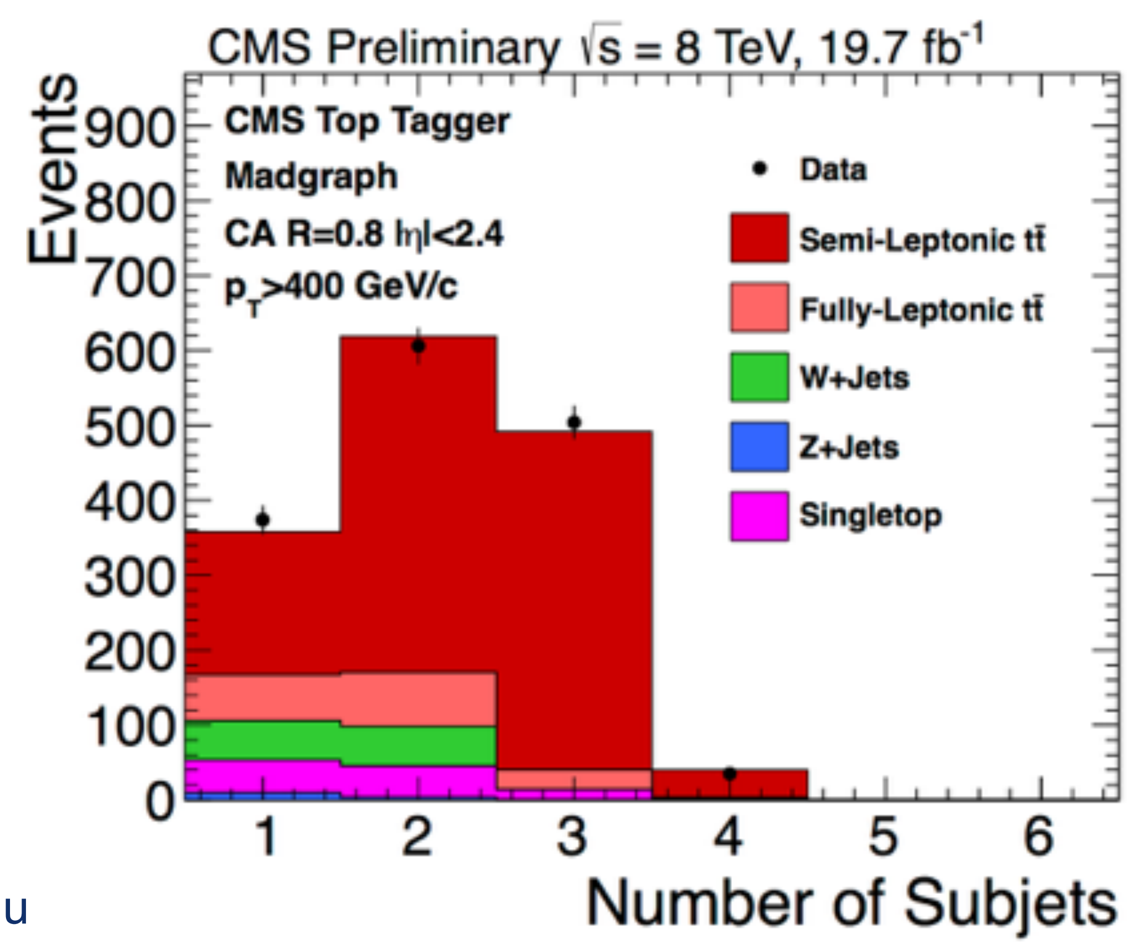


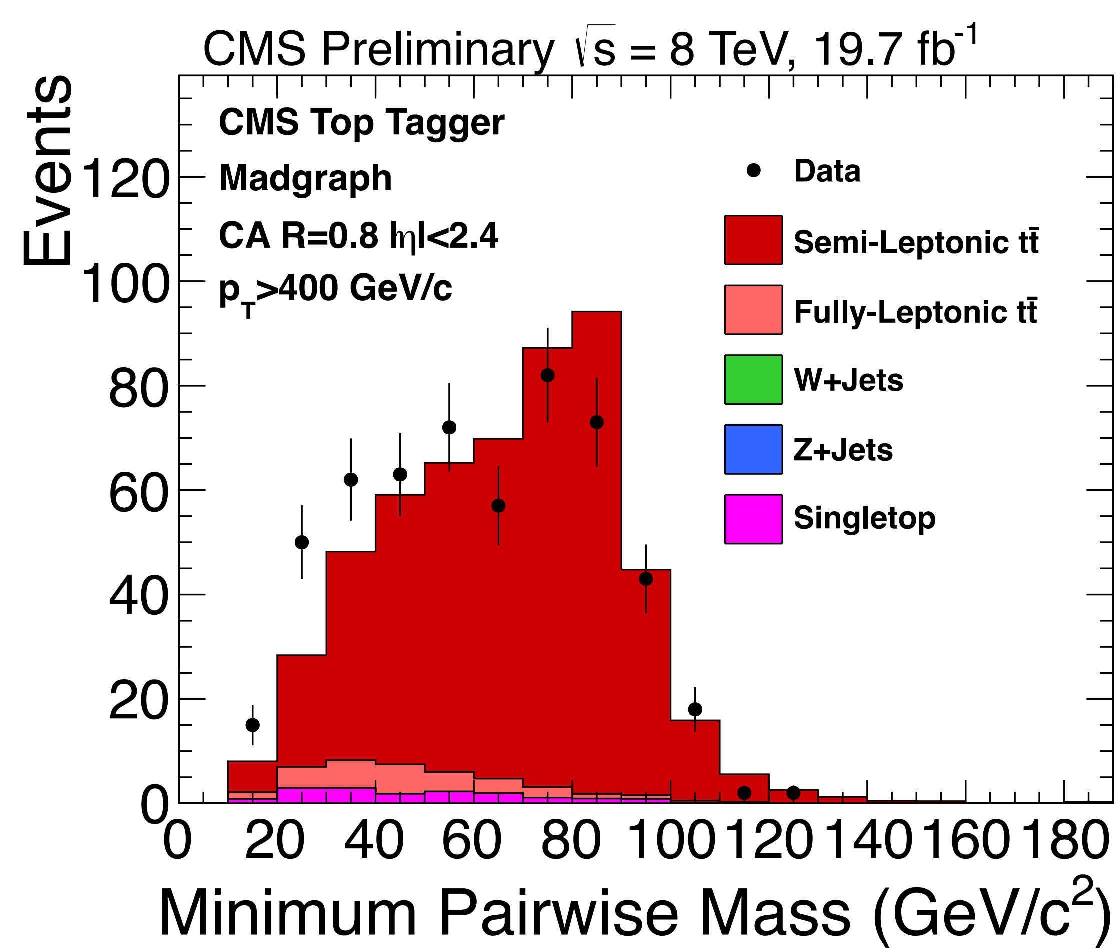
- HEP top tagger
  - Break up cluster sequence to get three or four subjets
  - Impose “Dalitz-like” cuts



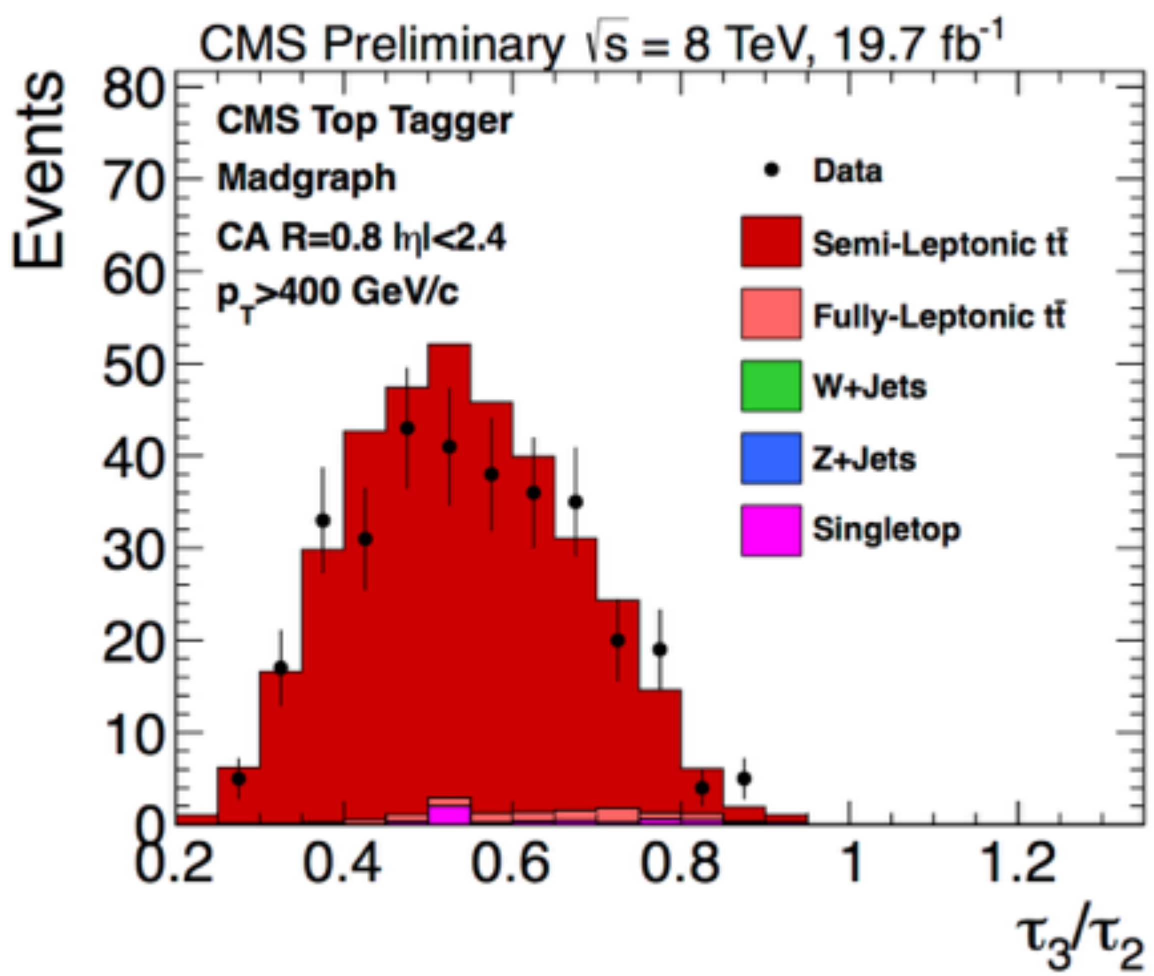
Top Tagging : JHU/CMS Tagger

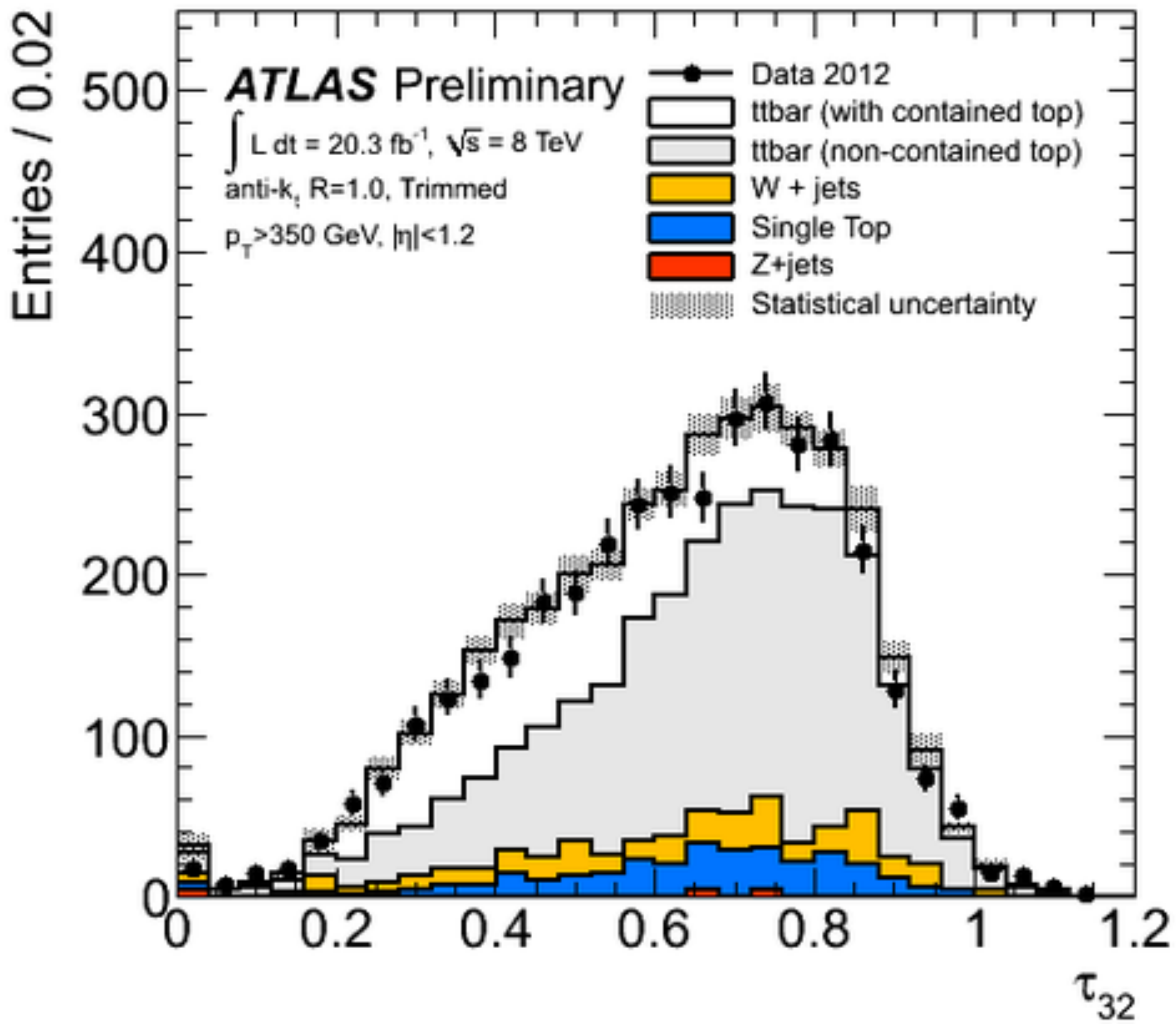




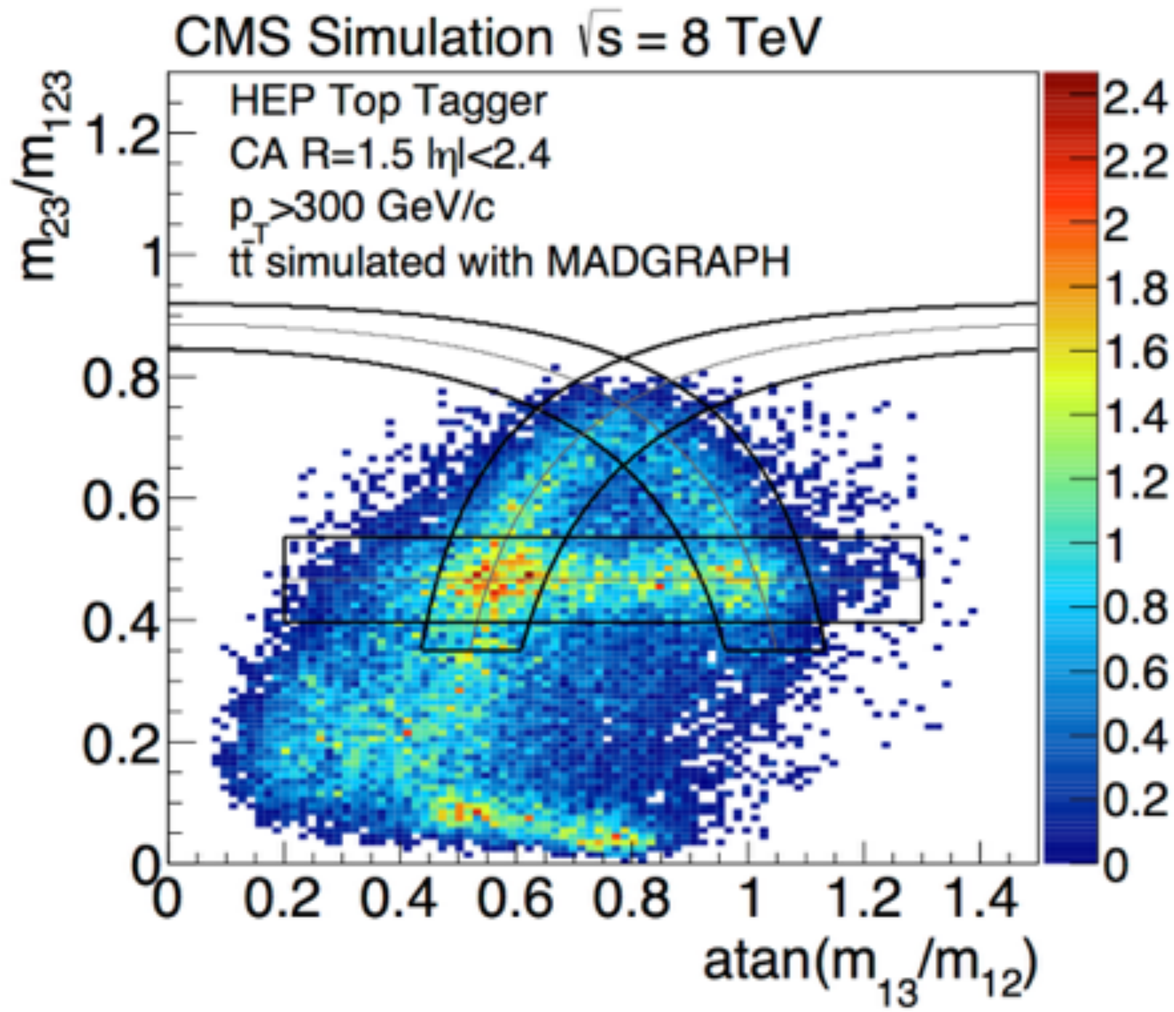


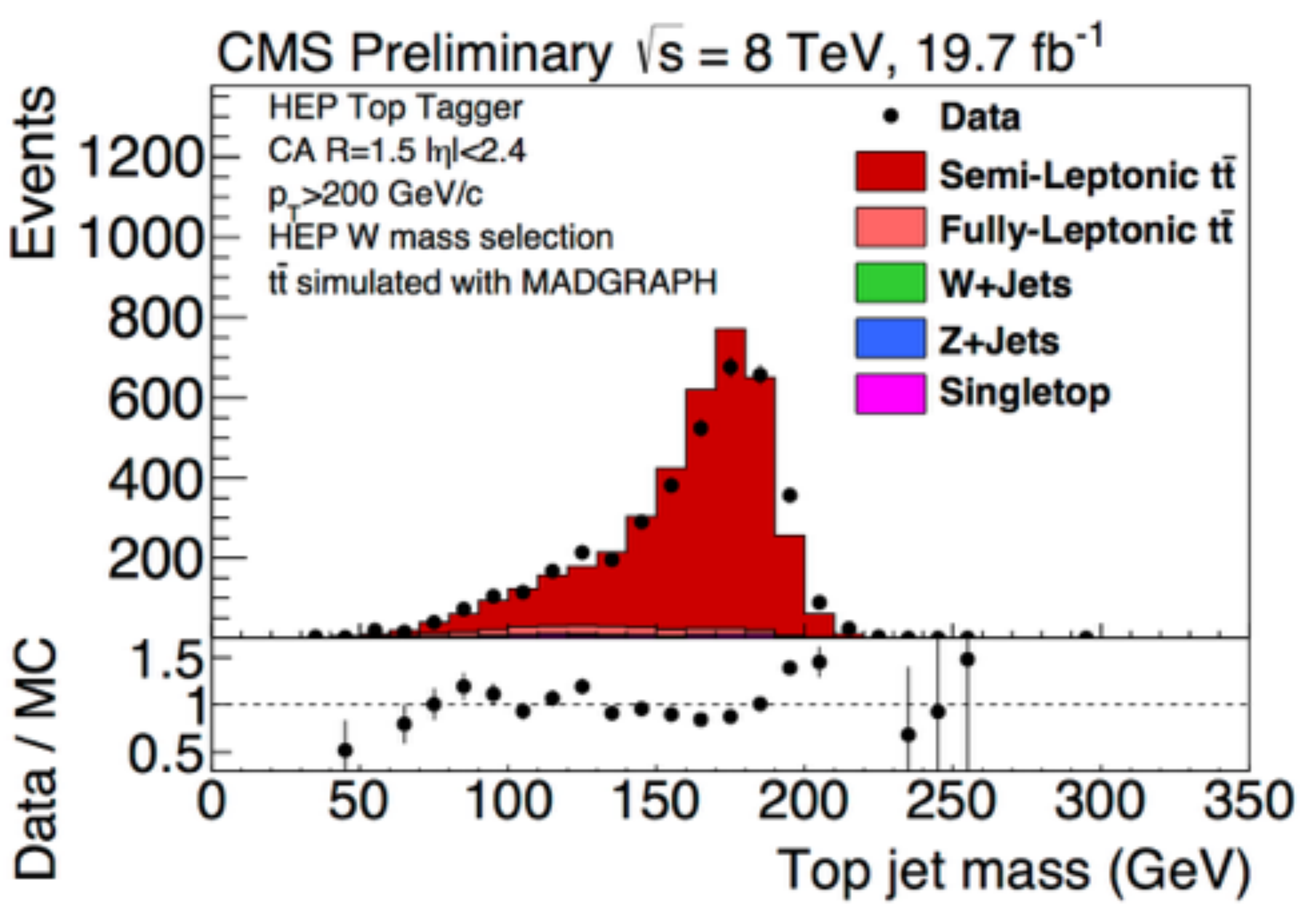
N-subjettiness

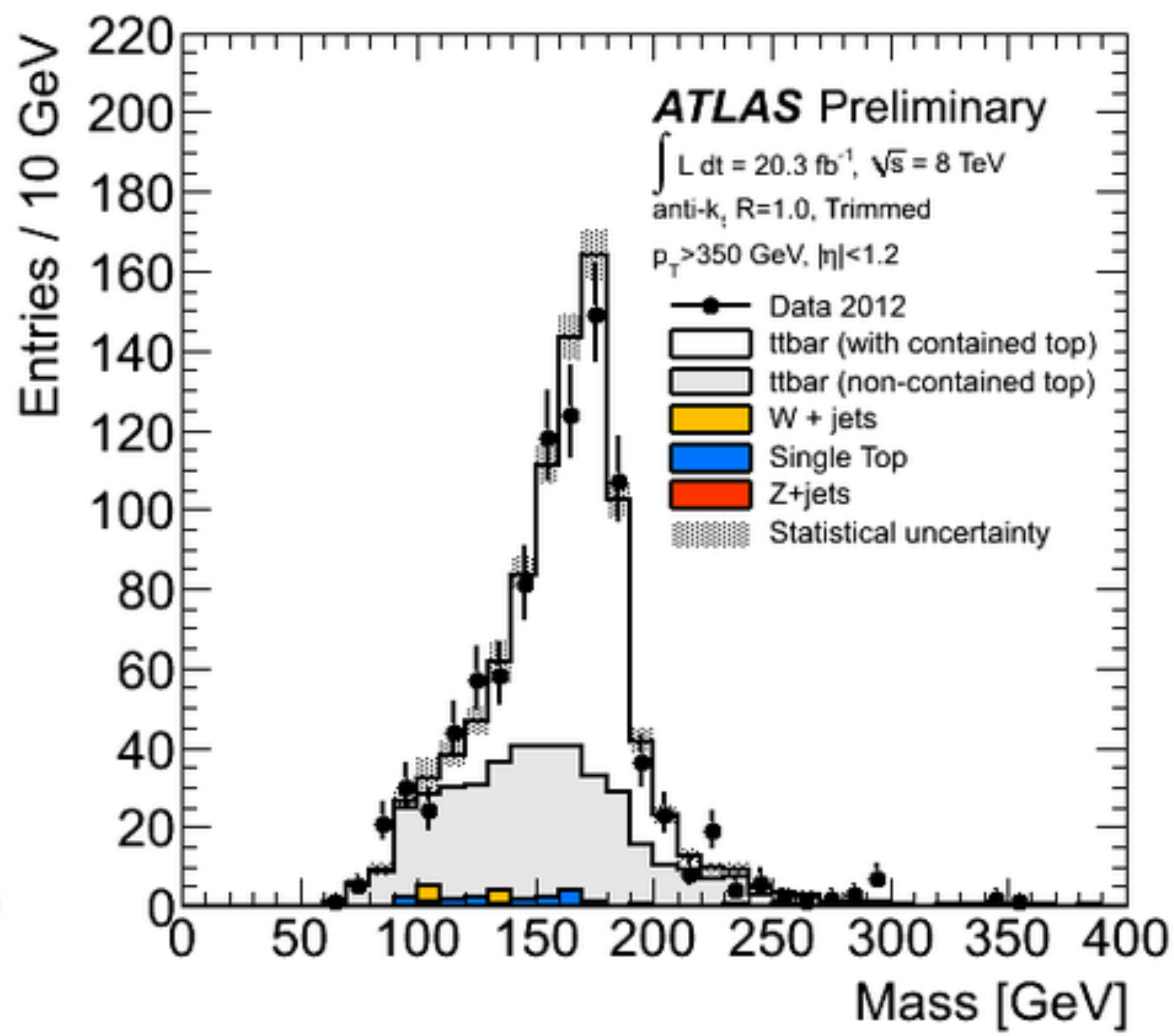




HEP Top Tagger

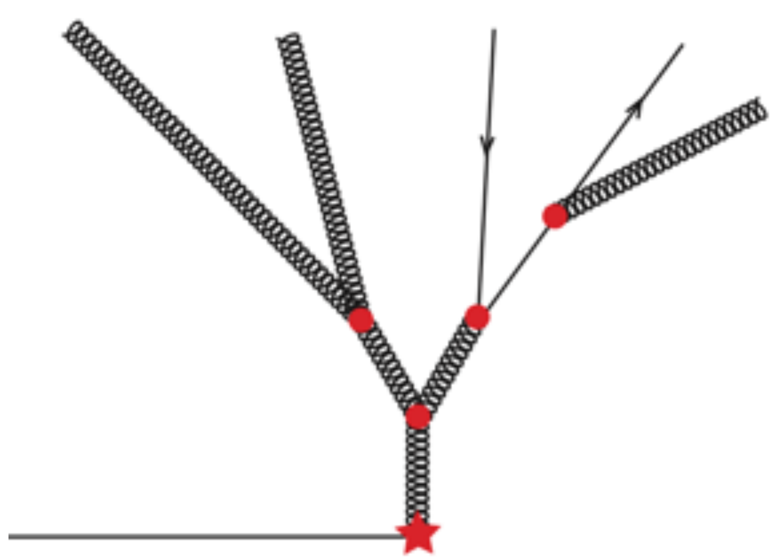


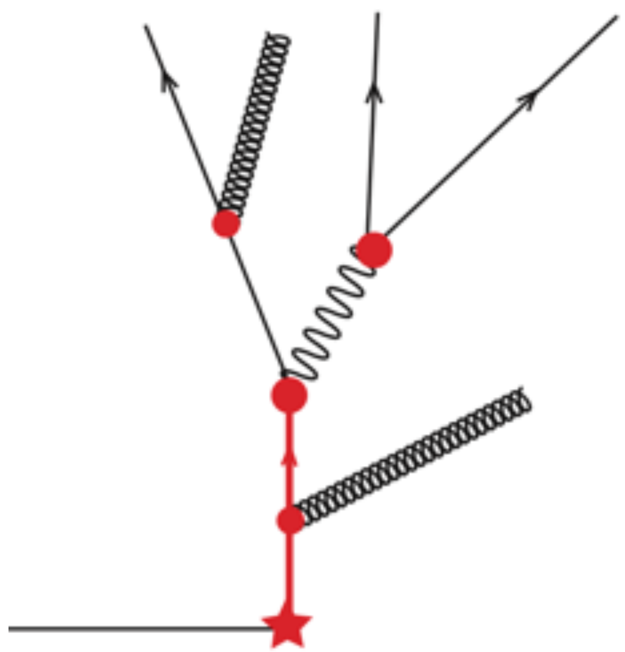




Top Tagging Analytics : Shower Deconstruction

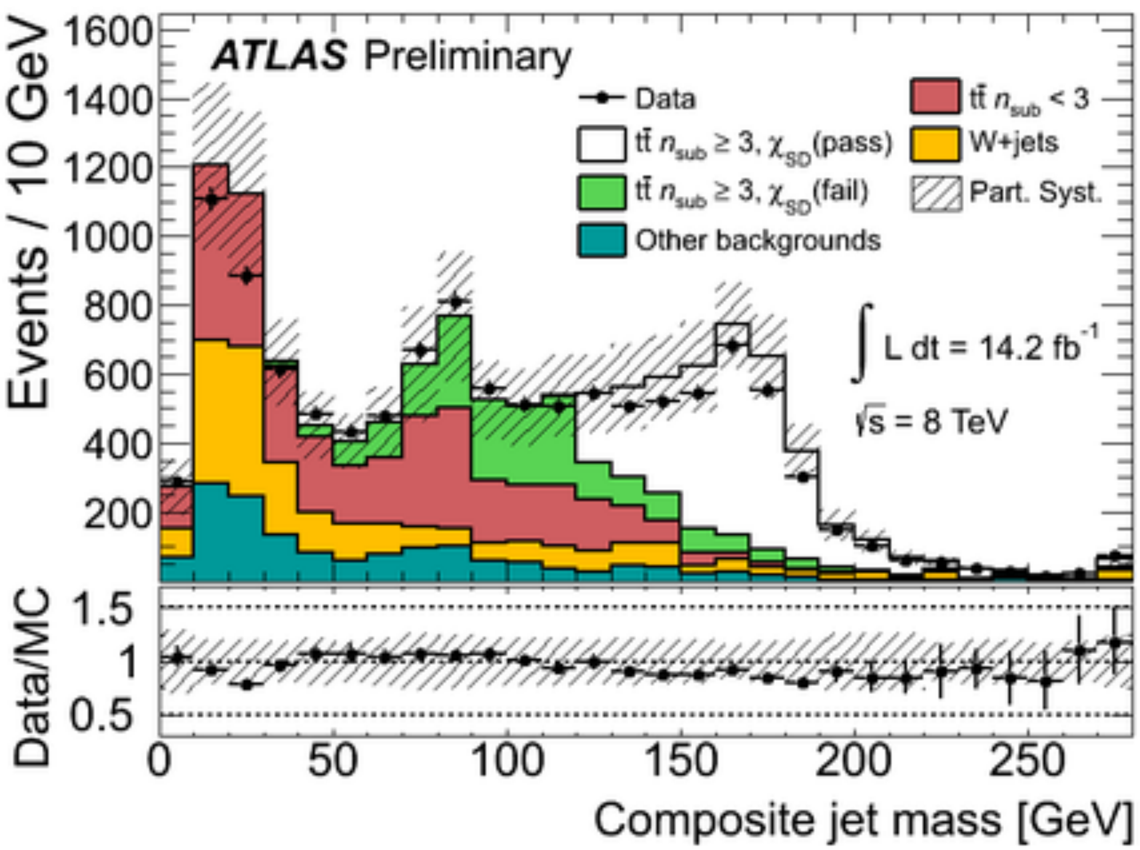
- Make "microjets" out of CA jet constituents
- Keep at most 9 microjets with pt > ptmin
- Approximate probability for observed particles to satisfy a "signal-like" shower, or a "background-like" shower
- Construct likelihood and compare

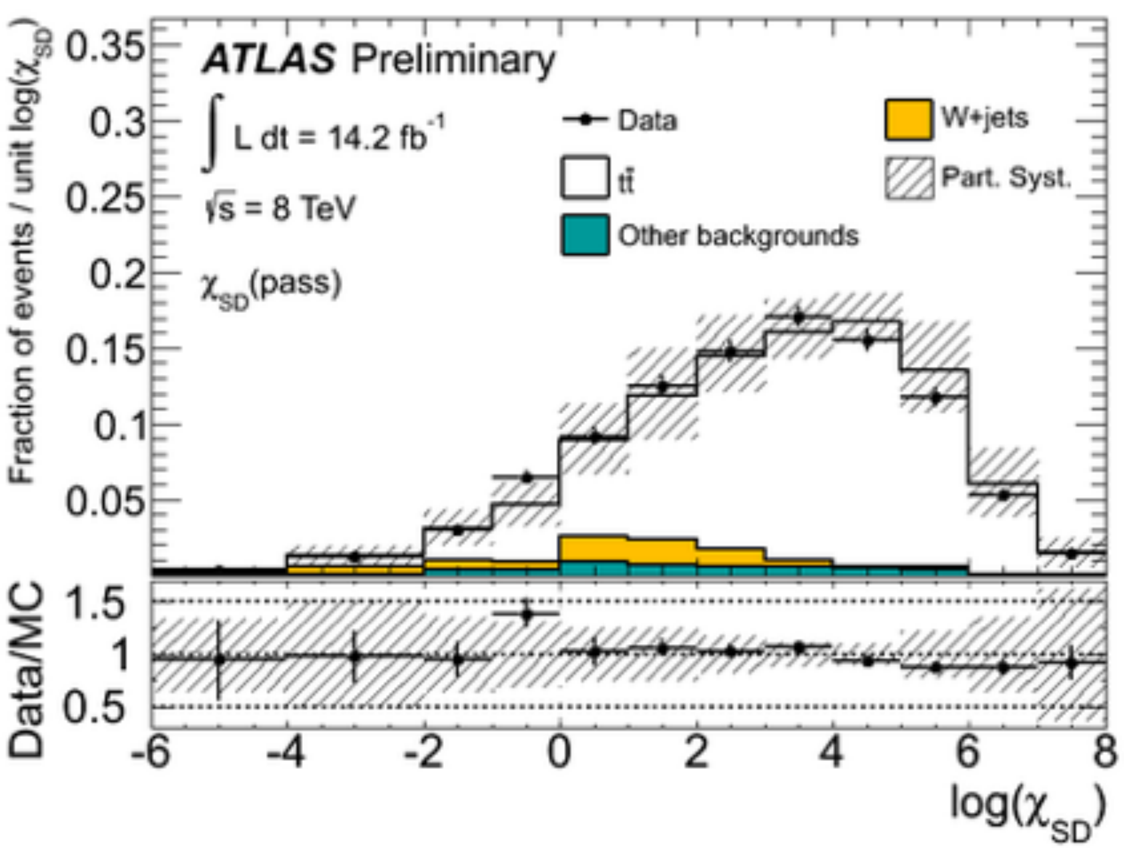




Soper and Spannowsky : Phys.Rev. D84 (2011) 074002

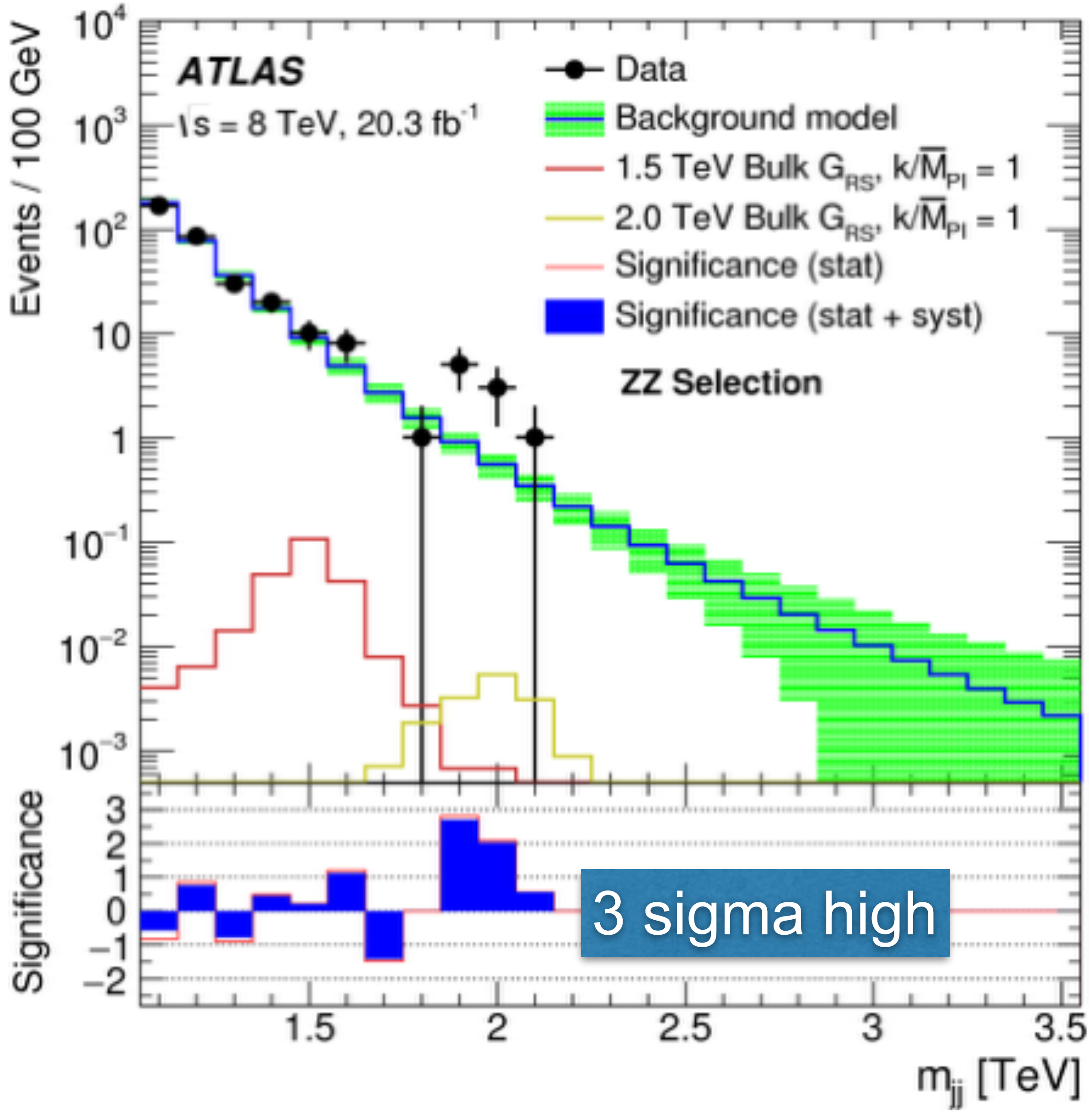
Top Tagging Analytics : Shower Deconstruction



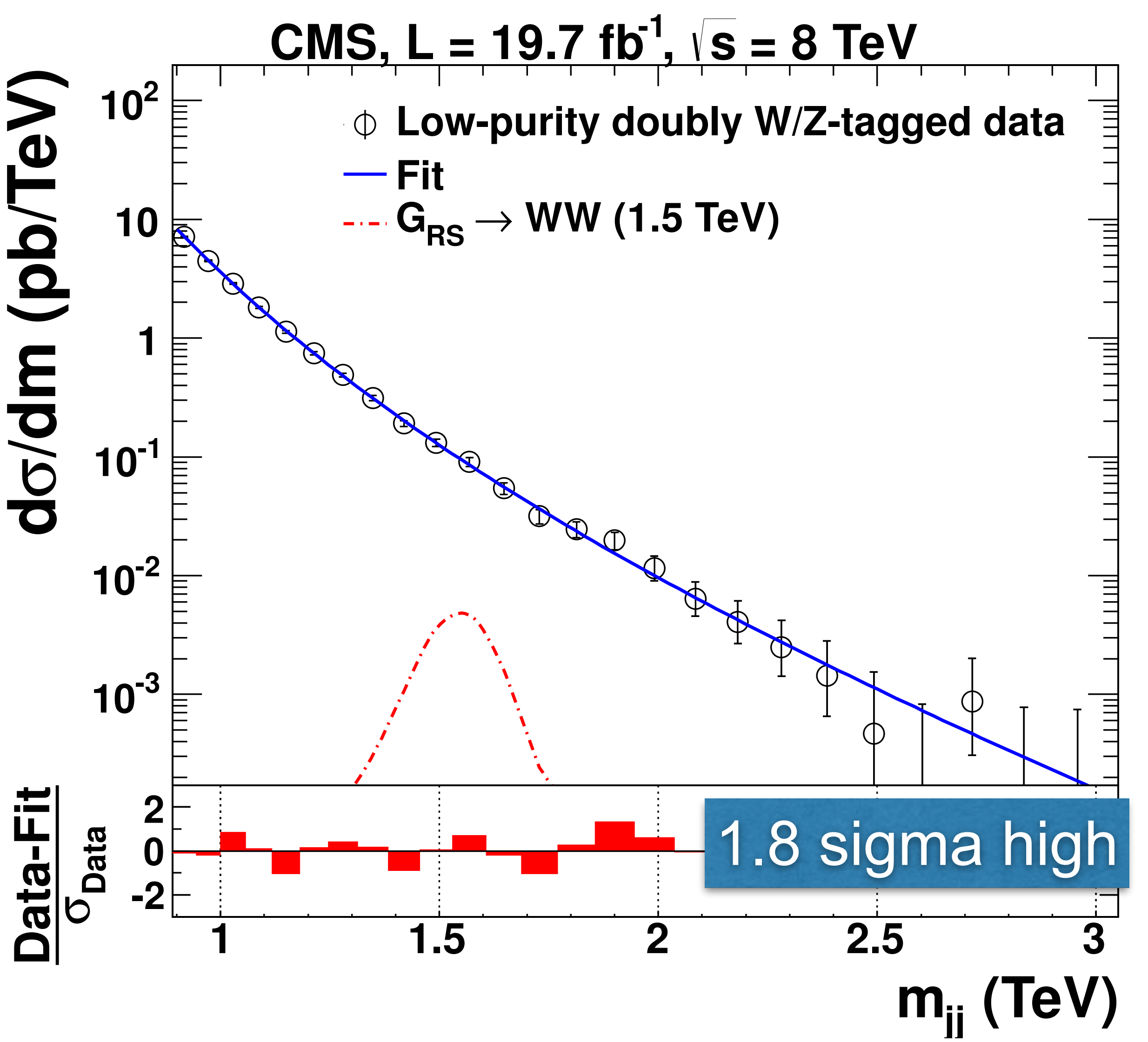


Excitement?

ATLAS X -> VV (hadronic)
EXOT-2013-08

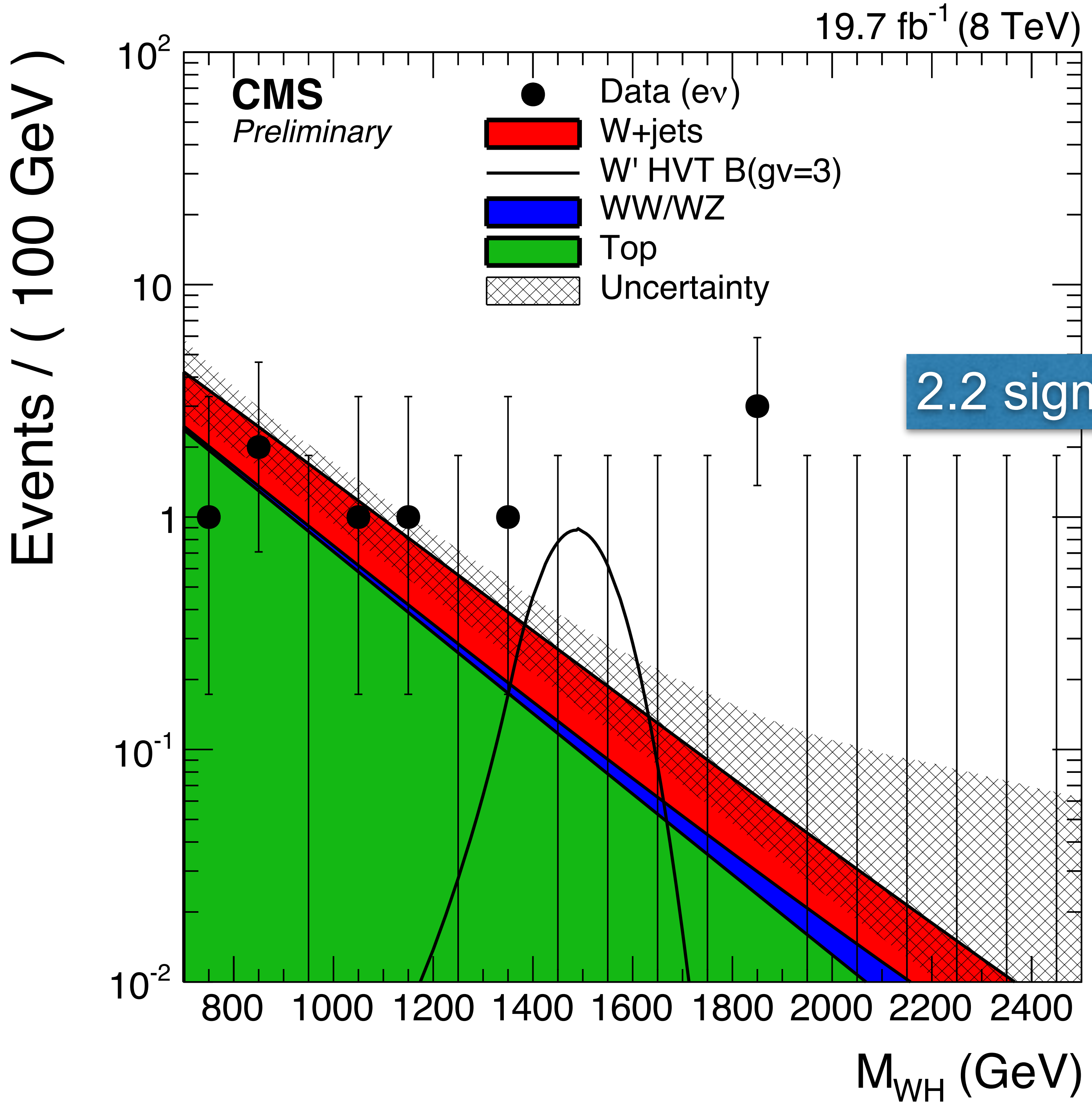


CMS X -> VV (hadronic)
EXO-2012-024



Excitement?

CMS X -> WH (lnubb)
EXO-2014-010



2.2 sigma high

Public Service Announcement