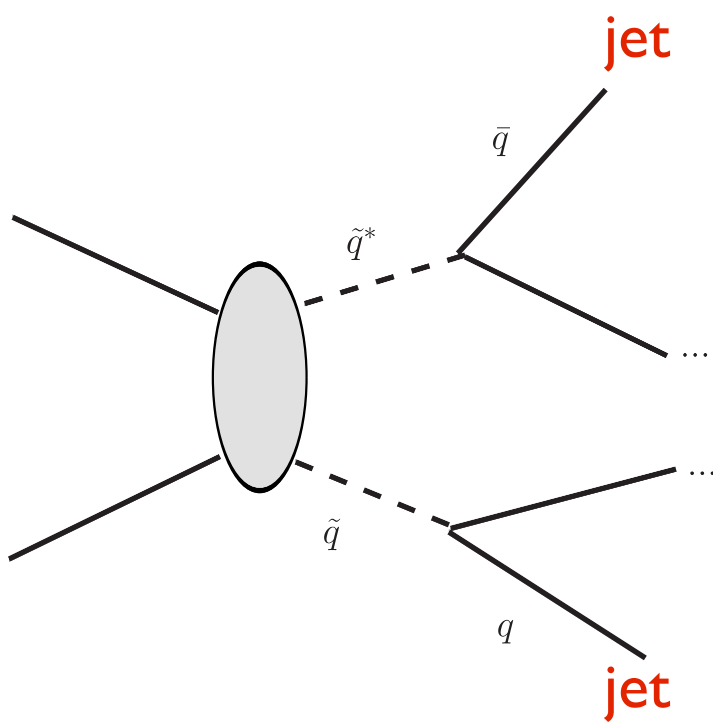
Jet subtructure, grooming.

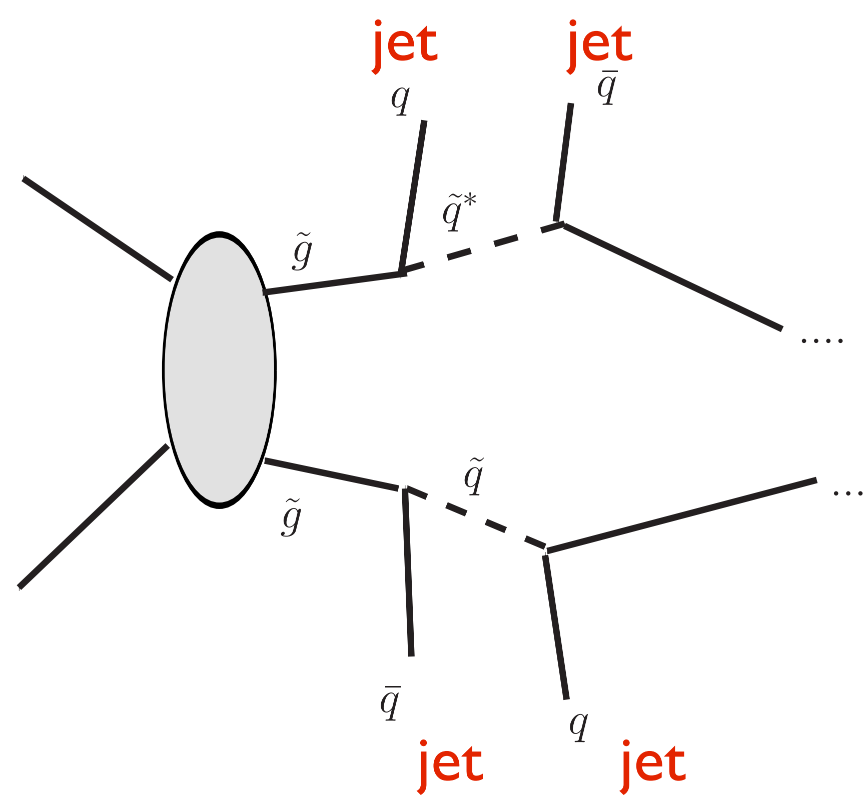
Lian-Tao Wang
University of Chicago

Snowmass QCD working group
Jan. 30, 2013

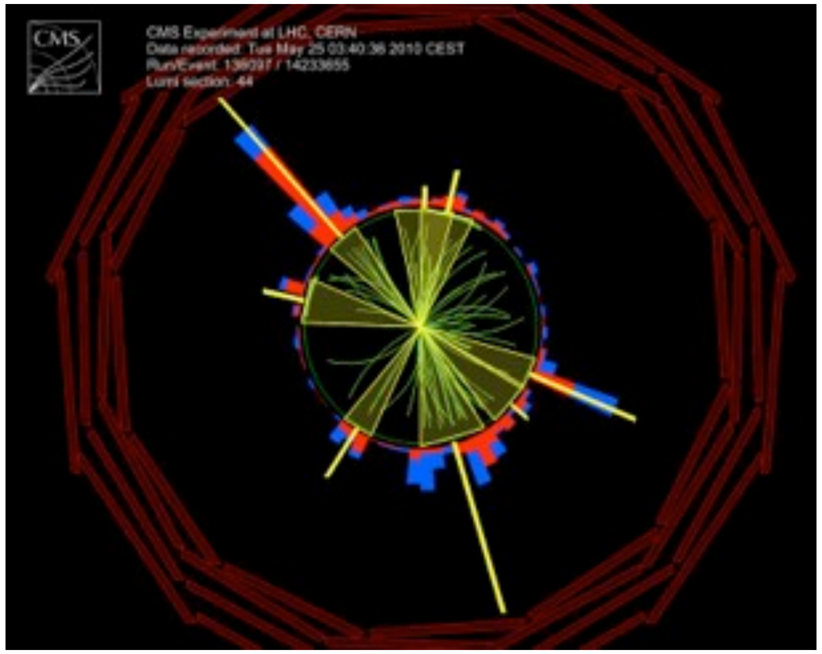
The importance of jets:

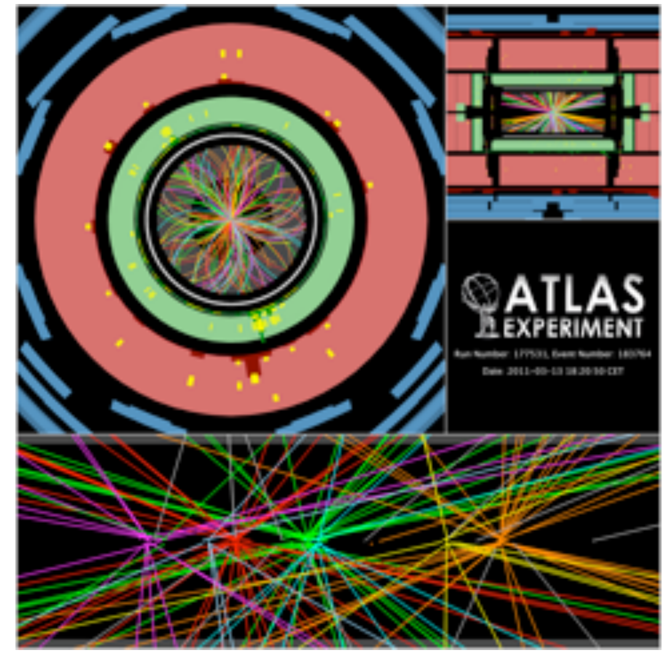
- “Everywhere” at hadron colliders. $p\ p,\ \text{or},\ p\bar{p}$ initial state.
- Present in (almost) all new physics signals.
  - Many of them only have hadronic channels.





Need new jet tools for the LHC.

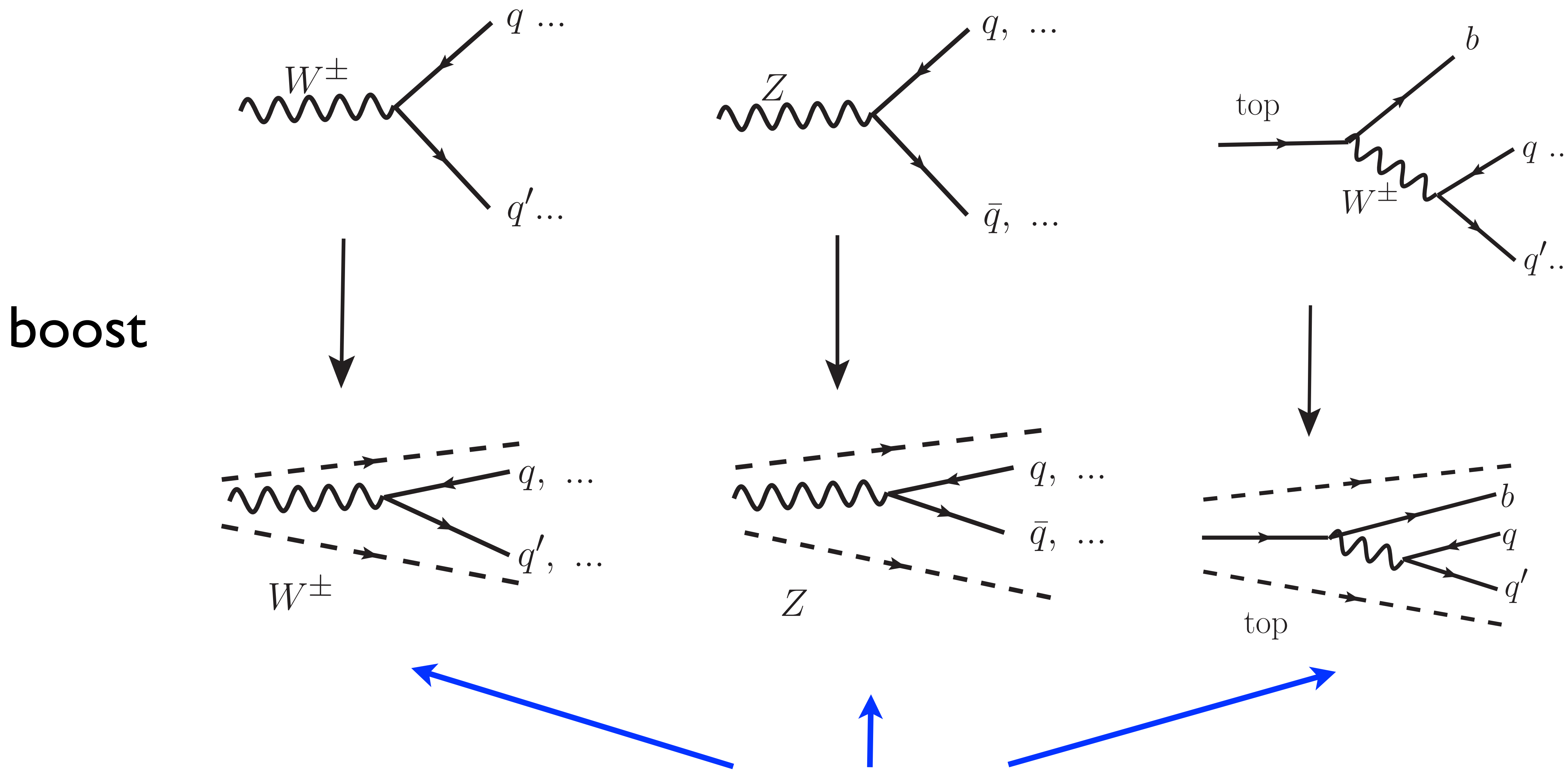




- More energetic, bigger, jet at the LHC.
  - LHC jet: 50 GeV - several TeV
  - Tevatron jet: 50 - 100s GeV
- Much higher "noise" level at the LHC.
  - LHC: 10-100 GeV / rapidity
  - Tevatron: 2-10 GeV / rapidity

Jet look likes

- When produced at TeV-scale energies, they have a large boost.



Jets with substructure.

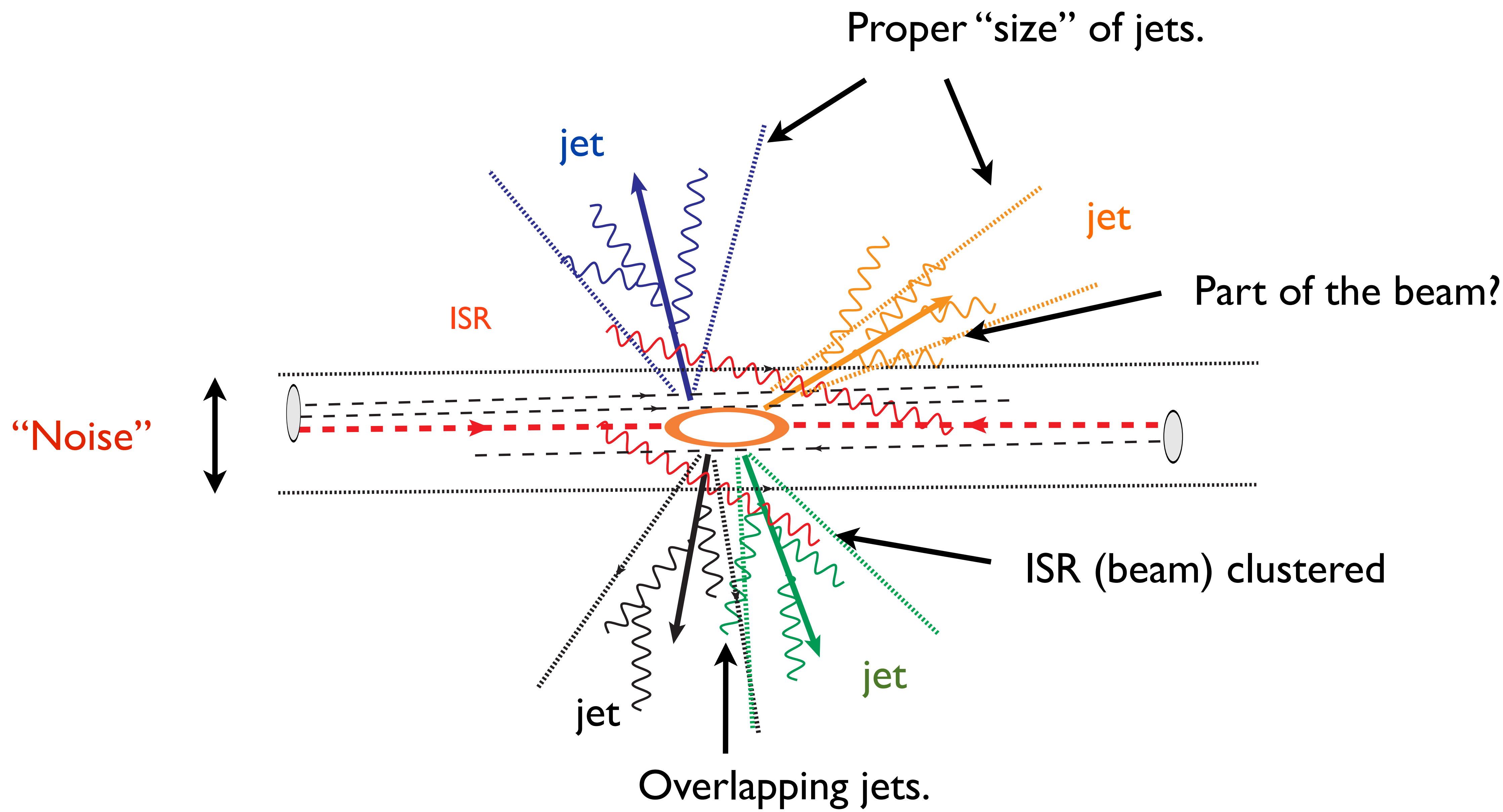
Challenge: distinguishing them from QCD jets (q and g).

Two closely related directions.

- Better QCD jet.
  - Smarter jet algorithm.
  - Noise suppression with jet grooming.
- Jet substructure.
  - Boosted objects: top, Higgs, Z, W, ...
- Review of some highlights. Hope to start discussion on related QCD issues.

Better QCD (q, g) jets.

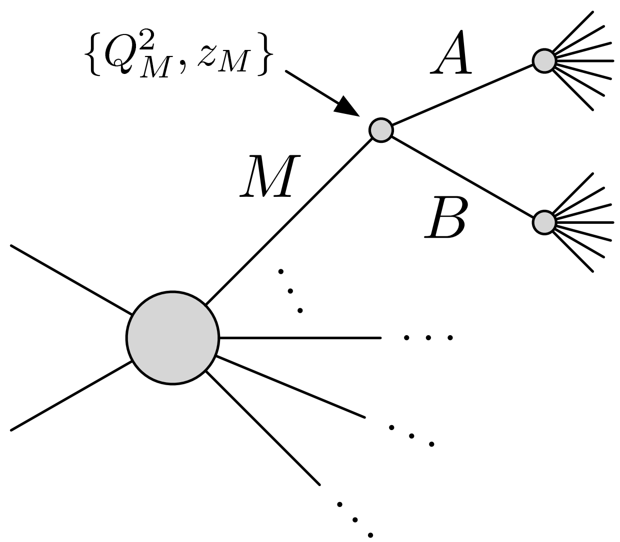
Why is it hard?



- To best preserve $p_{\text{jet}} \simeq p_{\text{parton}}$ we would like to:
  - Use "smart" jet shapes.
  - Reduce "noise".

Basic intuition: parton shower⇒QCD jet

- From the initial parton, a jet is built up by many radiations.



Branching $M \to A + B$ controlled by

Evolution variable: virtuality ("mass") Q_M, or , p_T

Energy fraction: $z = \min(E_A, E_B)/E_M$

$$d\sigma^{\rm QCD}_{M\to A+B} = \frac{dQ_M^2}{2\pi}\frac{1}{Q_M^2}\, dz \frac{\alpha(\mu)}{2\pi} P_{M\to AB}(z)\Delta(\mu_{\rm start},\mu)$$

Prefers collinear radiation

P ~ $(z)^{-1}$ prefers soft radiation

QCD jet: a cluster of radiation

a) relatively soft

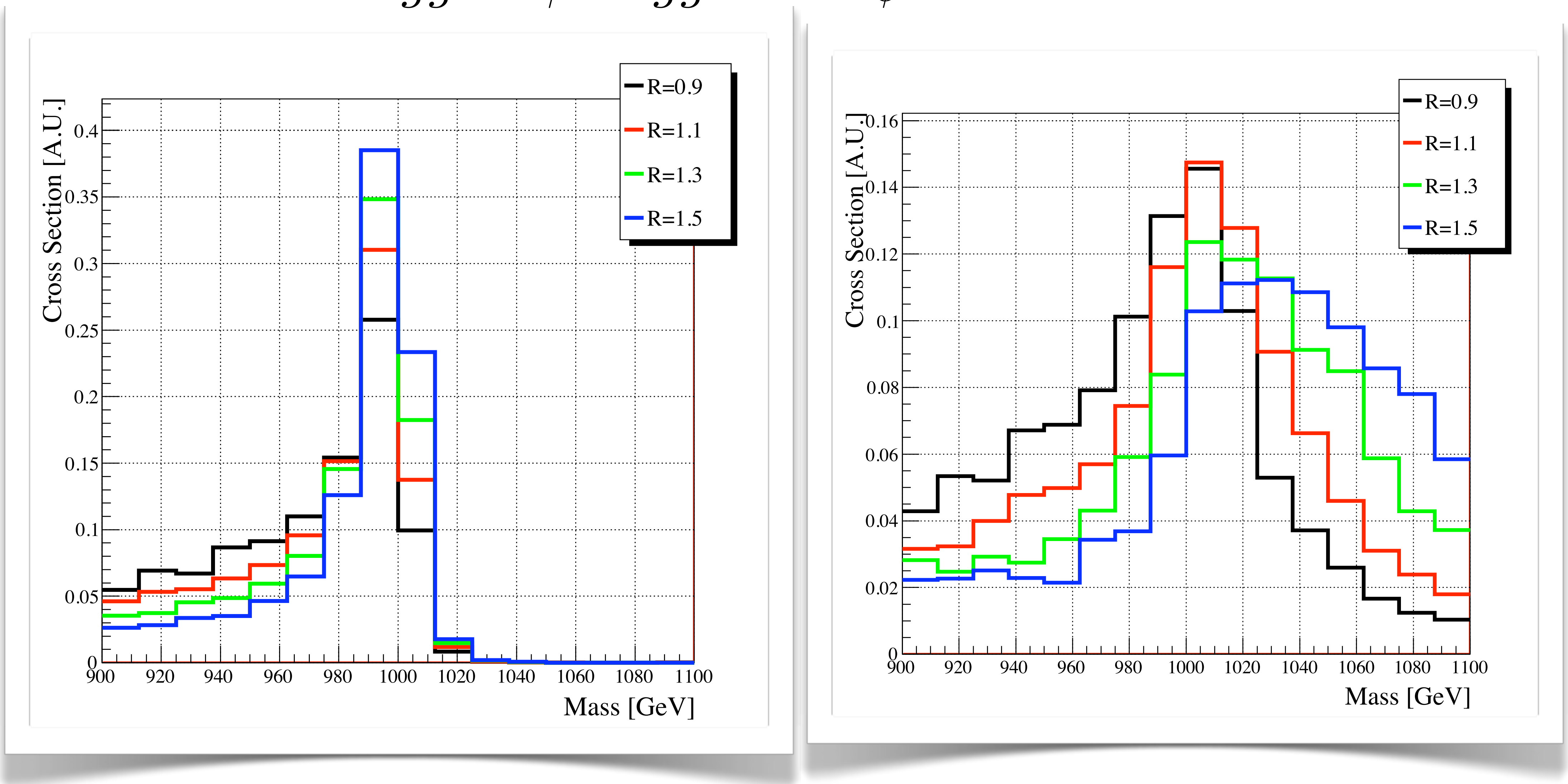
b) close to the direction of P_M

c) approximately symmetrical around P_M

"noise" control

- Noise: Initial state radiation (ISR), multiple interaction (MI), underlying events (UE), pile-up (PU).

$gg \to \phi \to gg \qquad m_\phi = 1 \text{ TeV}$



FSR only

Including ISR, MI, UE, pile-up

Room for improvement!

Jet trimming.

D. Krohn, J. Thaler, LTW, arXiv:0912.1342

- Introducing a "cut" on soft radiation.

- Our implementation.
  - Cluster all calorimeter data using any algorithm
  - Take the constituents of each jet and recluster with smaller radius Rsub (for example, Rsub = 0.2).
  - Discard the subjet i if $p_{Ti} < f_{\rm cut} \cdot \Lambda_{\rm hard}$

- Best choice of the hard scattering scale and fcut.
  - Process dependent.
  - Can be optimized experimentally.

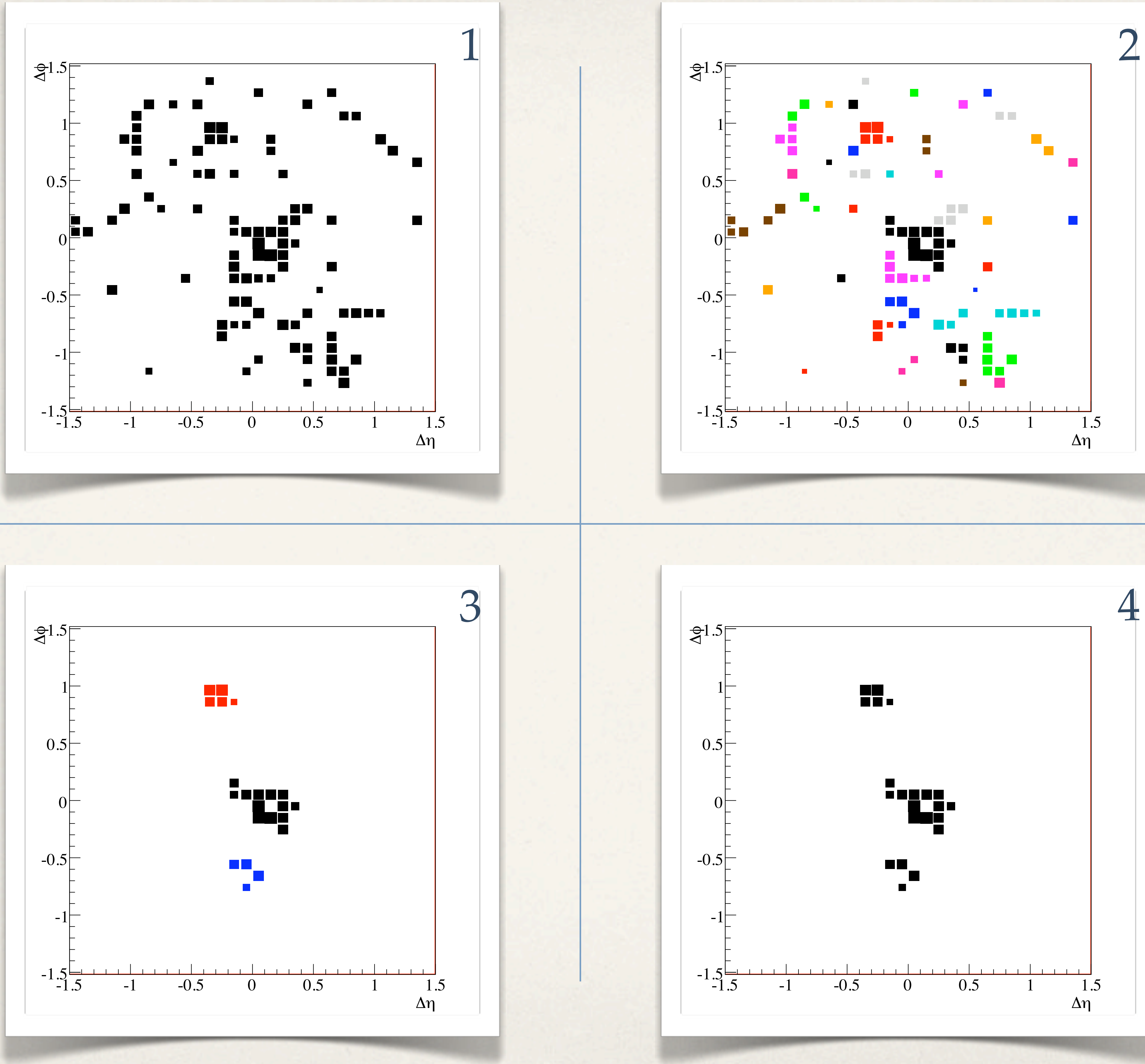
Related but different "jet grooming" approaches:
Filtering: J. Butterworth, A. Davison, M. Rubin, G. Salam, arXiv:0802.2470
Pruning: S. Ellis, C. Vermilion, J. Walsh, arXiv:0903.5081

Start

Cluster into subjets

Discard soft subjets

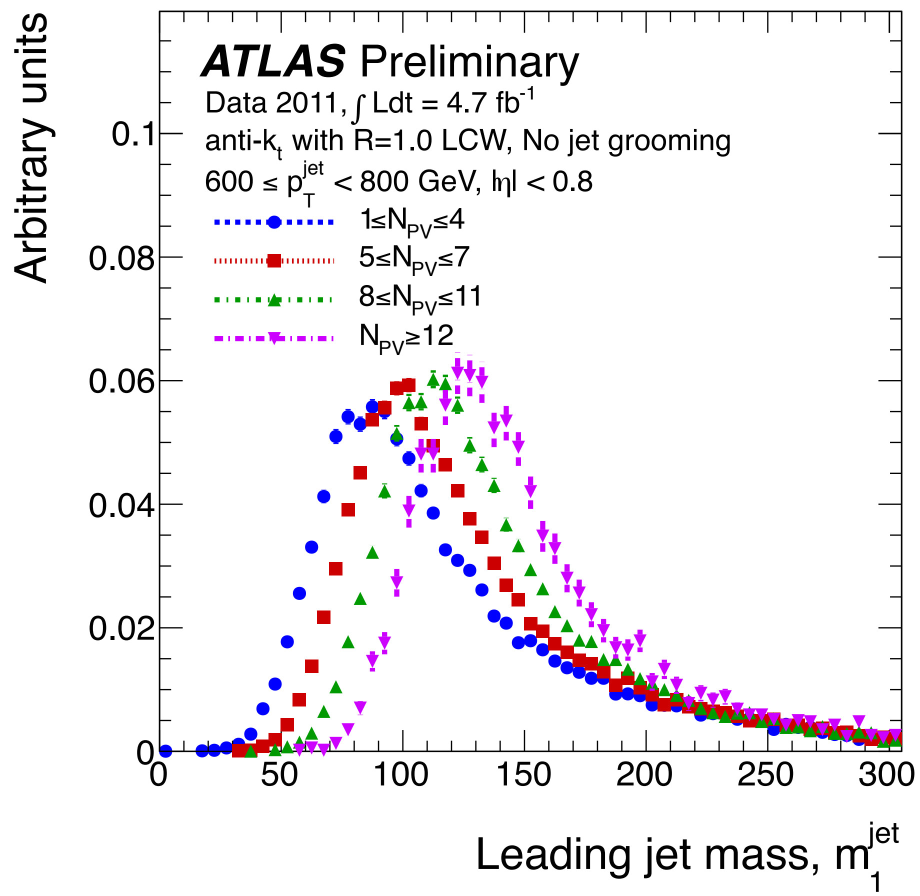
Reassemble



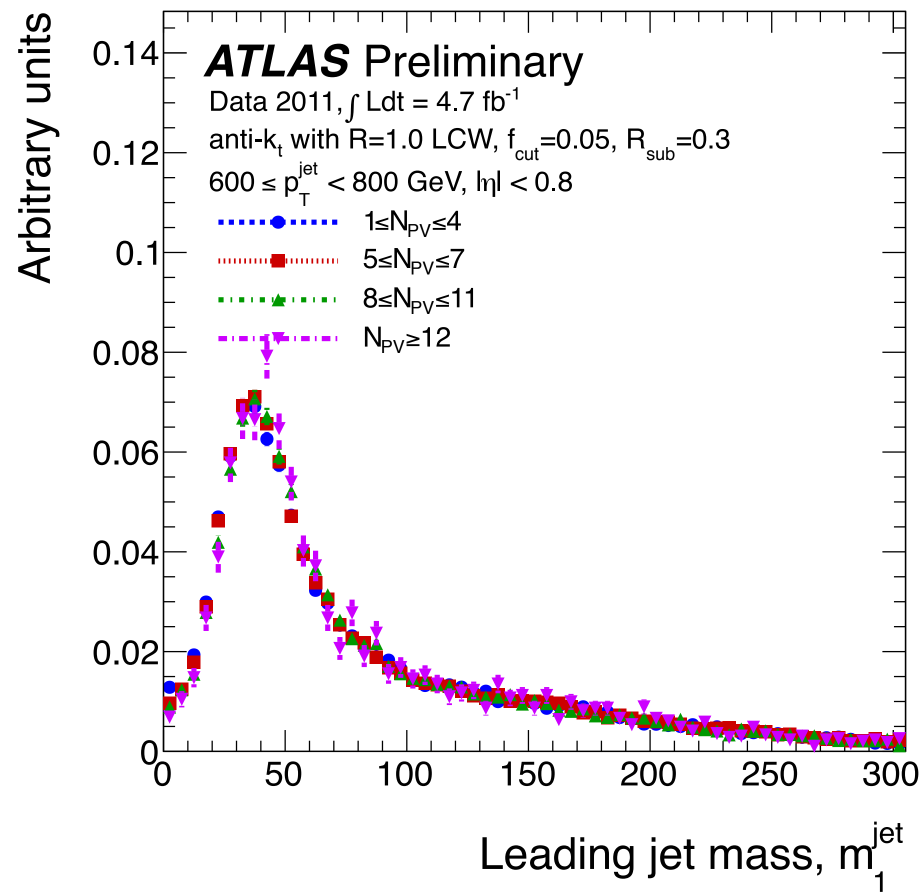
Comparisons of "fat" jet substructure with 5 fb^{-1} *(II)*

"Ungroomed" and trimmed jet mass: `PYTHIA` *vs.* `POWHEG+PYTHIA`

ATLAS-CONF-2012-066



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



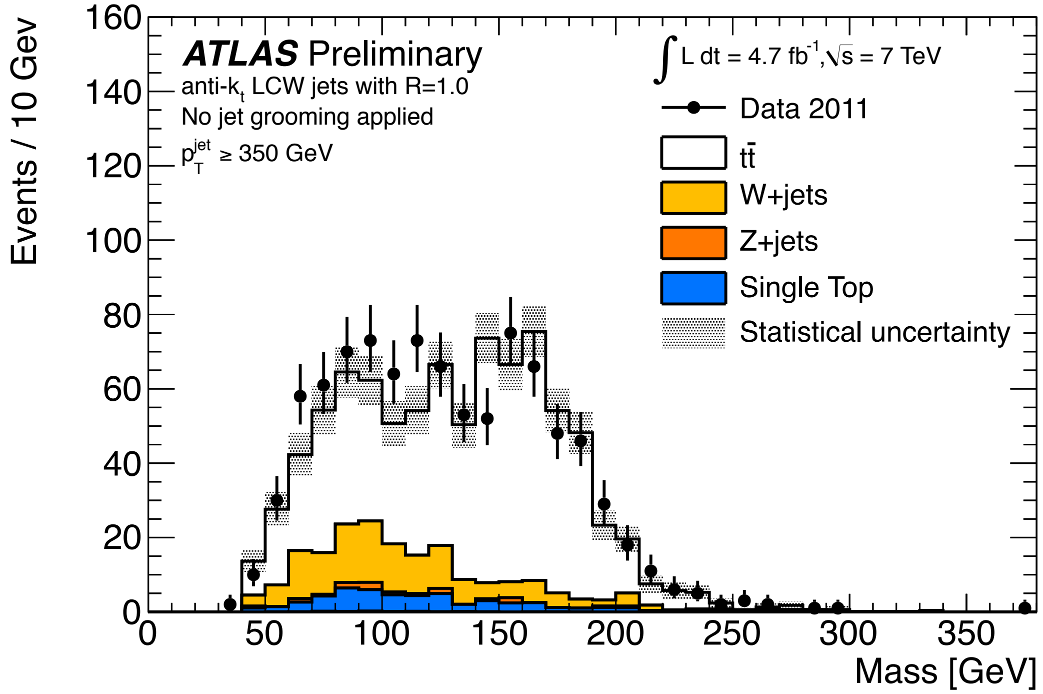
anti-k_t, $R = 1.0$ mass (trimmed)

- Grooming provides resilience against pile-up
- Full jet mass distribution exhibits significantly improved stability
- *Not* a jet mass correction but rather a robust **new definition** of a jet

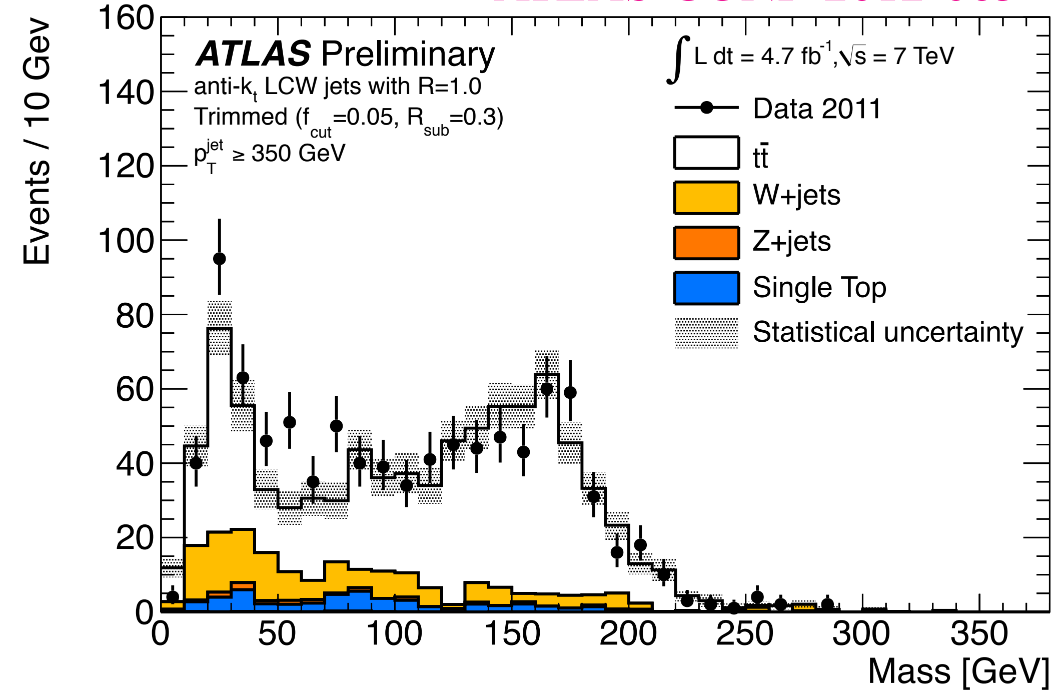
Commissioning boosted object tools with SM top quarks (I)

Enriched sample of boosted tops using semi-leptonic (μ) selection and high-p_T fat jets

ATLAS-CONF-2012-065



anti-k_t, $R = 1.0$, ungroomed



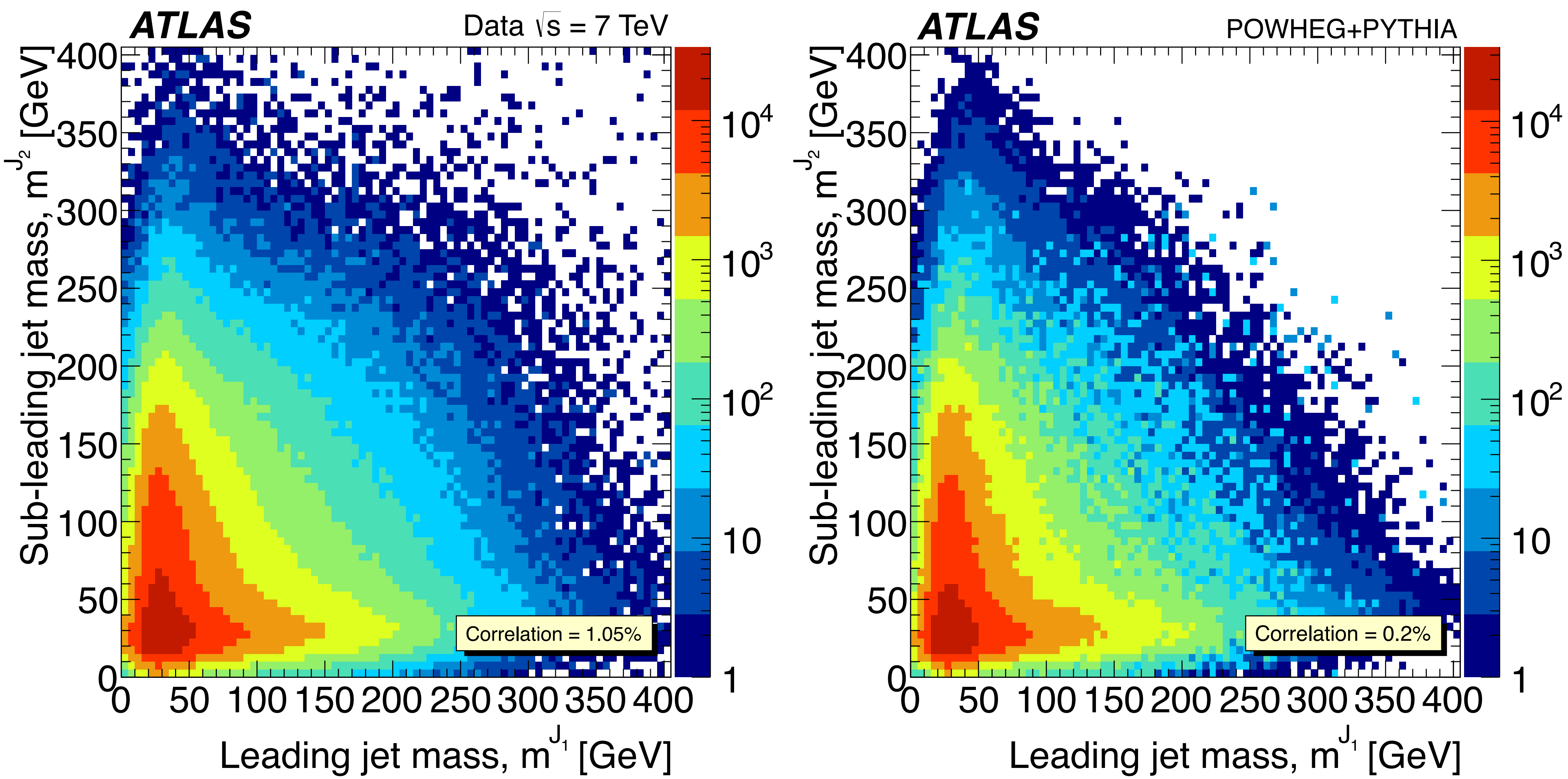
anti-k_t, $R = 1.0$, trimmed

- Significant increase in the purity of the top mass peak between $120 < M^{\text{jet}} < 200$ GeV
- Narrower top mass peak after trimming that is well described by the data
- Rate of boosted tops is well-predicted by MC@NLO top MC
  - *implicit limits from this plot alone*
- *b*-tag requirement highlights improvement in mass resolution from trimming

Trimmed jet mass correlations (inclusive)

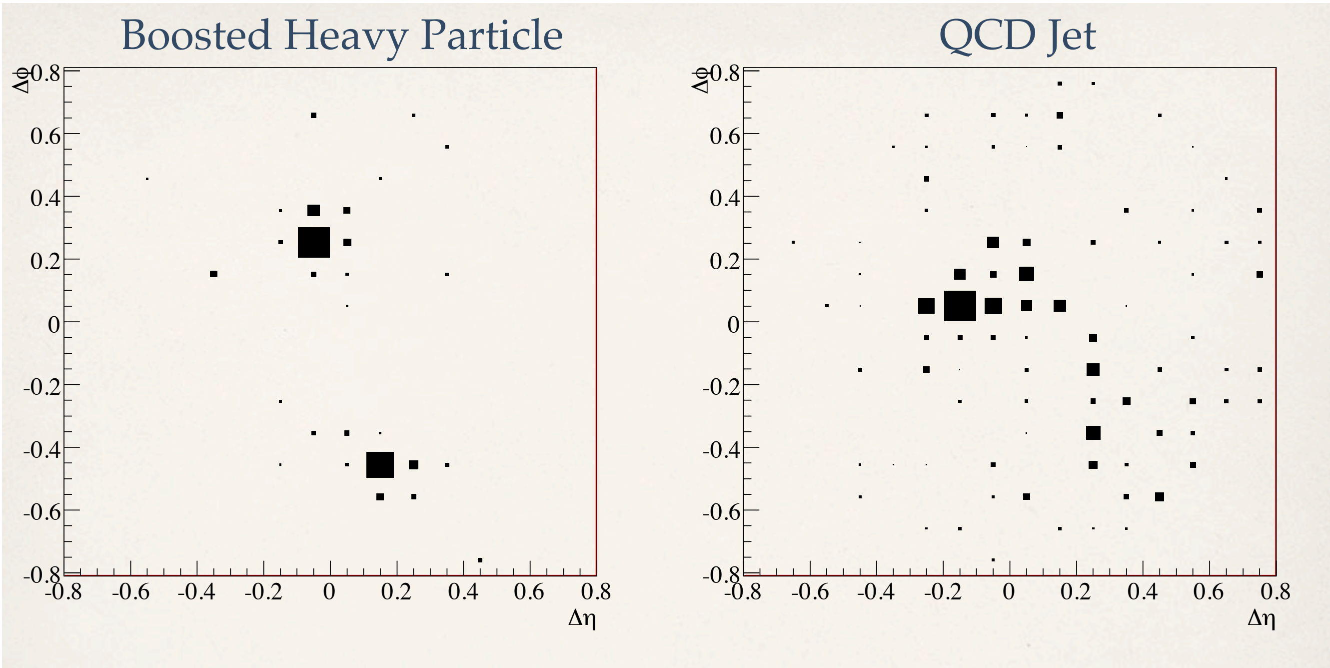
RPV gluino search with substructure (see N. Tran's talk, arXiv:1210.4813)

- Correlations between leading and subleading (in p_T^{jet}) jet masses



- **Data:** 1.05%, **`POWHEG`:** 0.2% (trimmed)
- Decent agreement between data and `POWHEG` (Correlation driven by low mass jets)

Jet substructure, and applications in new physics searches.



Pioneering work: M. Seymour 1991, 1993

Many ideas

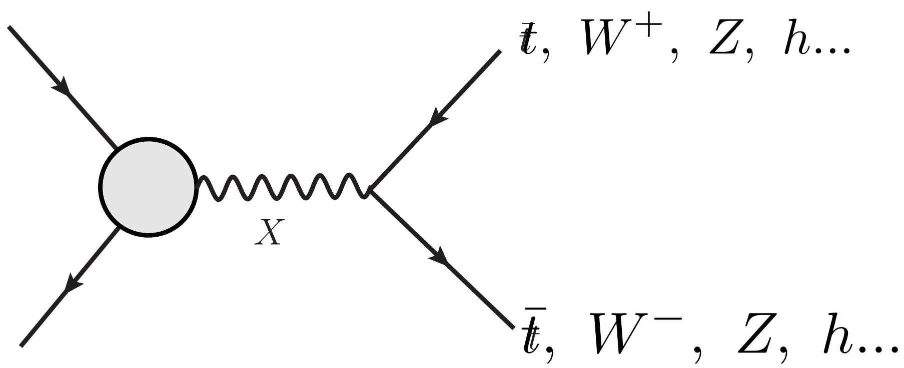
Jet Grooming	Filtering	Butterworth, Davison, Rubin, Salam (0802.2470)
	Pruning	Ellis, Vermilion, Walsh (0903.5081)
	Trimming	Krohn, Thaler, Wang (0912.1342)
2-pronged resonances	Mass-drop/Filtering and variations	Butterworth, Davison, Rubin, Salam (0802.2470)
		Plehn, Salam, Spannowsky (0910.5472)
		Kribs, Martin, TSR, Spannowsky; (0912.4731, 1006.1656)
3-pronged resonances	Y-splitter	Butterworth, Cox, Forshaw (hep-ph/021098)
	Johns Hopkins tagger	Kaplan, Rehermann, Schwartz, Tweedie (0806.0848)
	HEP tagger	Plehn, Spannowsky, Takeuchi, Zerwas (1006.2833)
	tree-less approach	Jankowiak, Larkoski (1104.1646)
General procedures	Template method	Almeida et al. (1006.2035)
	N-subjettiness	Thaler, Van Tilburg (1011.2268); Kim (1011.1493)
	Multi-variate approach	Gallicchio, Schwartz (1106.3076)
	Shower deconstruction	Spannowsky, Soper 1102.3480)
	Qjets	Ellis, Hornig, Krohn, TSR, Schwartz (1201.1914)

slide from T. Roy

When to consider substructure

- Have to consider the boosted objects.

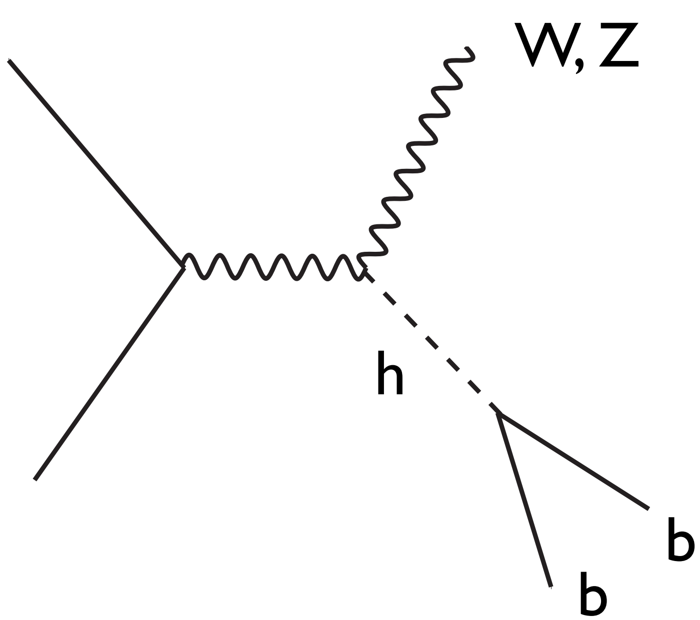
e.g.



For example, boost tops
Brooijmans; Lillie, Randall, LTW; Thaler, LTW;
D. Kaplan, K. Reherman, M. Schwartz, B. Tweedie;
L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, J. Virzi
S. Chekanov and J. Proudfoot.
...

- It is beneficial to consider the boosted objects.

e.g.



Lower combinatorics,
SM background boost differently.

Butterworth, Davidson, Bubin, Salam

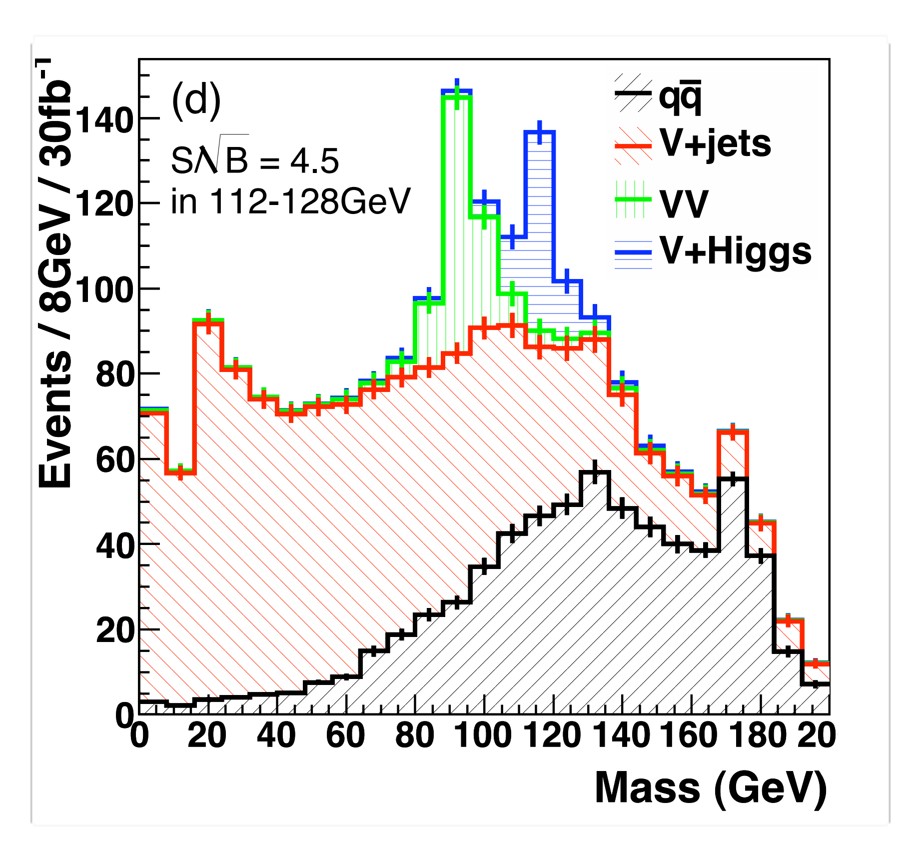
For a summary of recent developments: C. Vermilion, 1001.1335

BDRS (filtering)

Butterworth, Davison, Rubin, Salam, 0802.2470

- Z+H and W+H with H->bb.
- Considered boosted Higgs.
  - Better acceptance. $p_T^h > 200$ GeV
  - backgroud such as ttbar boost differently.

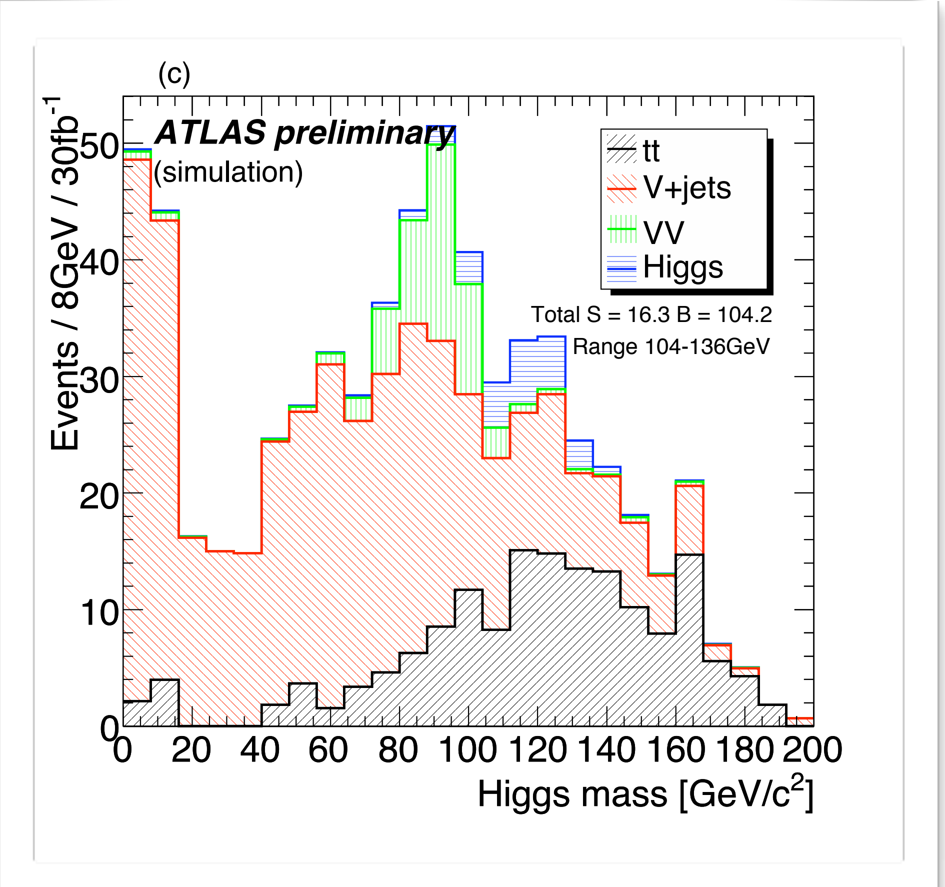
$4.5(8.2)\sigma$ with 14 TeV and 30(100) fb^{-1}



- Similar result reproduced by ATLAS

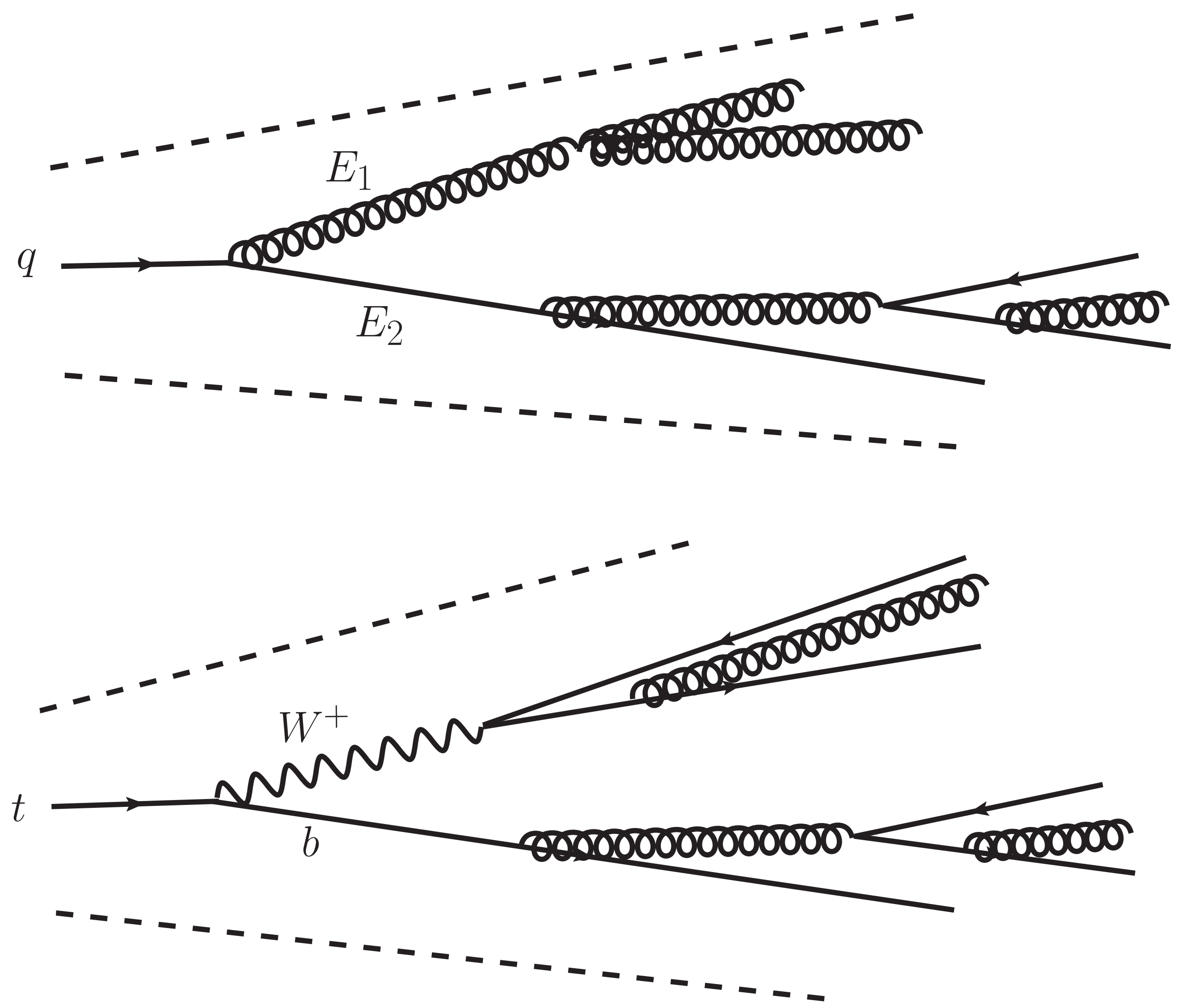
3.7σ with 14 TeV and 30 fb^{-1}

ATL-PHYS-PUB-2009-088



Example: boosted top tagging at the LHC

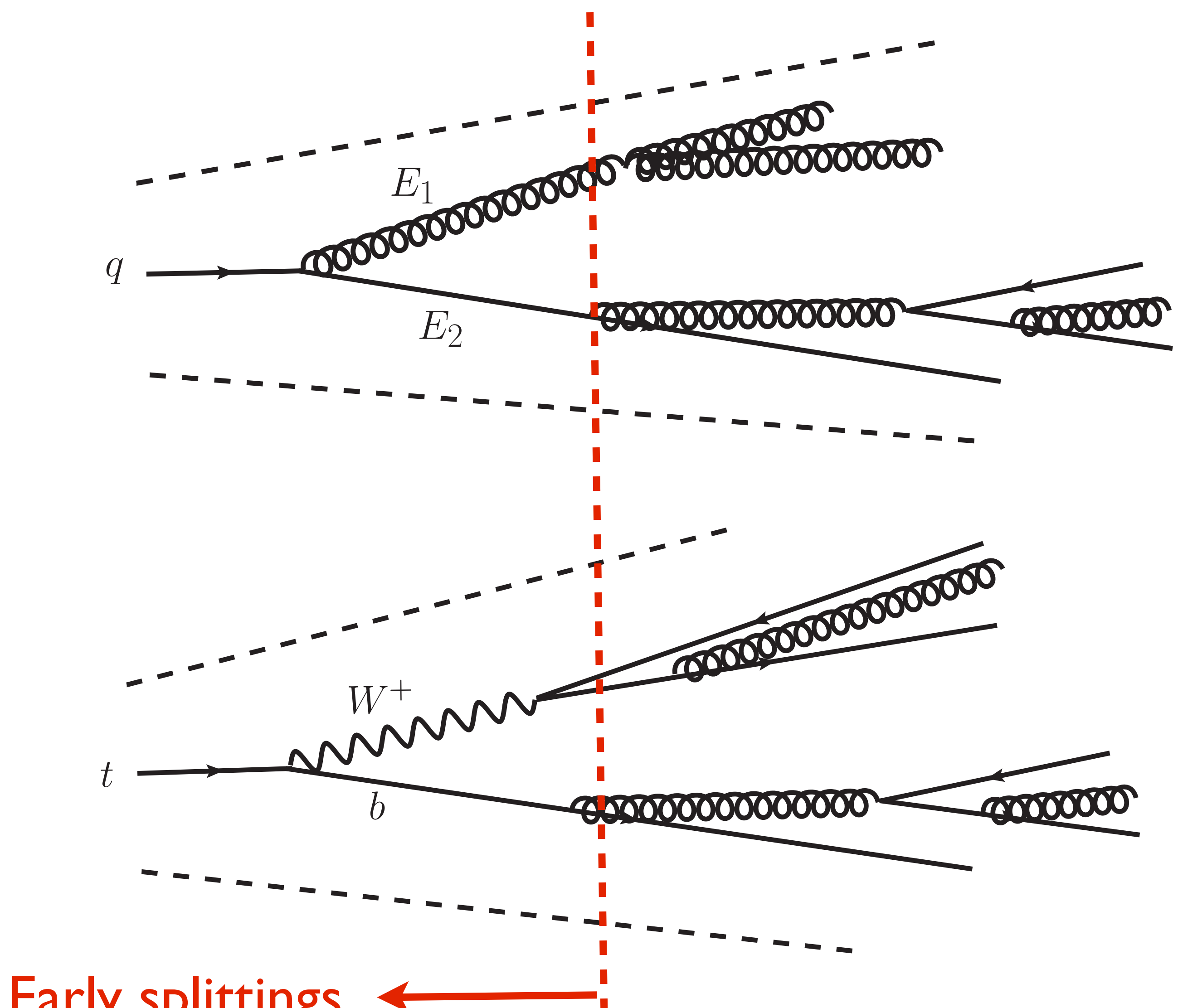
- Fully collimated tops look like QCD jets.



Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Zooming in near the first splitting

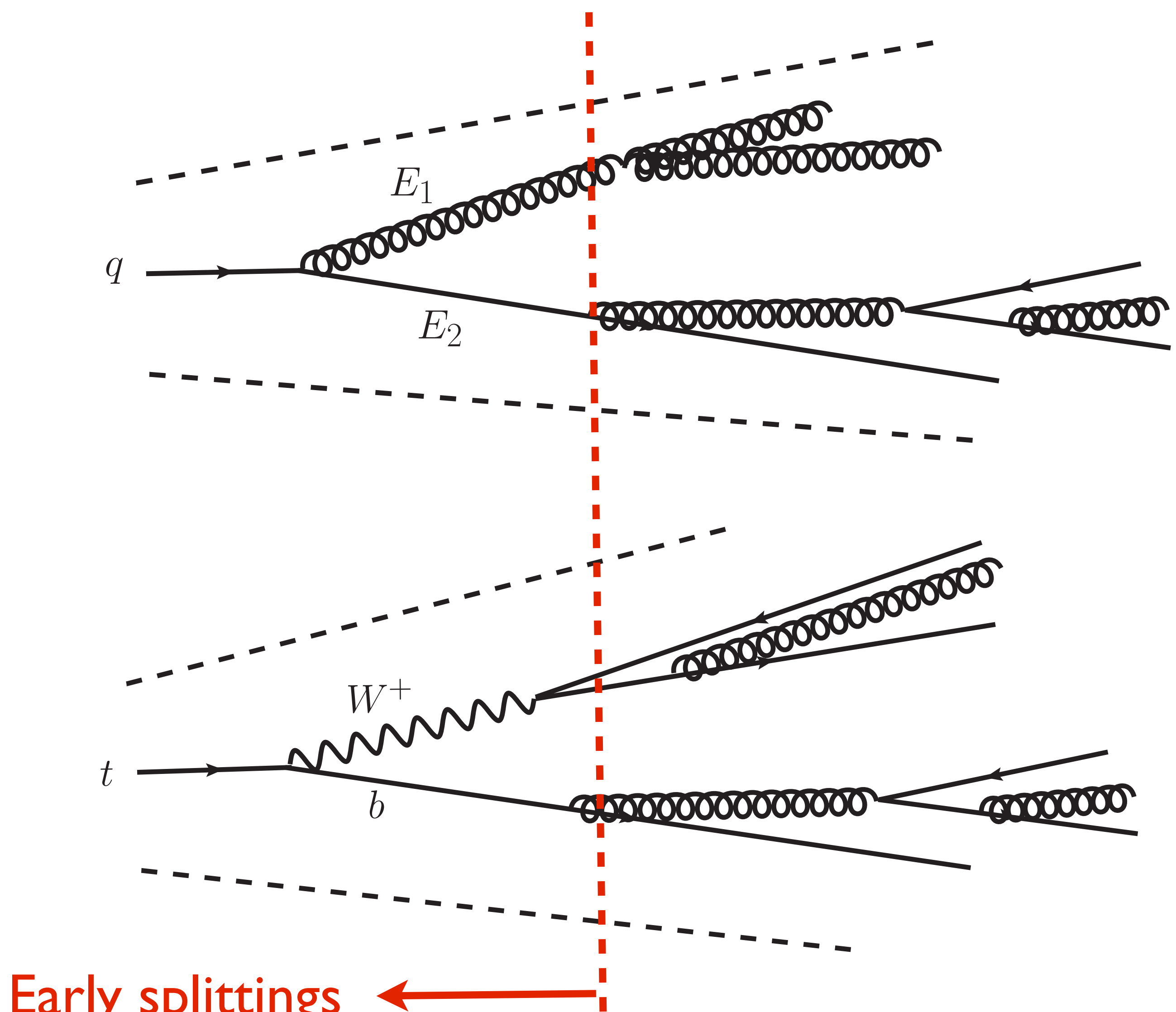
QCD. Soft radiation: $z = \frac{\mathrm{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

Top. Decay: $z = \frac{\mathrm{Min}(E_\mathrm{W}, E_b)}{E_\mathrm{W} + E_b} \to \text{finite}$

Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:
- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Zooming in near the first splitting

QCD. Soft radiation: $z = \dfrac{\mathrm{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

Jet mass: $d\sigma \propto \dfrac{1}{m_{\rm jet}^2}$

Top. Decay: $z = \dfrac{\mathrm{Min}(E_{\rm W}, E_b)}{E_{\rm W} + E_b} \to \text{finite}$

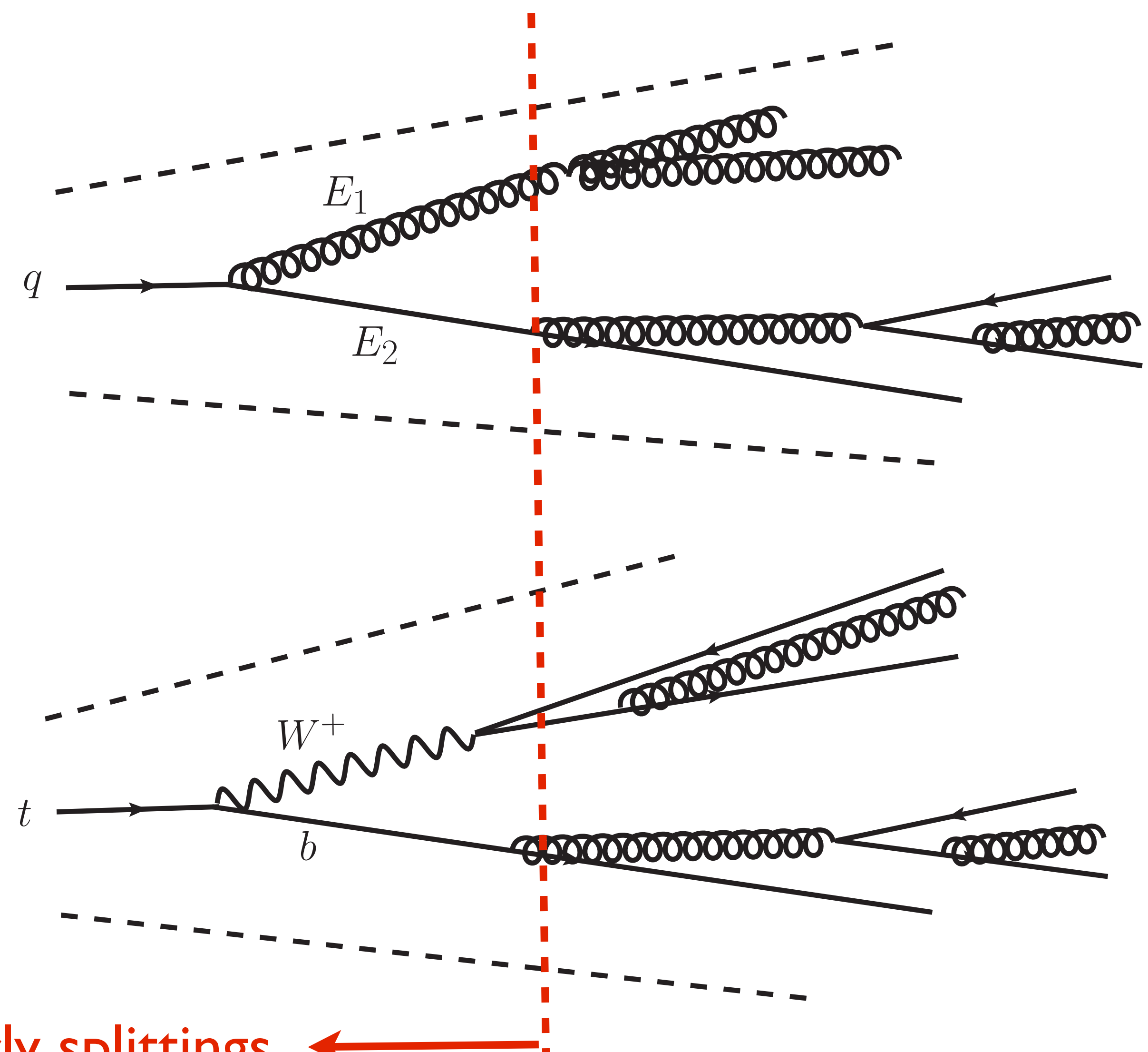
Jet mass: $m_{\rm jet} \simeq m_{\rm top}$

Example: boosted top tagging at the LHC

- Fully collimated tops look like QCD jets.

Basic distinction:

- QCD: radiation.
- Top decay: $t \to bW(\to qq')$ 3 hard objects.



Early splittings

Zooming in near the first splitting

QCD. Soft radiation: $z = \frac{\mathrm{Min}(E_1, E_2)}{E_1 + E_2} \to 0$

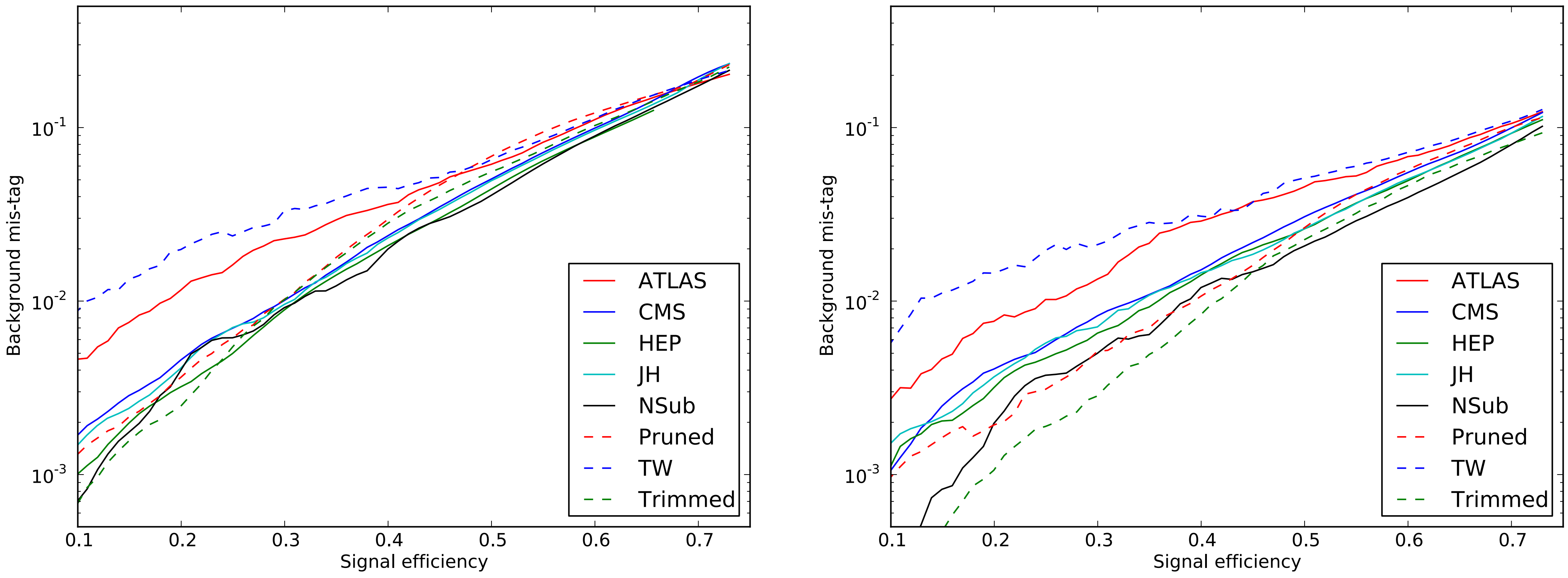
Jet mass: $d\sigma \propto \frac{1}{m_{\mathrm{jet}}^2}$

Top. Decay: $z = \frac{\mathrm{Min}(E_{\mathrm{W}}, E_b)}{E_{\mathrm{W}} + E_b} \to \text{finite}$

Jet mass: $m_{\mathrm{jet}} \simeq m_{\mathrm{top}}$

microscope: jet substructure variables

Various top taggers



(a) all p_T, optimised

(b) p_T 500–600 GeV, optimised

Applied to various searches.

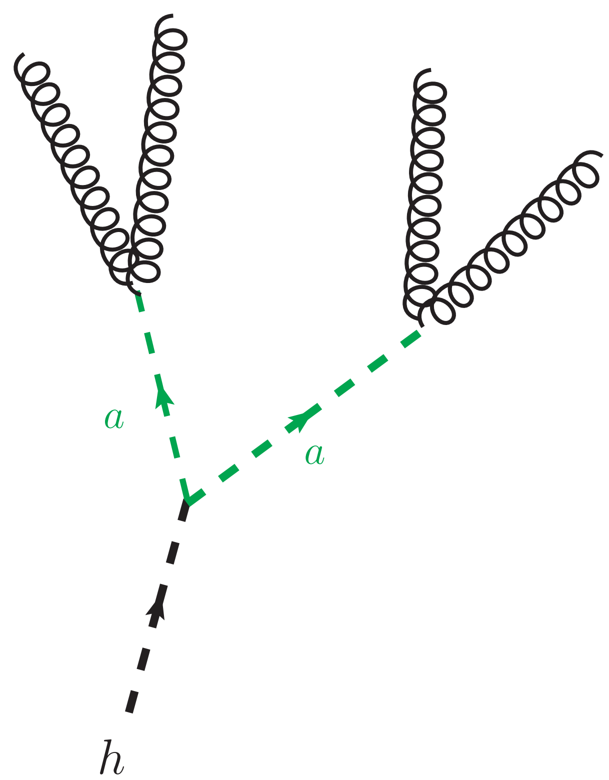
C. Vermillion, et al, 1201.0008

G. Brooijmans, arXiv:0802.3715;
CMS Coll. CMS PAS JME-09-001
J. Thaler, LTW, arXiv:0806.0023
D. Kaplan, K. Reherman, M. Schwartz, B. Tweedie, arXiv: 0806.0848.
L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, J. Virzi, arXiv:0807.0243
L. Almeida, S. Lee, G. Perez, G. Sterman, I. Sung, arXiv:1006.2035
Barger, Huang, 1102.3183

Unbury the Higgs.

$h \to aa \to gggg$, "buried"!

For example:
B. Bellazzini, C. Csaki, A. Falkowski, A. Weiler, arXiv:0910.3210, arXiv:0906.3026



Soft gluon jets,
considered impossible.

$h \to aa \to c\bar{c}c\bar{c}$, "charmful"?

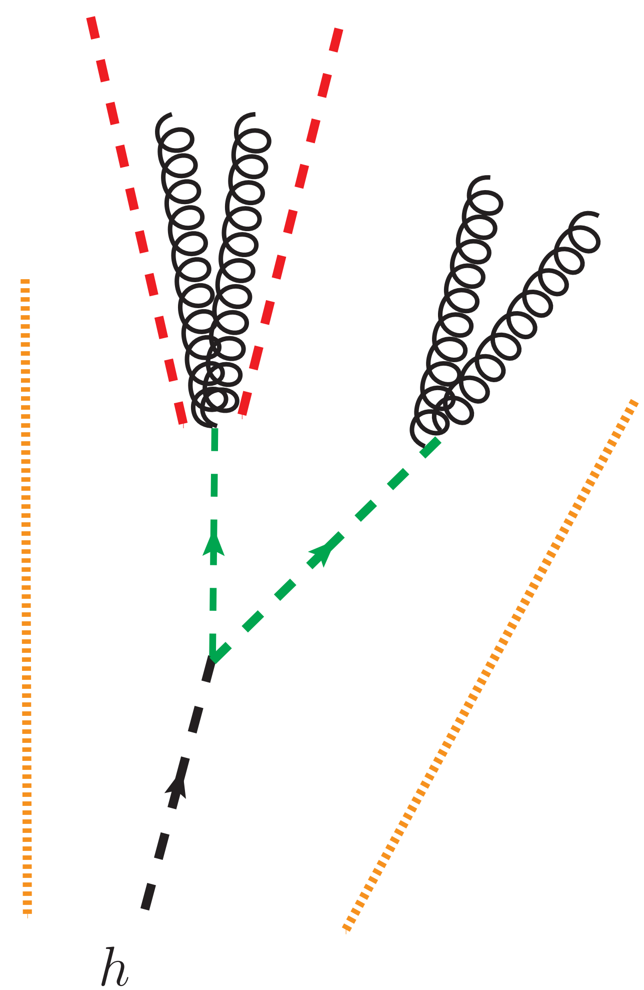
$h \to aa \to 4\tau,\ 4b,\ \bar{b}b\bar{\tau}\tau$

For example:
P. Graham, A. Pierce, J. Wacker, hep-ph/0605162
M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

Unbury the Higgs.

$h \rightarrow aa \rightarrow gggg,$ "buried"!

For example:
B. Bellazzini, C. Csaki, A. Falkowski, A. Weiler, arXiv:0910.3210, arXiv:0906.3026



Boosting the Higgs.

$h \rightarrow aa \rightarrow c\bar{c}c\bar{c},$ "charmful"?

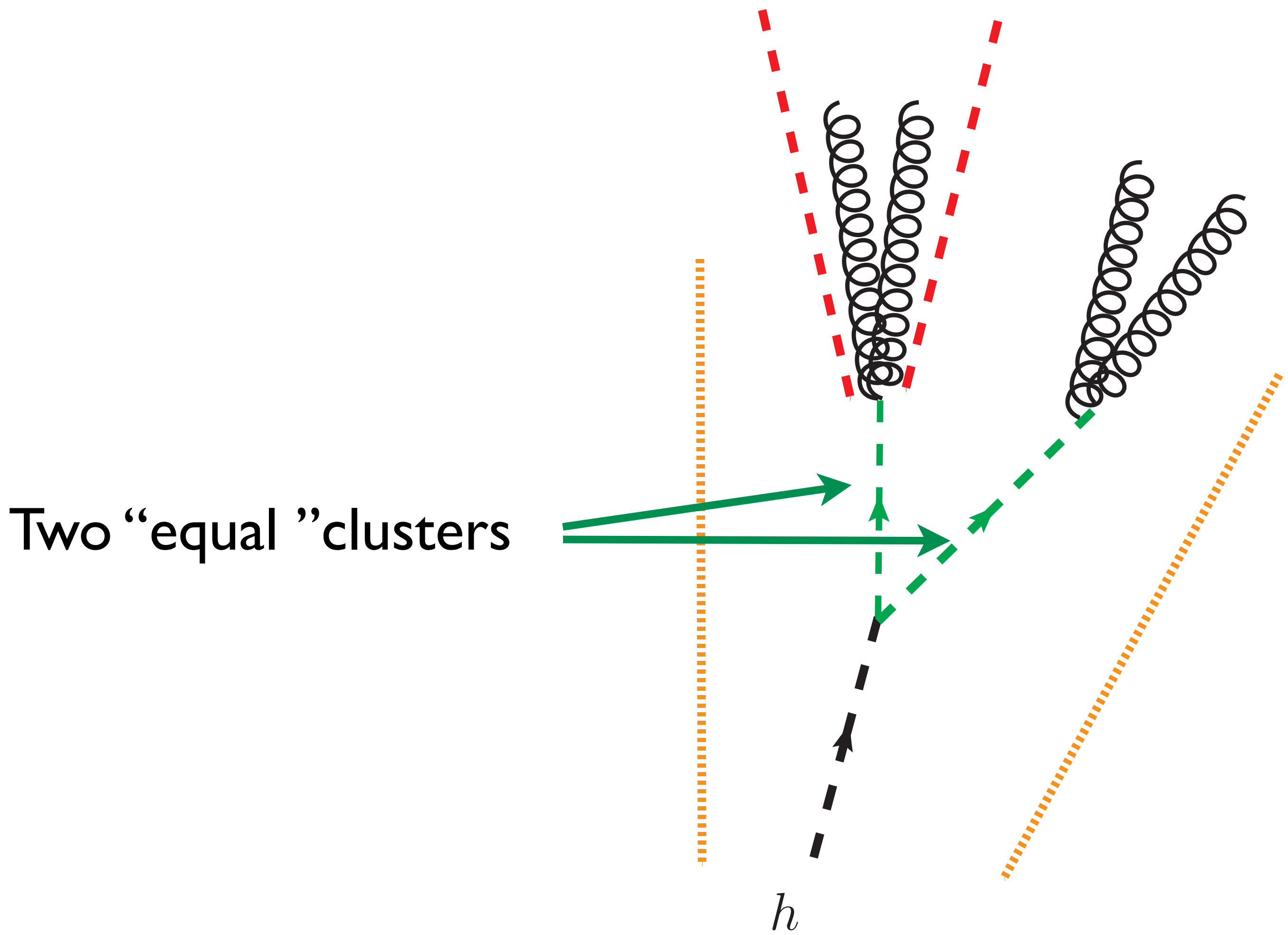
$h \rightarrow aa \rightarrow 4\tau,\ 4b,\ \bar{b}b\bar{\tau}\tau$

For example:
P. Graham, A. Pierce, J. Wacker, hep-ph/0605162
M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

Unbury the Higgs.

$h \to aa \to gggg$, "buried"!

For example:
B. Bellazzini, C. Csaki, A. Falkowski, A. Weiler, arXiv:0910.3210, arXiv:0906.3026



Boosting the Higgs.

$h \to aa \to c\bar{c}c\bar{c}$, "charmful"?

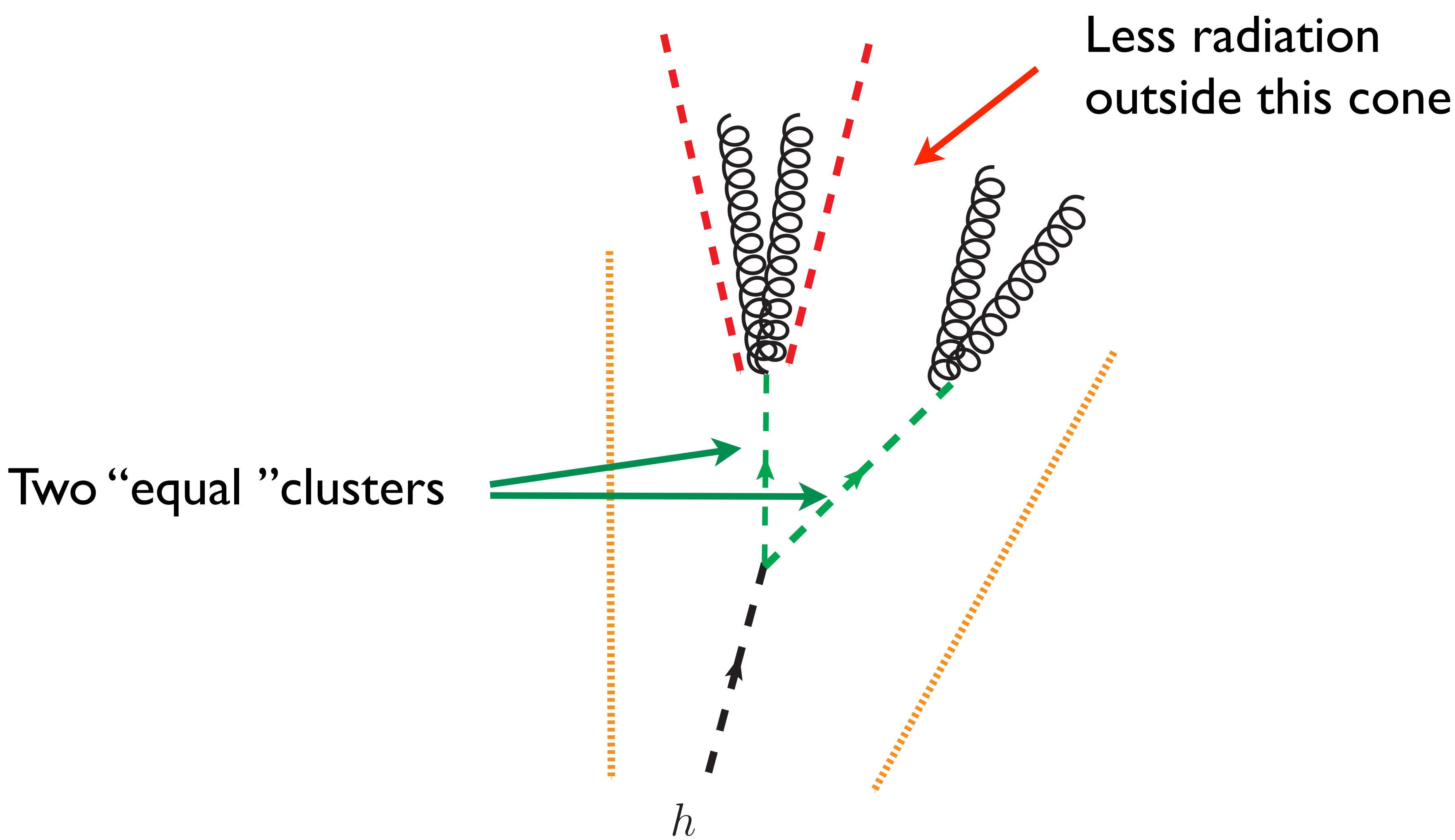
$h \to aa \to 4\tau,\ 4b,\ \bar{b}b\bar{\tau}\tau$

For example:
P. Graham, A. Pierce, J. Wacker, hep-ph/0605162
M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

Unbury the Higgs.

$h \to aa \to gggg$, "buried"!

For example:
B. Bellazzini, C. Csaki, A. Falkowski, A. Weiler, arXiv:0910.3210, arXiv:0906.3026



Boosting the Higgs.

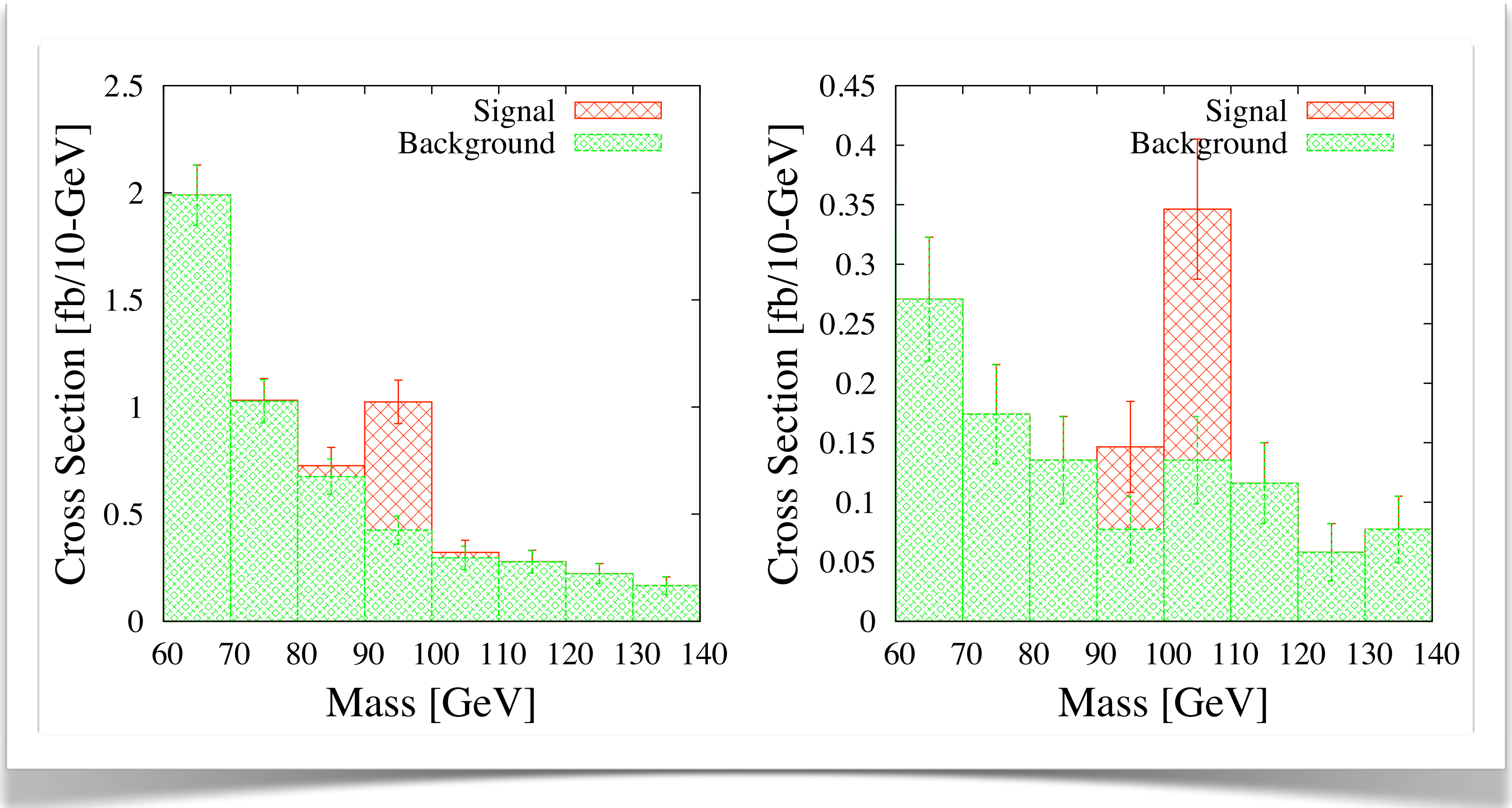
$h \to aa \to c\bar{c}c\bar{c}$, "charmful"?

$h \to aa \to 4\tau,\ 4b,\ \bar{b}b\bar{\tau}\tau$

For example:
P. Graham, A. Pierce, J. Wacker, hep-ph/0605162
M. Carena, T. Han, G. Huang, C. Wagner, arXiv:0712.2466

Encouraging results.

$$> 5\sigma \text{ at } 100 \text{ fb}^{-1}$$



W/Z+h

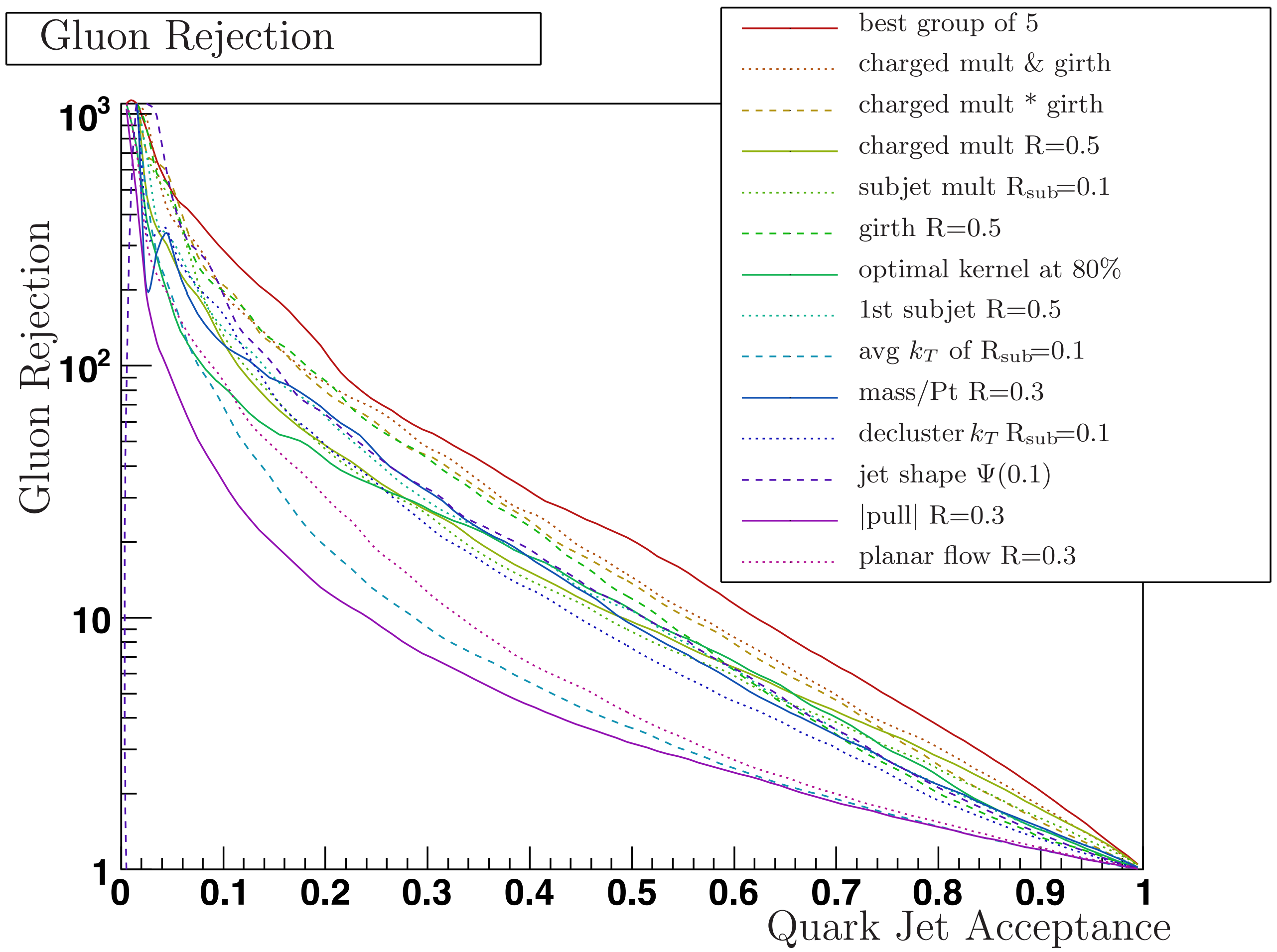
ttbar+h

Chen, Nojiri, Sreethawong, 1006.1151

Falkowski, Krohn, Shelton, Thalapillil, and LTW, 1006.1650

q/g of jet

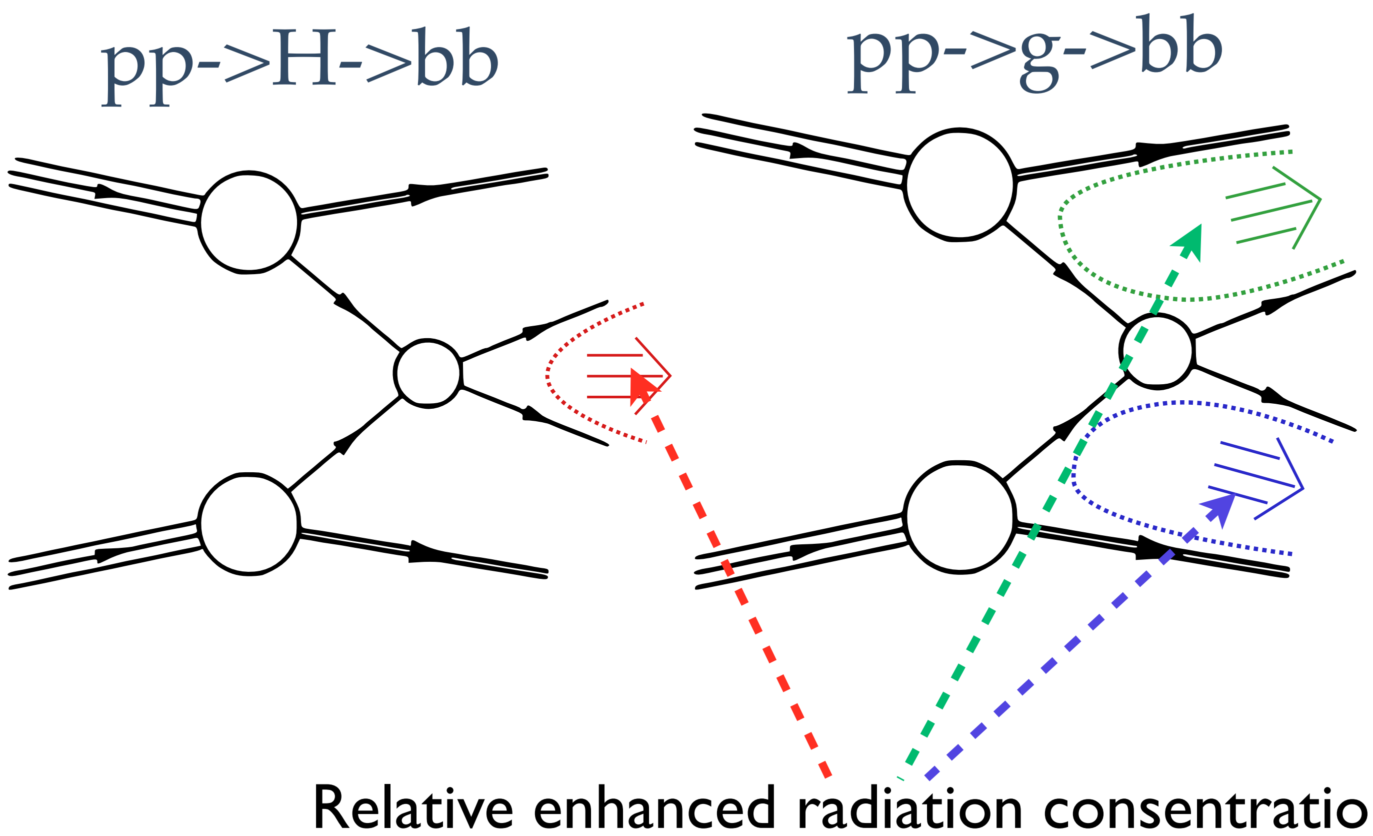
- Addition properties of jet: quark or gluon.
  - Can improve, for example, SUSY gluino signal...
- Quark gluon taggers.



Gallicchio, Schwartz, 1211.7038

Superstructure

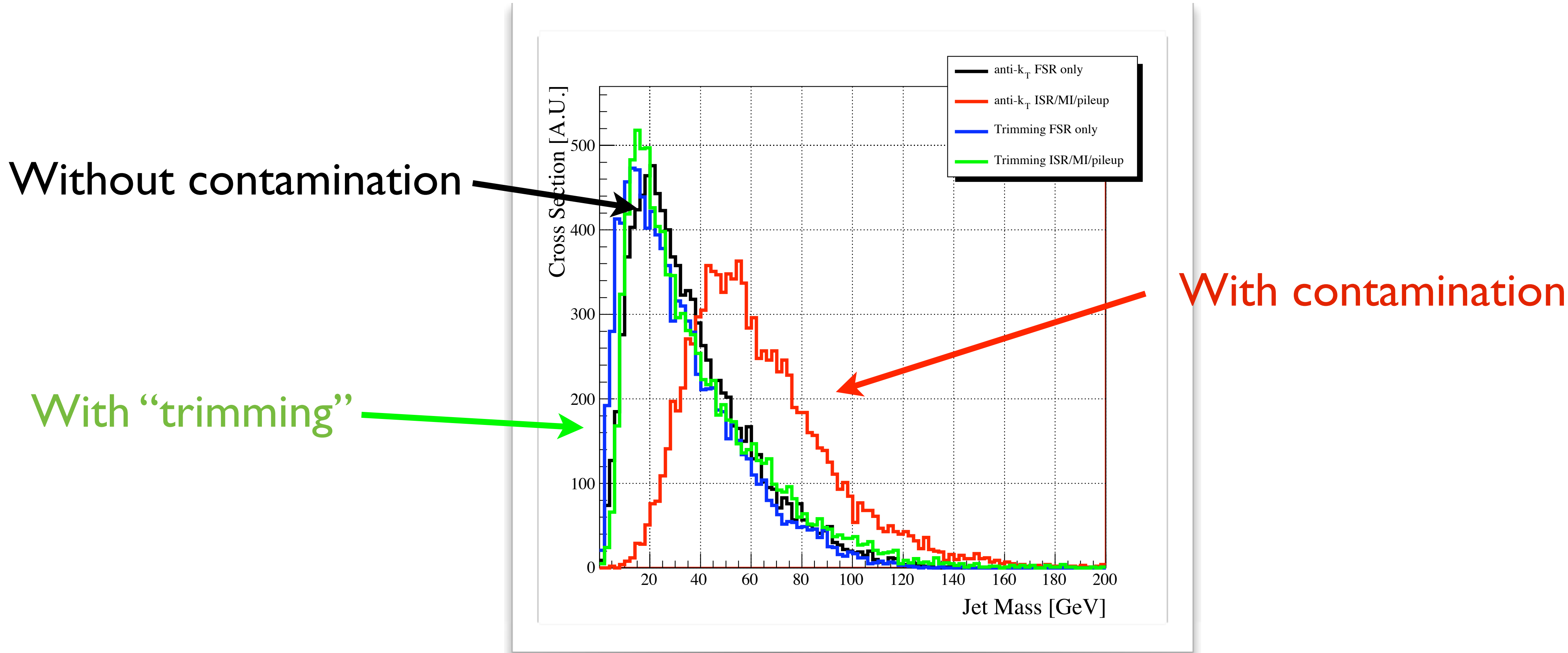
Gallicchio, Schwartz, 1001.5027



- Using more global information.
- Applications to other channels as well.
- Can this be done at LHC?

Substructure vs grooming

Jet mass: help from new jet algorithm



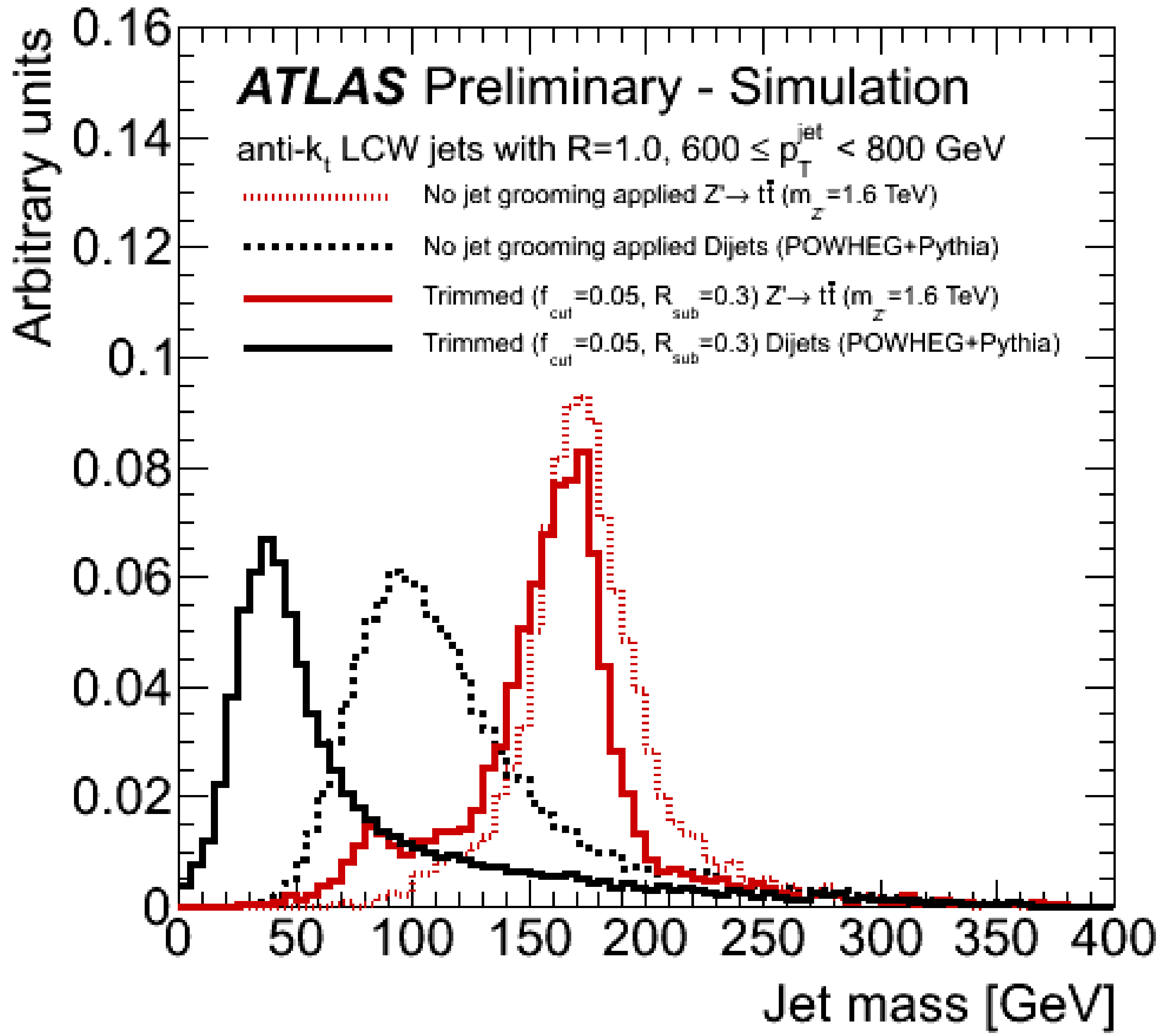
More faithful (smaller) jet mass for the background.

- Effect of radiation contamination on the jet mass

$$\langle \delta M^2 \rangle \simeq (\Lambda_{\rm soft} + p_T^{\rm ISR}) p_T^j \left(\frac{(\Delta R)^4}{4} + \ldots \right)$$

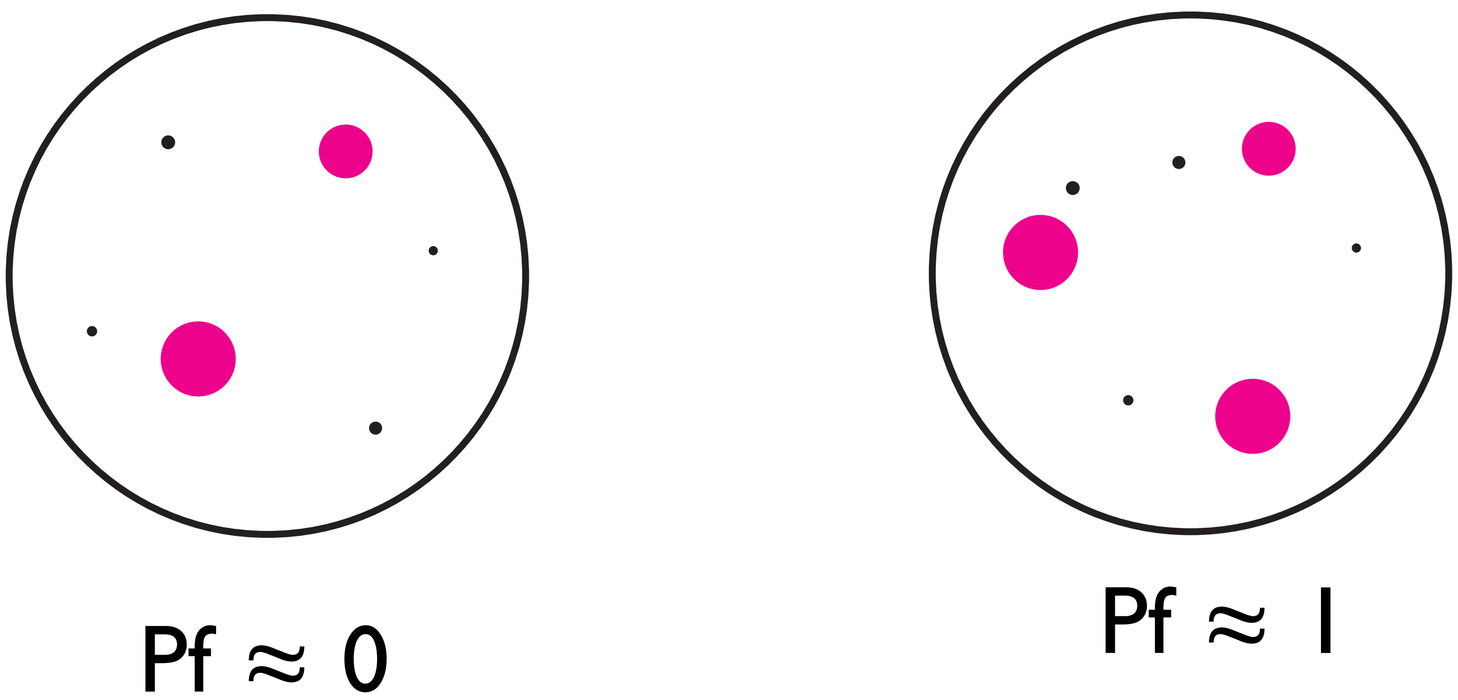
- Trimming gives large improvement by reducing effective jet size significantly.

Enhance signal significance



More jet shape variables.

- Top decay is more like 3-body. Span a "plane" perpendicular to the jet axis.
  - Transverse sphericity, or planar flow



$$I_w^{kl} = \sum_i w_i \frac{p_{i,k}}{w_i} \frac{p_{i,l}}{w_i}$$

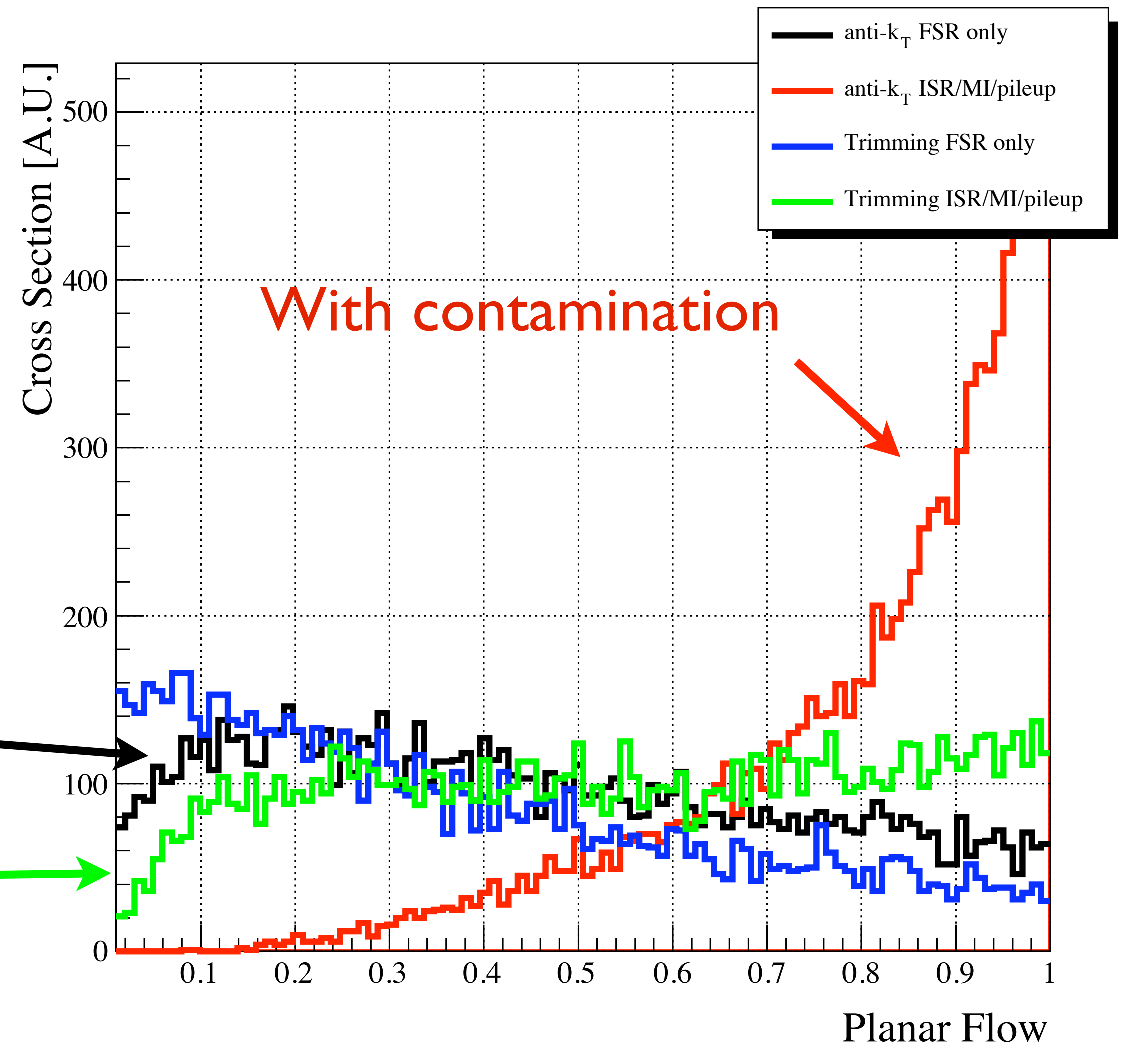
$$\lambda_1,\ \lambda_2 :\ 2 \text{ eigenvalues of } I_w^{kl}$$

$$\text{Pf} = \frac{4\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2}$$

Thaler and LTW, arXiv:0806.0023.
Almeida, Lee, Perez, Sterman, Sung, Virzi, arXiv:0807.0234

Grooming gives better jet shape

Planar flow



With no contamination

With “trimming”

- Can be used to further improve top tagging. An additional factor of several possible.
- Interesting to compare with improved QCD calculation, using modern technologies such as SCET.

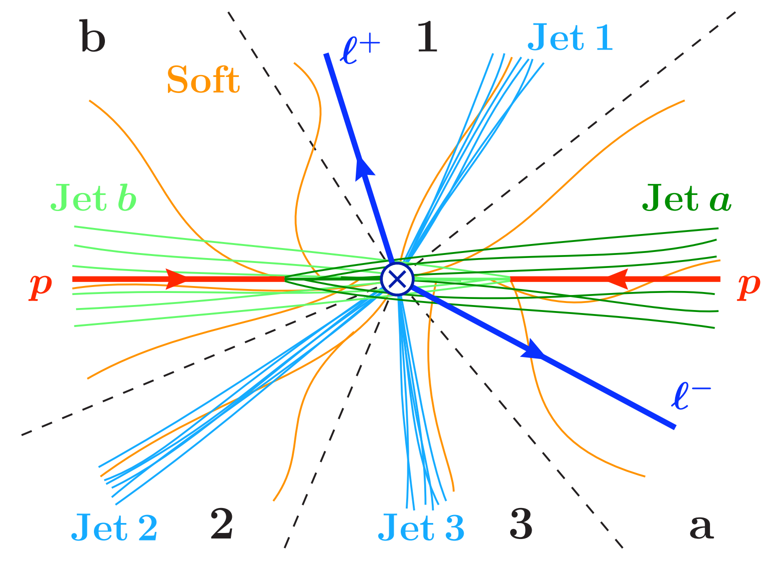
New developments: N-(sub)jettiness

Stewart, Tackmann, Waalewijn, 1004.2489

$$\mathcal{T}_N = \sum_k |\vec{p}_{kT}| \min\{d_a(p_k), d_b(p_k), d_1(p_k), d_2(p_k), \ldots, d_N(p_k)\}$$
$$\equiv \mathcal{T}_N^a + \mathcal{T}_N^b + \mathcal{T}_N^1 + \cdots + \mathcal{T}_N^N$$

N-jet like event

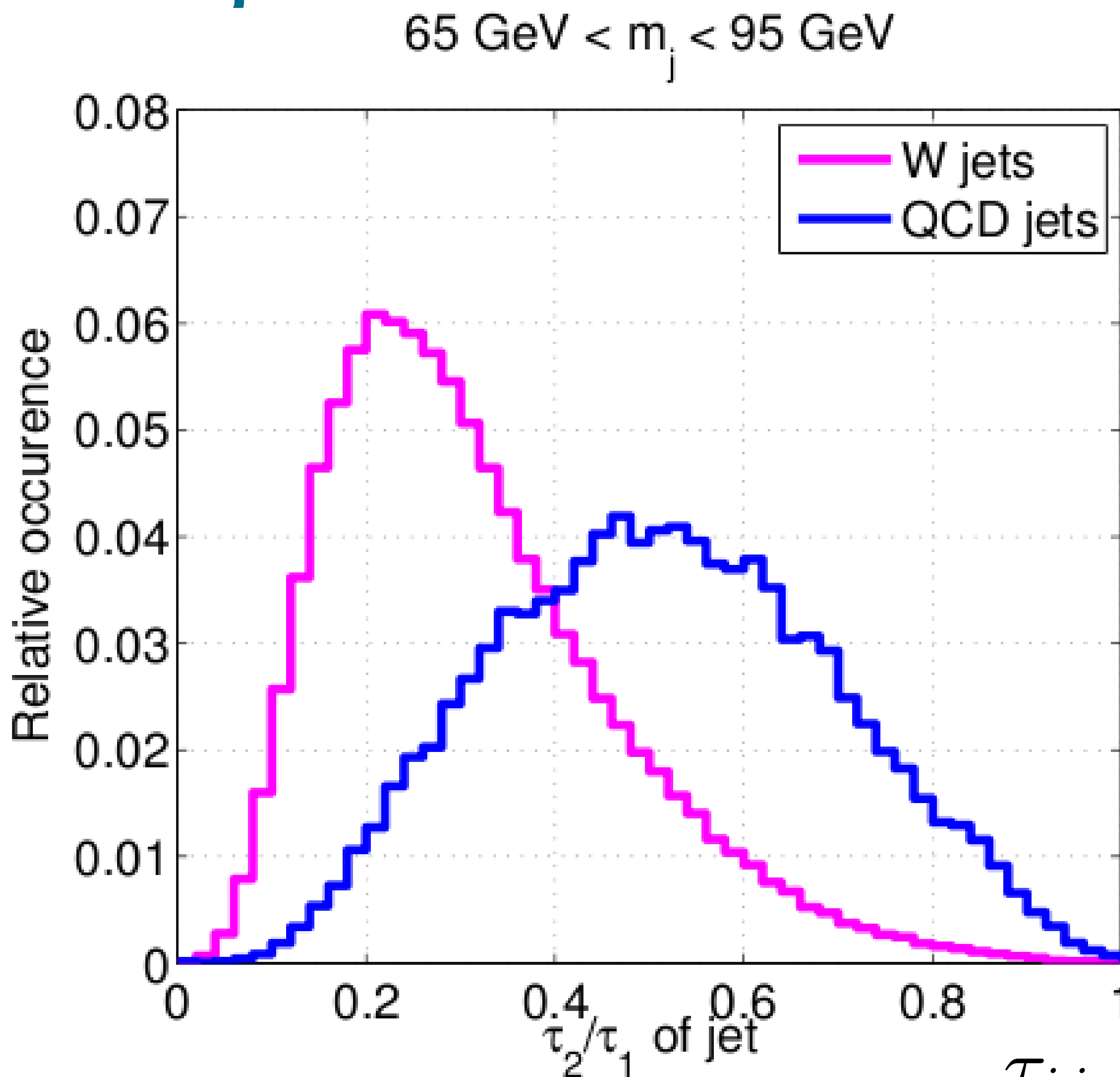
$$\tau_N \longrightarrow 0$$



- Using event shape instead of clustered jets.
- Allowing better QCD (SCET) treatments.
- Examples: nsub-jettiness. top tagger, etc.

N-subjettiness

Thaler, Tilburg, 1011.2268, 1108.2701

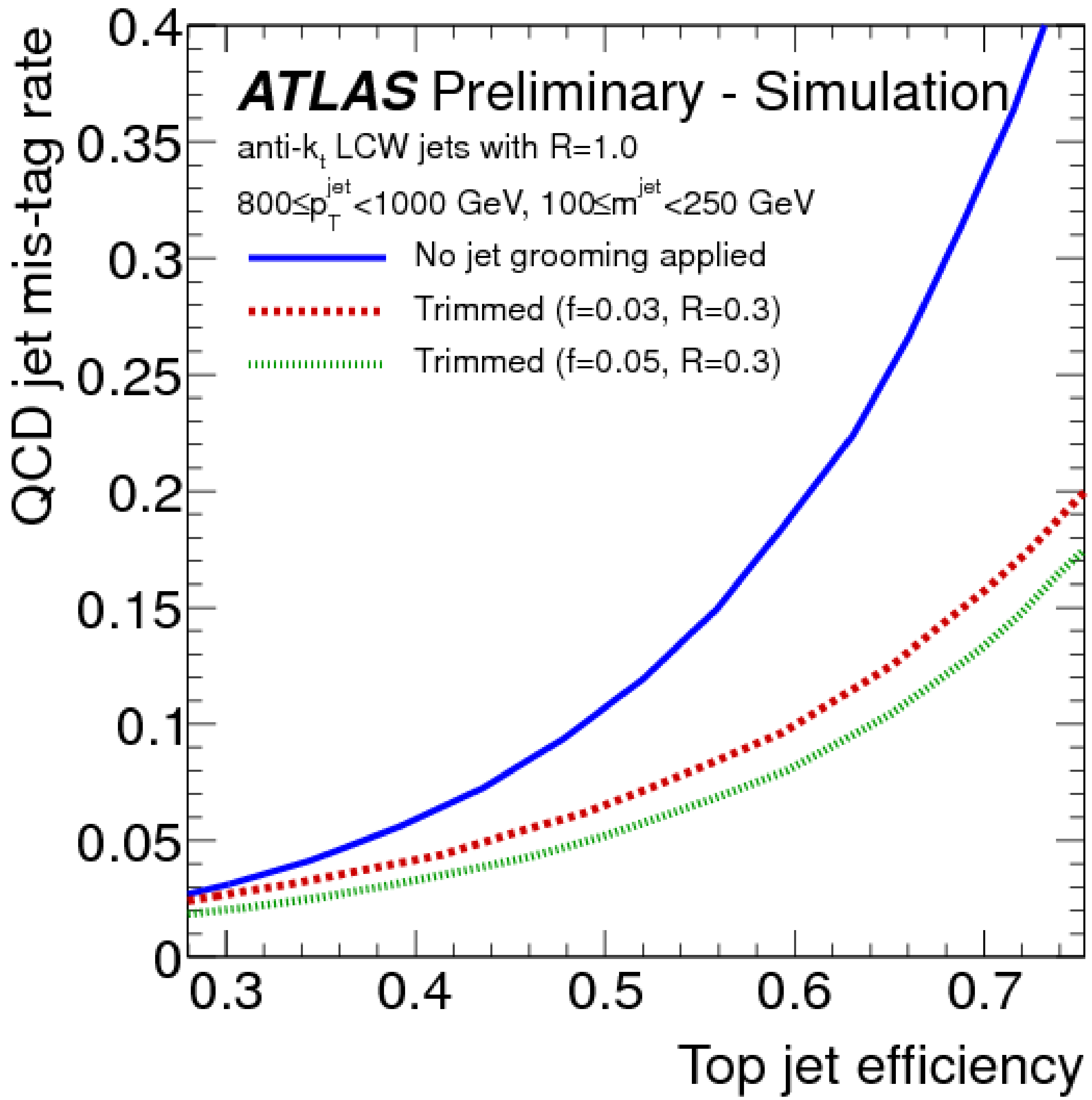


$$\tau_{ij} = \tau_i/\tau_j$$

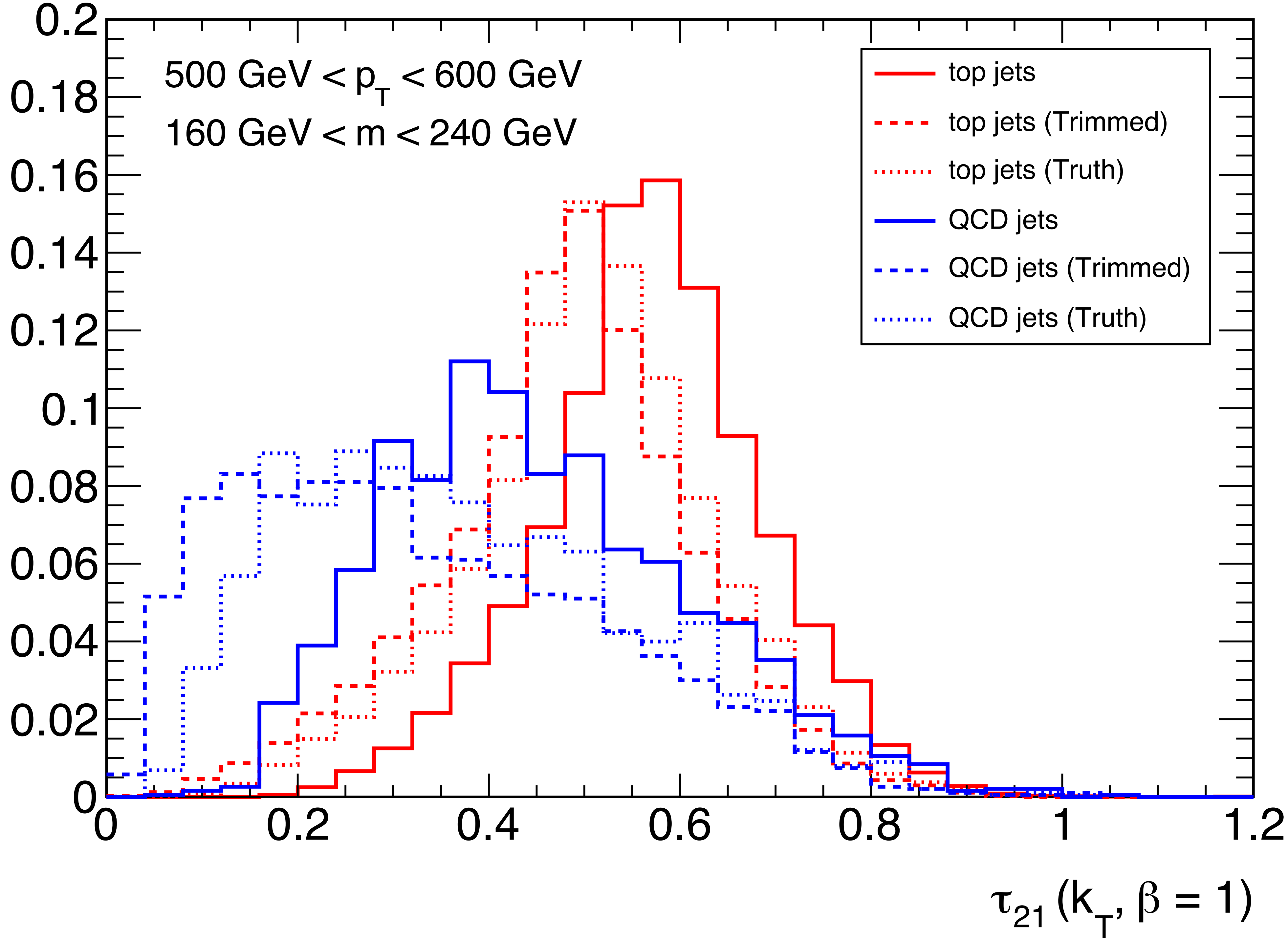
$$\tilde{\tau}_N^{(\beta)} = \frac{1}{d_0} \sum_i p_{T,i} \min\left\{(\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta\right\}.$$

In top tagging vs grooming.

800 < pT < 1000 GeV (100 < M < 250 GeV)



grooming vs nsub-jettiness



closer to truth, S/B?

Conclusion

- Many new developments, applied to actual exp searches already.
- Most results based very simple intuitions of QCD radiation.
- QCD calculation of various jet shape/structure/grooming still not quite catching up.
- MC + measurement from data still crucial.
- The current status may well change at higher energy and luminosity.

Conclusions

- Better understanding the properties of more shape variables.
- How would performance change with luminosity energy.
  - Effectiveness of substructure variables.
  - The agreement between MC and groomed jets.
  - Use of better QCD calculation. Better understanding the observables (well under control?)

Boosted by boosting

Boosted by boosting

- ttbar + h. Plehn, Salam, Spannowsky, 0910.5472

Boosted by boosting

- ttbar + h. Plehn, Salam, Spannowsky, 0910.5472
- Combining several grooming methods. Soper, Spannowsky, 1005.0417

Boosted by boosting

- ttbar + h. Plehn, Salam, Spannowsky, 0910.5472
- Combining several grooming methods. Soper, Spannowsky, 1005.0417
- Using N-subjettiness. Kim 1011.1493, Thaler, Tilburg 1011.2268

Boosted by boosting

- ttbar + h. Plehn, Salam, Spannowsky, 0910.5472
- Combining several grooming methods. Soper, Spannowsky, 1005.0417
- Using N-subjettiness. Kim 1011.1493, Thaler, Tilburg 1011.2268
- boosted H $\to\tau\tau$, di-tau jet. Katz, Son, Tweedie, 1011.4523

Boosted by boosting

- ttbar + h. Plehn, Salam, Spannowsky, 0910.5472
- Combining several grooming methods. Soper, Spannowsky, 1005.0417
- Using N-subjettiness. Kim 1011.1493, Thaler, Tilburg 1011.2268
- boosted H $\to \tau\tau$, di-tau jet. Katz, Son, Tweedie, 1011.4523
- Heavy H$\to$ ZZ, boosted Z jet. Hackstein, Spannowsky, 1008.2202

Substructure can also be useful for

- From top/W/Z/Higgs from NP decay, early LHC prospects.
  - Resonance ttbar.
  - SUSY. — Kribs, Martin, Roy, Spannowsky , 0912.4731, 1006.1656
  - Top partner to Higgs. — Kribs, Martin, and Roy, 1012.2886
  - Z' to WW, Zh... — Cui, Han, Schwartz, 1012.2077; Katz, Son, Tweedie, 1010.5253
- Boosted NP particles.
  - Neutralino + RPV — Butterworth, Ellis, Raklev, Salam, 0906.0728
  - Boosted gluino from squark. — Fan, Krohn, Mosteiro, Thalapillil, 1102.0302

Substructure can also be useful for

- From top/W/Z/Higgs from NP decay, early LHC prospects.
  - Resonance ttbar.
  - SUSY. Kribs, Martin, Roy, Spannowsky , 0912.4731, 1006.1656
  - Top partner to Higgs. Kribs, Martin, and Roy, 1012.2886
  - Z' to WW, Zh... Cui, Han, Schwartz, 1012.2077
    Katz, Son, Tweedie, 1010.5253
- Boosted NP particles.
  - Neutralino + RPV Butterworth, Ellis, Raklev, Salam, 0906.0728
  - Boosted gluino from squark. Fan, Krohn, Mosteiro, Thalapillil, 1102.0302
- And

Substructure can also be useful for

- From top/W/Z/Higgs from NP decay, early LHC prospects.
  - Resonance ttbar.
  - SUSY. Kribs, Martin, Roy, Spannowsky , 0912.4731, 1006.1656
  - Top partner to Higgs. Kribs, Martin, and Roy, 1012.2886
  - Z' to WW, Zh... Cui, Han, Schwartz, 1012.2077; Katz, Son, Tweedie, 1010.5253
- Boosted NP particles.
  - Neutralino + RPV Butterworth, Ellis, Raklev, Salam, 0906.0728
  - Boosted gluino from squark. Fan, Krohn, Mosteiro, Thalapillil, 1102.0302
- And
  - boosted di-jet resonance of mass 150 GeV.

Substructure can also be useful for

- From top/W/Z/Higgs from NP decay, early LHC prospects.
  - Resonance ttbar.
  - SUSY. Kribs, Martin, Roy, Spannowsky , 0912.4731, 1006.1656
  - Top partner to Higgs. Kribs, Martin, and Roy, 1012.2886
  - Z' to WW, Zh... Cui, Han, Schwartz, 1012.2077; Katz, Son, Tweedie, 1010.5253
- Boosted NP particles.
  - Neutralino + RPV Butterworth, Ellis, Raklev, Salam, 0906.0728
  - Boosted gluino from squark. Fan, Krohn, Mosteiro, Thalapillil, 1102.0302
- And
  - boosted di-jet resonance of mass 150 GeV.
  - 2 TeV axi-gluon decaying into boosted tops.

Simulations by theorists.

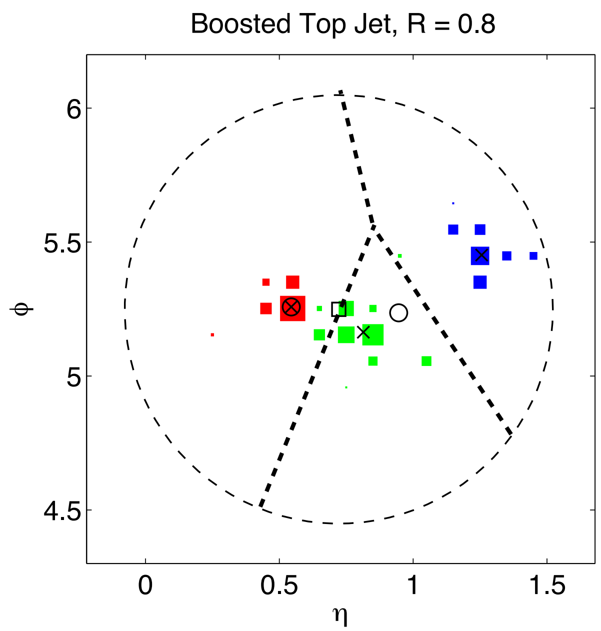
- Parton level Signal and background:
  - Madgraph, Alpgen, ...
- ME+PS matching, UE, Pileup:
  - Pythia, Herwig, Sherpa, ...
- Some detector effect, in particular, granularity 0.1x0.1
  - PGS, Delphes, "by hand".
- Jet tools.
  - Fastjet. http://www.lpthe.jussieu.fr/~salam/fastjet/
  - SpartyJet http://projects.hepforge.org/spartyjet/

New jet tools: going forward

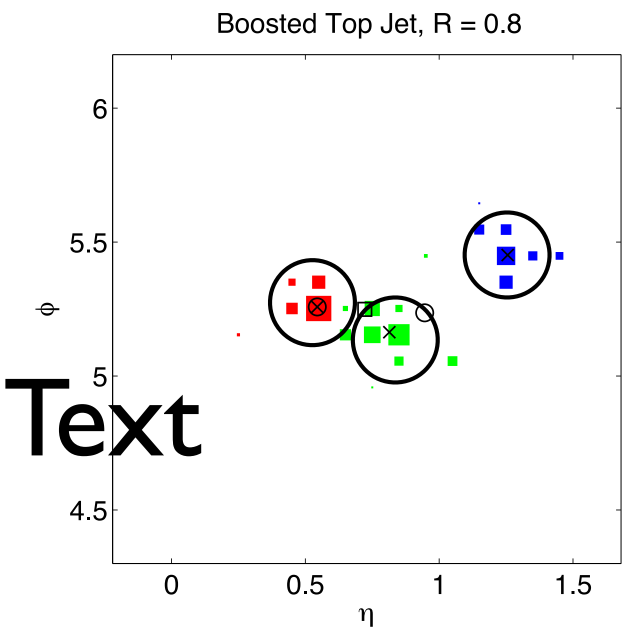
- Room for new ideas
  - More flexible, dynamical algorithms.
- Better theoretical understanding
  - Impact of grooming techniques
  - size, shape...
- Testing with early LHC data, 7 TeV, O(fb^{-1})
  - NP searches using jet substructure: SUSY, Zprime...
  - Training on known particles, boosted W/Z/t.
  - Measuring jet shapes with LHC data.

Jet Grooming

Start with a large jet (capturing all decay products)



vs.



loss of mass resolution from out-of-cone effects

A grooming procedure removes radiation which is more likely to be contamination from the Underlying events and Pile-up

Improved mass resolution expected.

Mass drop/Filtering

Butterworth, Davison, Rubin, Salam
0802.2470

Start with high-p_t jet C/A, R=1.2

1. Undo last stage of clustering ($\equiv$ reduce R): $J \to J_1, J_2$
2. If $\max(m_1, m_2) \lesssim 0.67m$, call this a **mass drop** [else goto 1]
3. Require $y_{12} = \frac{\min(p_{t1}^2, p_{t2}^2)}{m_{12}^2} \Delta R_{12}^2 \simeq \frac{\min(z_1, z_2)}{\max(z_1, z_2)} > 0.09$ [else goto 1]

dimensionless rejection of asymmetric QCD branching

Filter the jet

- Reconsider region of interest at smaller $R_{\text{filt}} = \min[0.3, \frac{\Delta R_{j_1,j_2}}{2}]$
- Take **3** hardest subjets