Goings-on in Jets and Jet Substructure

Summary of the
Joint Theoretical-Experimental Workshop
University of Washington, January 11-15, 2010
http://silicon.phys.washington.edu/JetsWorkshop/

Jonathan Walsh
University of Washington

Joint Theoretical-Experimental Workshop on Jets and Jet Substructure

Goals:

- Bring theorists and experimentalists together in a forum to discuss challenging aspects of jet substructure:
  - Experimental challenges in using jet substructure
  - Comparison between jet substructure techniques
  - Development of common, useful software tools
  - Theoretical progress in understanding jets

Working groups:

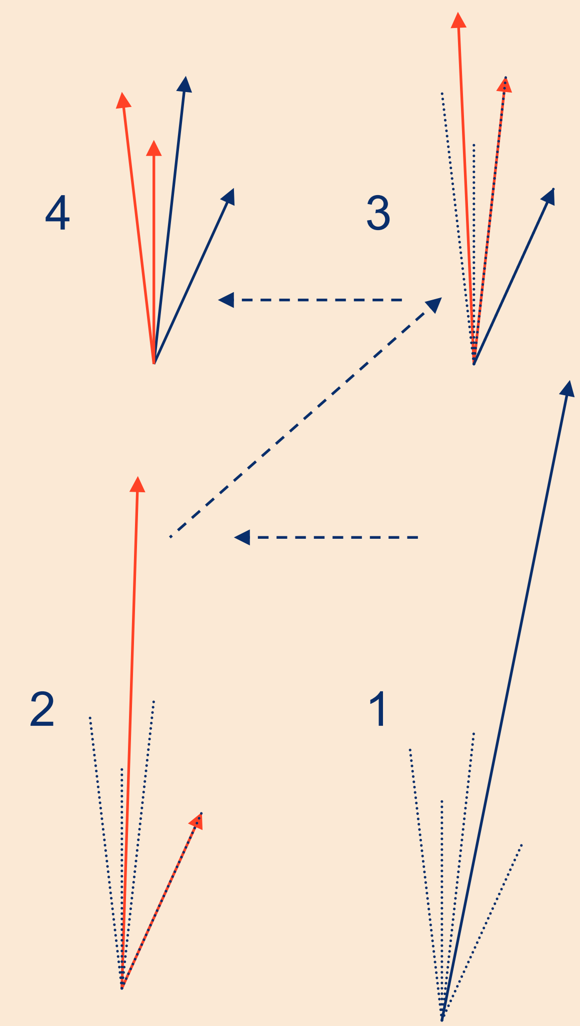
- Jet substructure techniques, implementation, calibration
- Software tools for jet studies
- Theory developments in understanding jets

Topics for this summary talk

- Experimental challenges in using jet substructure
  - Calibration of subjets, understanding effects of pileup and underlying event, MC issues
  - Examples from some jet substructure techniques
- Example comparison of jet substructure tools: trimming and pruning
  - Words on software development
- Jet superstructure
  - Complementary to jet substructure, may be broadly useful

Top Tagging : Kaplan et al

- Cluster jets with C-A
  - R : Function of sum pT of event
- Reverse cluster sequence
  - Throw out soft clusters
    - Fraction of hard jet pt < δp
- Repeat on clusters until one of:
  - Both subjets are harder than δp (PASS)
  - Both subjets are softer than δp (FAIL)
  - Subjets are too close (FAIL)
    - $|\delta\eta| + |\delta\varphi| < \delta r$
  - There is only one cell left (FAIL)
- Apply cuts:
  - Total mass consistent with m_{top}
  - 2 subjets consistent with m_W
  - W helicity consistent with top decay



Sal Rappoccio

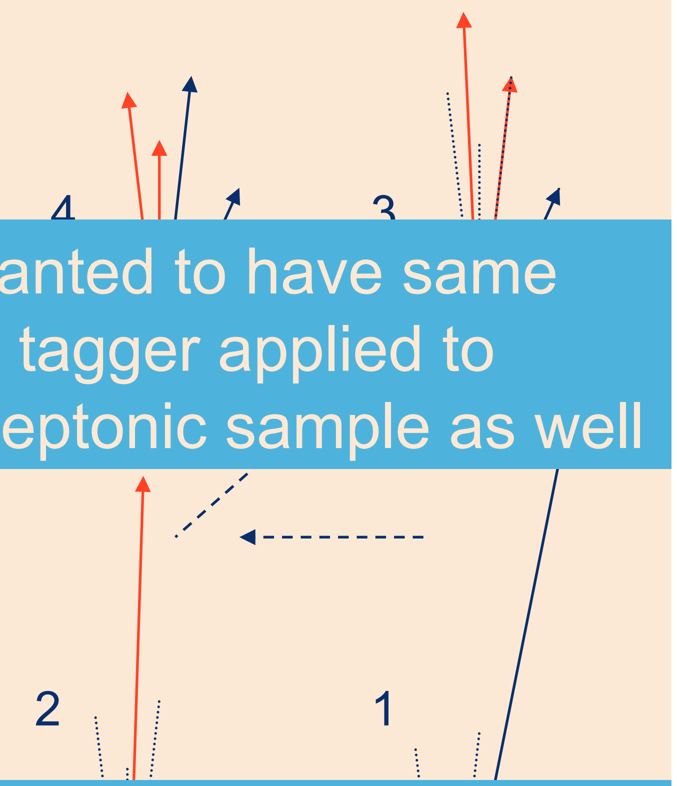
U of Washington Jet Workshop

Top-tagger designed to identify hard subjets in CA jet substructure, identify tops from these subjets

CMS study on top-tagging (from Johns Hopkins group) highlights substructure issues - and successes in implementing jet substructure techniques

Top Tagging : CMS

- Cluster jets with C-A
  - R : ~~Function of sum pT of event~~
- Reverse cluster sequence
  - Throw out soft clusters
    - Fraction of hard jet pt < δp
- Repeat on clusters until one of:
  - Both subjets are harder than δp (PASS)
  - Both subjets are softer than δp (FAIL)
  - ~~Subjets are too close (FAIL)~~
    - ~~$|\delta\eta| + |\delta\varphi| < \delta r$~~
  - There is only one cell left (FAIL)
- Apply cuts:
  - Total mass consistent with m_{top}
  - ~~2 subjets consistent with m_W~~
  - ~~W helicity consistent with top decay~~
  - Minimum mass pairing of subjets consis with W



Wanted to have same tagger applied to semileptonic sample as well

Found di-subjet min mass pairing more experimentally accessible (consulting with Brock)

Sal Rappoccio

U of Washington Jet Workshop

Procedure augmented by CMS for use - but key elements still retained

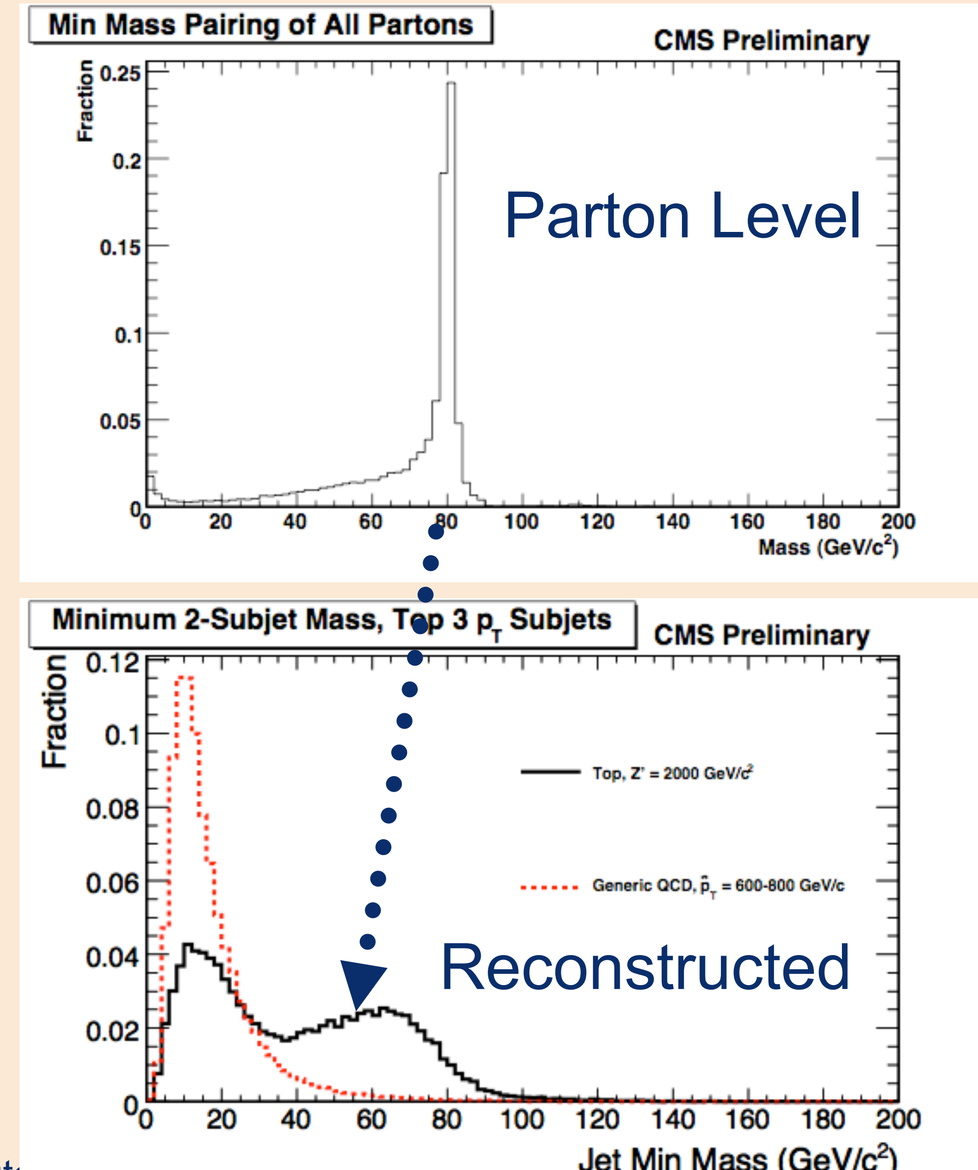
Expect to happen for all substructure methods - important to communicate between theory and experiment!

CMS study on top-tagging (from Johns Hopkins group) highlights substructure issues - and successes in implementing jet substructure techniques

Top Tagging : CMS

- Discriminate **top jets** against **non-top jets**
  - Top mass
  - **W mass ~ min di-subjet mass**

bq	j1+j2
bq' →	j1+j3
qq'	j1+j3



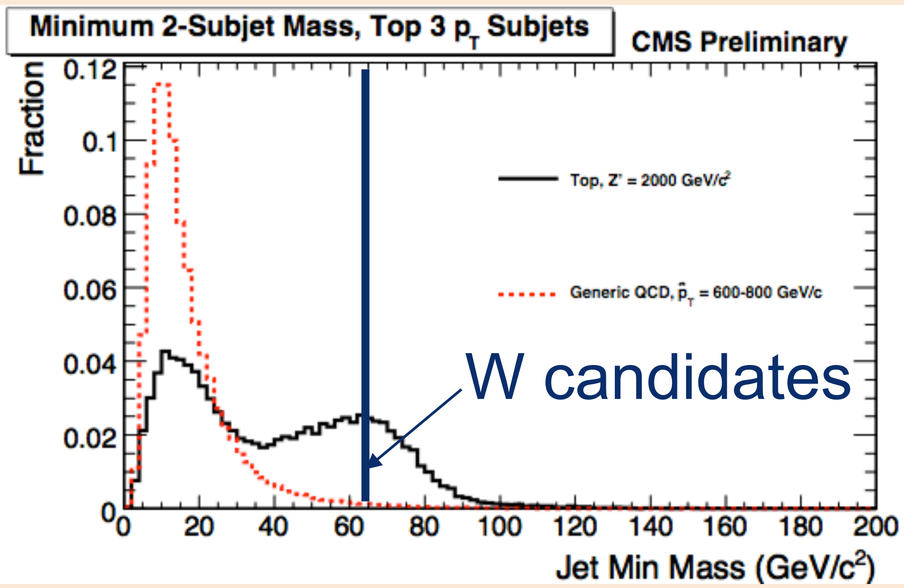
Top reconstruction efficiency does not suffer from decreased resolution! - cuts and background rejection changes, but tag rate robust

Parton level kinematics very different from detector level - worry that correlations in the substructure can be lost

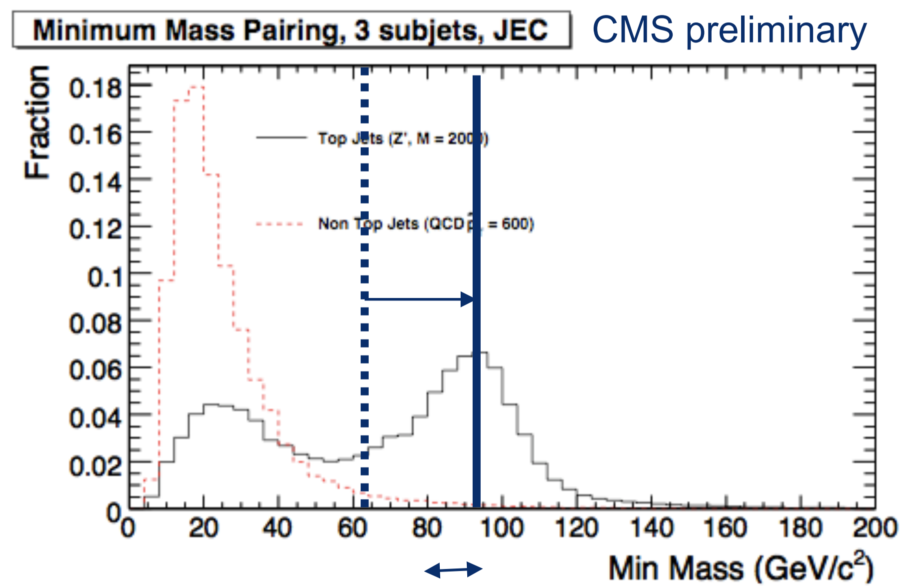
Experimental Challenges for Subjets

How can we calibrate subjets?

Uncorrected



Corrected



Significant over-correction

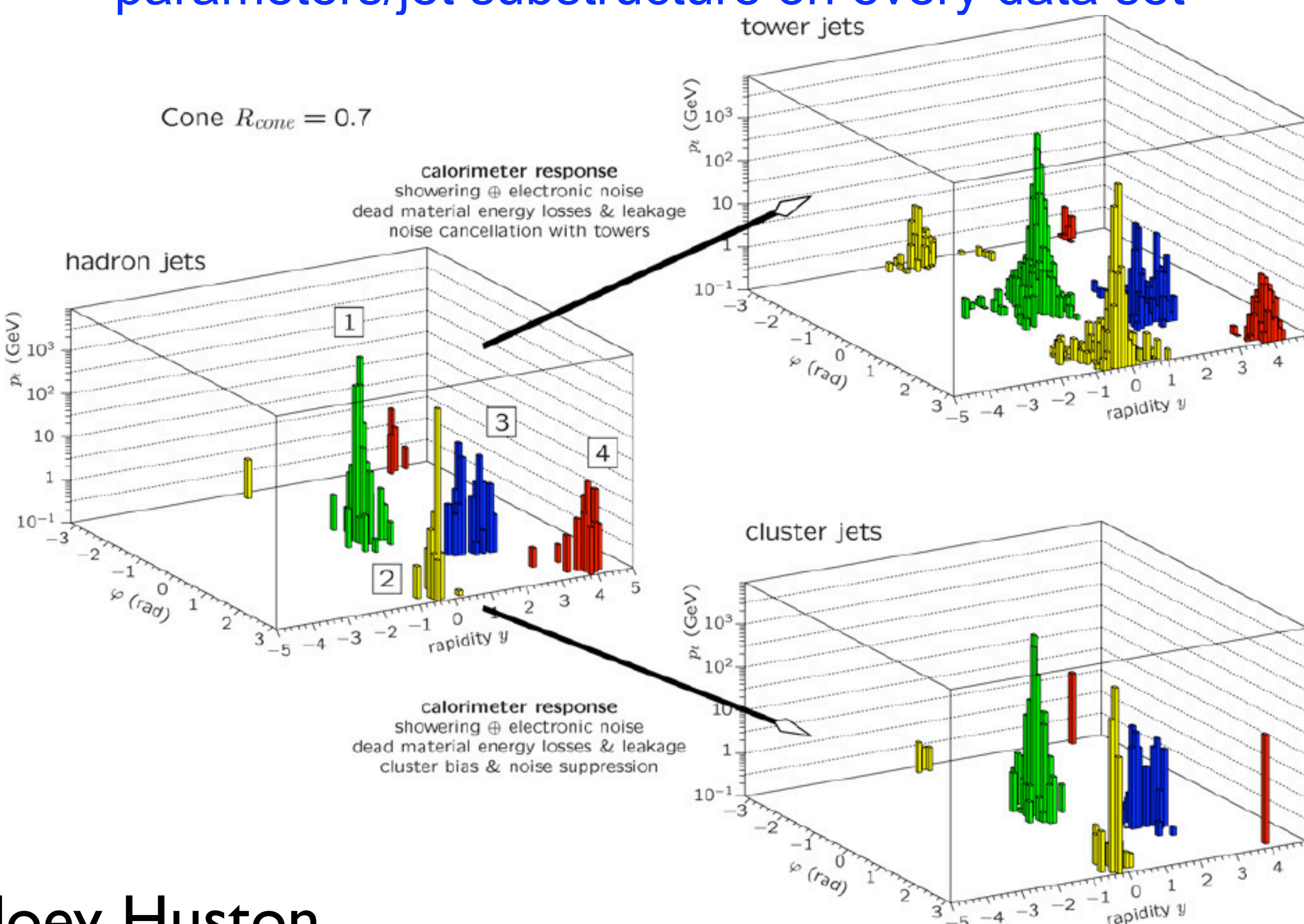
Can calibration through standard candle channels work? e.g., EW+jet; can also add in heavy flavor tags to jets

Jet energy scale correction does not apply to subjets - e.g., soft physics has been removed, smaller area

Possible that improved theory calculations can help here?

ATLAS jet reconstruction

- Using calibrated topoclusters, ATLAS has a chance to use jets in a dynamic manner not possible in any previous hadron-hadron calorimeter, i.e. to examine the impact of multiple jet algorithms/ parameters/jet substructure on every data set



blobs of energy in the calorimeter correspond to 1/few particles (photons, electrons, hadrons); can be corrected back to hadron level

rather than jet itself being corrected

similar to running at hadron level in Monte Carlos

These "dynamic" techniques with calibrated topo-clusters could be tested with calibrated jets at known energies - Z decays or γ+jet

Joey Huston

Joey Huston emphasizes that using topo-clusters can offer a local calibration to study different algorithms and substructure methods on data sets

Issues surrounding UE and pile-up effects on jets

Both UE and pile-up will affect the jet finding, reconstruction, calibration and hence **jet substructure analyses**, the extent to which we can only begin to asses. Here are several of the issues we will talk about today:

- **Jet energy scale (JES) and mass distortions**
  - UE will augment (raise) the parton-level JES
  - Pile-up will augment (raise) the particle-level JES
  - Detector signal shaping for pile-up in ATLAS can also reduce the JES
- **Angular smearing**
  - Angular resolution is degraded by the presence of additional soft radiation
  - Uncorrelated pile-up affects also particle-level pointing resolution
  - "Back-reaction" occurs when this smearing is sever enough to add or subtract particles that otherwise would not have contributed
- **Diffuse and point-like contributions**
  - Diffuse background radiation will also flatten structure like planar flow
  - Point-like component of min. bias (MB) will hinder the diffuse approximation
- **Spurious jets**
  - Pure MB jets will affect jet multiplicity and isolation criteria
  - Close-by jets will increase, limiting JES precision even for sub-jets

What we think we know about pile-up and substructure

...and what we'd like to know

So far...

- Pile-up will be an issue for substructure analyses, but will it be dominant?
- At the luminosities for $H \rightarrow bb$: **yes**. Next year? No, but there will be enough to start testing tools
- Tracks provide a huge tool-box for finding, augmenting, and improving calorimeter jets, even without the need for 1-1 track-particle correlations

For the future

- Demonstrate–in data–the correlations and efficacy of track-based corrections
- Measure the level of correlation between pruning scales (z_{cut}, D_{cut}) and track based quantities
- Evaluate the scale(s) at which pile-up induces relevant sub-structures. Are they reducible?
- Can we use tracks to identify which jets and then prune?
- Can we use JVF to tune z_{cut}?

Comparing Pruning and Trimming

- Both techniques democratic regarding channel
- Comparison gives insight into how algorithms really operate on jet substructure
  - Both methods based on well motivated theory principles - but their action on real jets is complex
  - Identifying benefits and deficits can lead to better understanding of jet substructure
  - Experimentally implemented jet substructure tools likely to be some mix of techniques, adapted to the detector and analysis
- FastJet plugins exist - easy to construct comparative analyses
  - Driven improvements in analysis tools

Trimming

Krohn, Thaler, Wang
hep-ph/0912.1342

- Run a jet algorithm (kT) on a found jet, with angular scale R_{sub} smaller than the R value used for the initial jet
- Discard all subjets with pT < $f_{cut}\,\Lambda_{hard}$
  - f_{cut} is a dimensionless parameter, Λ_{hard} a hard scale
- The remaining subjets form the new (trimmed) jet
- Filtering: same procedure as trimming, but keep the N hardest subjets (instead of a pT cut on subjets)

Trimming designed for use on QCD jets
e.g., heavy particle decay (with low boost) to 2 jets

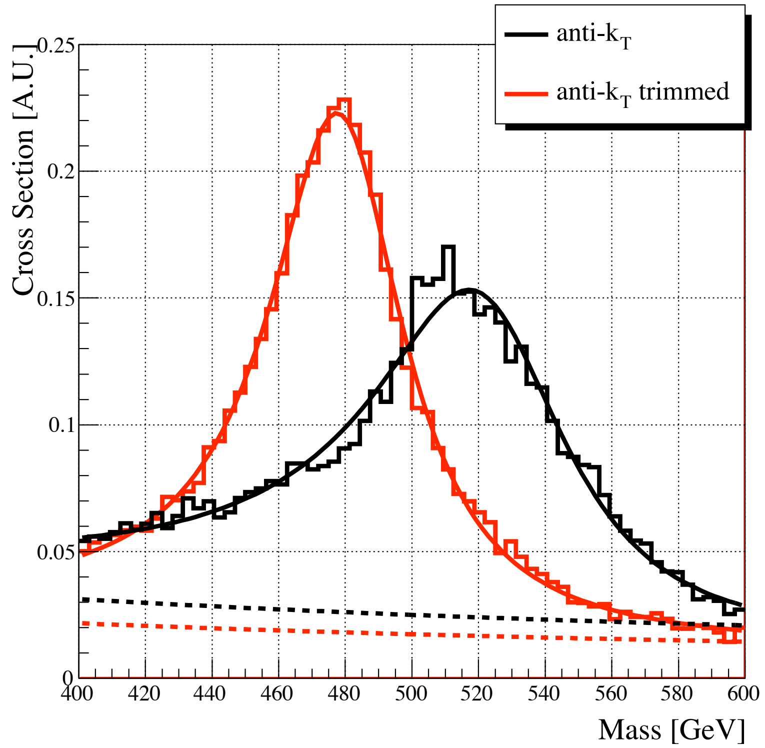
Trimming

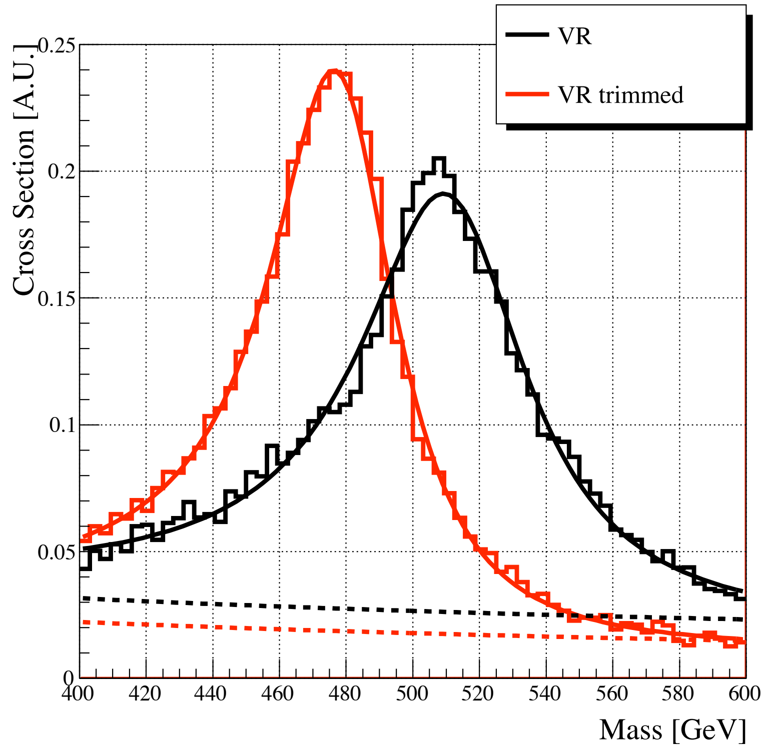
	Improvement	$f_{\rm cut}$, $N_{\rm cut}$	$R_{\rm sub}$	R_0, ρ	Γ [GeV]	M [GeV]
anti-k_T	-	-	-	1.0^*	71	522
anti-k_T (N)	40%	5^*	0.2^*	1.5^*	62	499
anti-k_T $(f,\, p_T)$	59%	$3\times10^{-2*}$	0.2	1.5	52	475
anti-k_T $(f,\, H)$	61%	$1\times10^{-2*}$	0.2	1.5	50	478
VR	30%	-	-	200^* GeV	62	511
VR (N)	53%	5	0.2	275^* GeV	53	498
VR $(f,\, p_T)$	68%	3×10^{-2}	0.2	300^* GeV	49	475
VR $(f,\, H)$	73%	1×10^{-2}	0.2	300^* GeV	47	478
Filtering	27%	2	$R_0/2$	1.3^*	61	515

$$S(m) = \alpha\left[\frac{1+\beta(m-M)}{(m^2-M^2)^2+\Gamma^2M^2}\right],$$

$$B(m) = \delta + \gamma/m,$$

$$\Delta \equiv S(M) = \frac{\alpha}{\Gamma^2M^2},$$



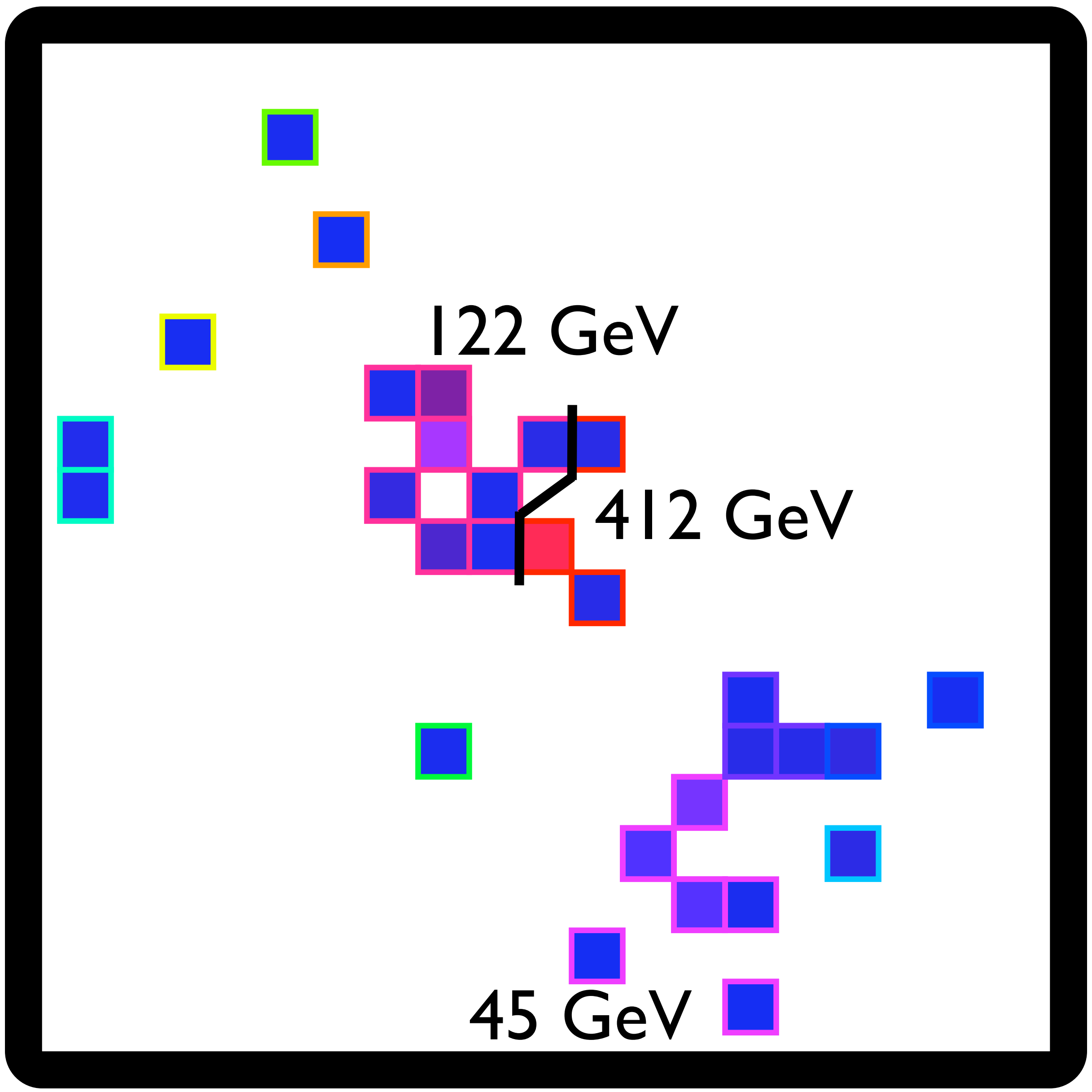


Trimming paper

Color octet scalar
$\phi \to gg$: 2j final state

Trimming yields improvements in pulling S from B

Trimming Example: QCD Jet



0.1x0.1 cells (y, ϕ)

Trimming designed to remove soft, isolated subjets through "pre-clustering"

11 subjets found
3 remain after pT cut:
$pT_{subjet} > 0.03\ pT_{jet} = 18$ GeV

pT of subjets removed:
7, 6, 4, 4, 1, 1, 1, 1 GeV

Pruning

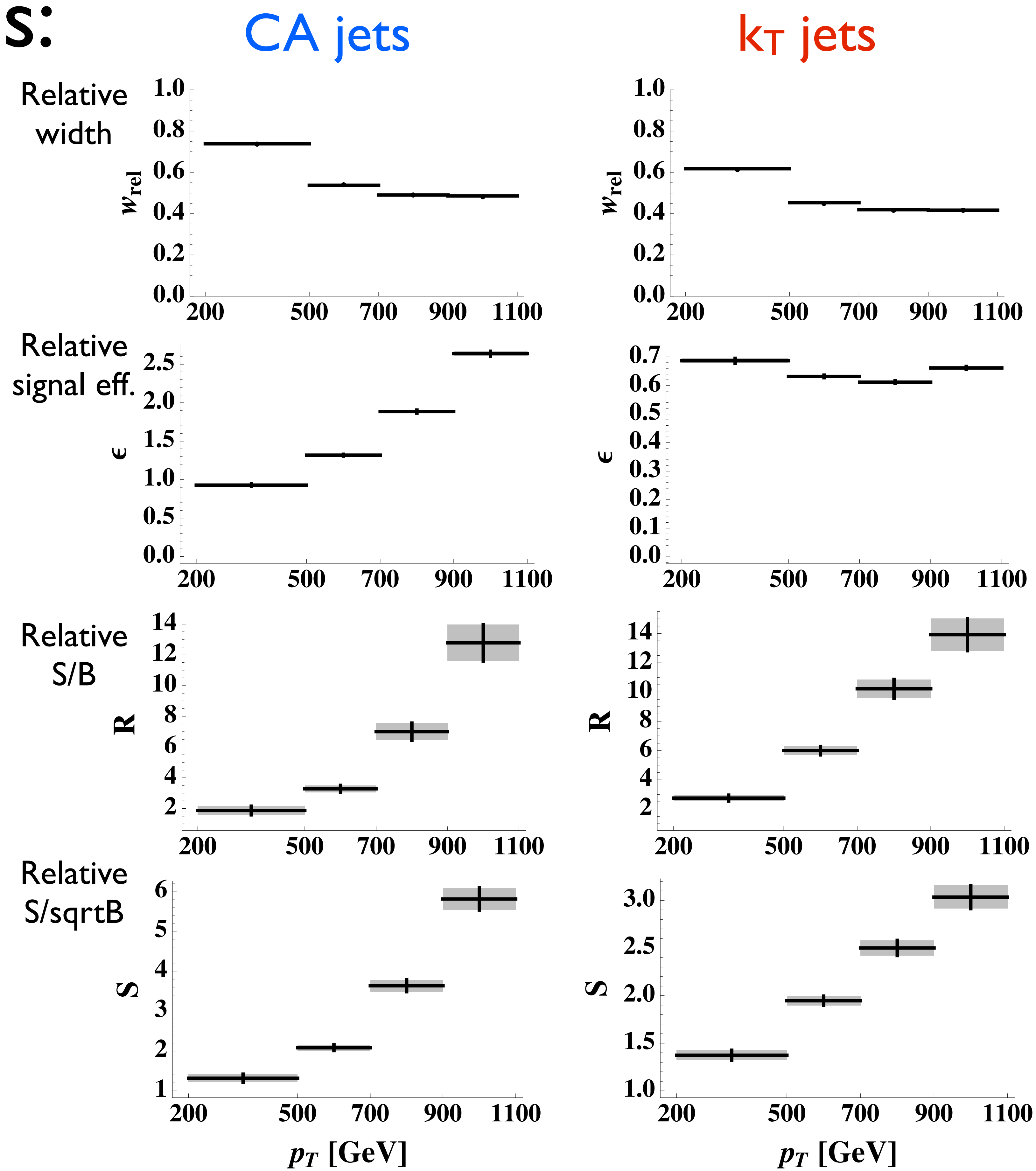
Ellis, Vermilion, JW
hep-ph/0903.5081
hep-ph/0912.0033

- Run CA or kT algorithm on the found jet, and at each recombination test if:
  - $z < z_{cut}$ and $\Delta R > D_{cut}$
    - $z_{cut} = 0.10$ for CA, 0.15 for kT; $D_{cut} = m_J/pT_J$
- If so, veto on the recombination - discard the lower pT daughter and continue
- The resulting jet is the new (pruned) jet

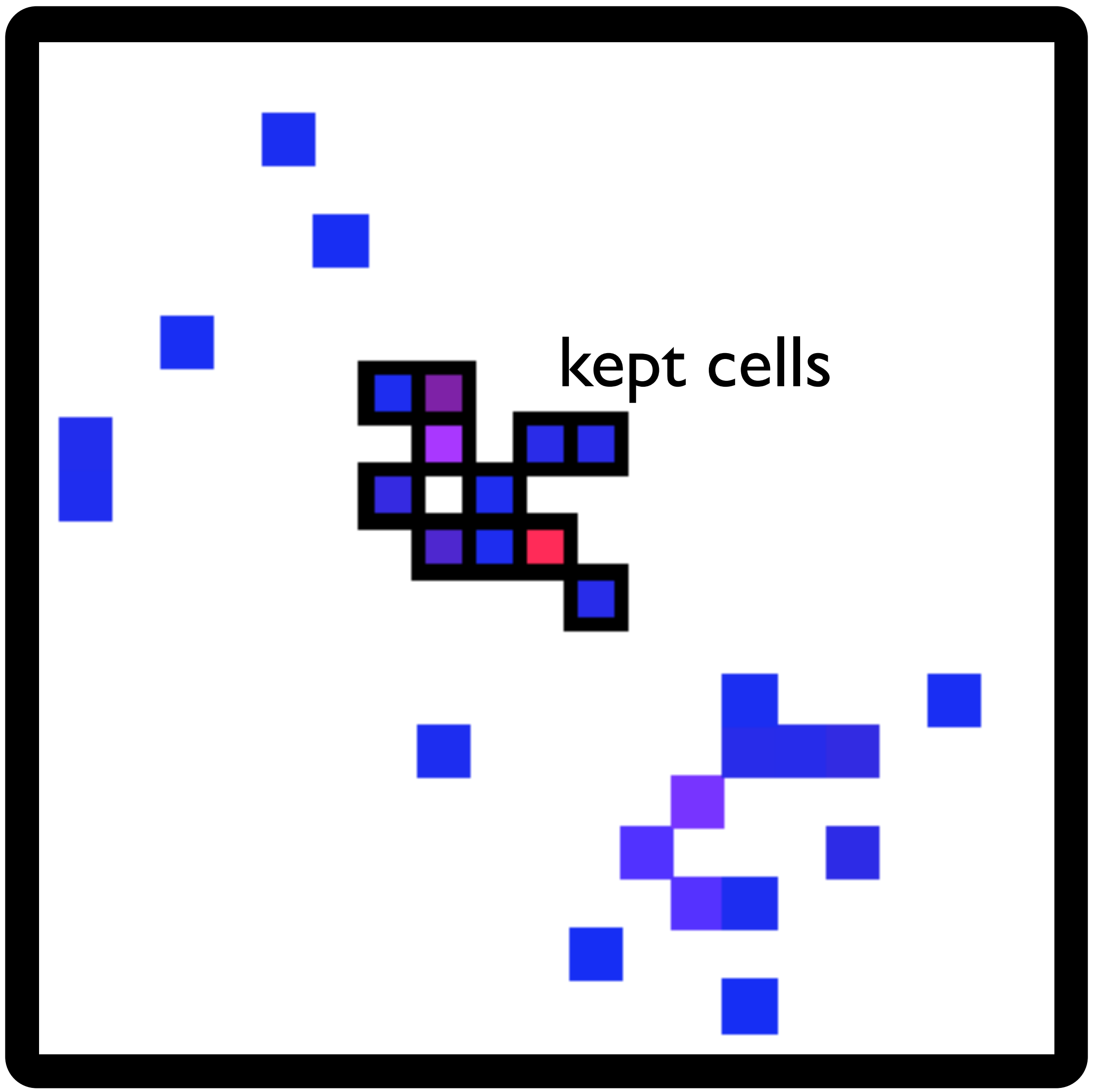
Pruning designed to identify jets from boosted heavy particles
i.e., heavy particle decay (with large boost) to 1 jet

Pruning on tops:

- Plot relative width, efficiency, S/B, S/sqrt(B) using a constant D = 1.0
- Variables are relative: improvements for pruning over not pruning
- Relative width is the improvement in mass resolution - at high p_T, pruned top mass width is 40% of unpruned
- Pruning shows consistent improvements, dramatically increasing at high p_T
- Statistical error bars shown



Pruning Example: QCD Jet



0.1x0.1 cells (y, ϕ)

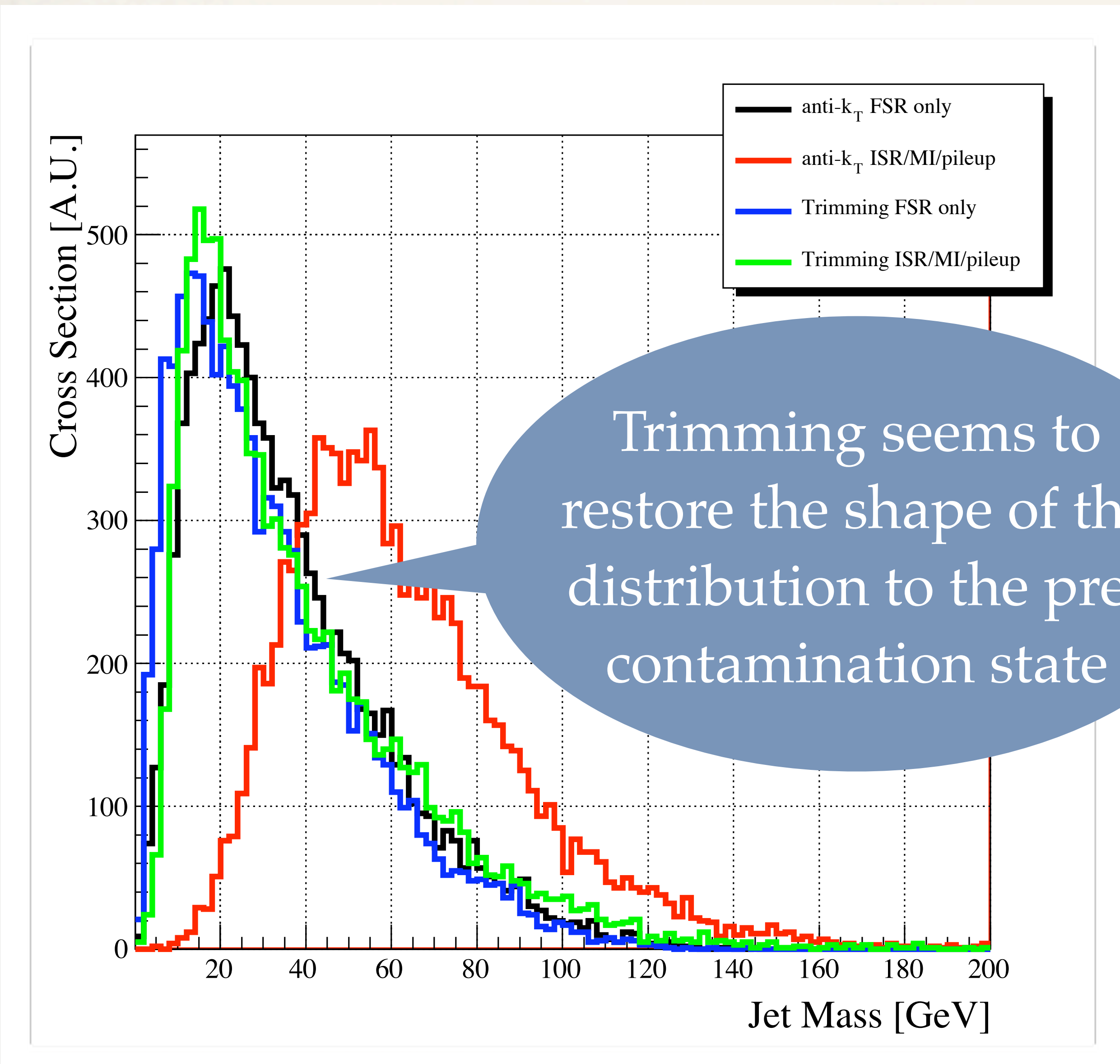
Pruning designed to remove soft, isolated parts of jet substructure during re-clustering

Keeps hard core of QCD jet, removes softer, wide angle radiation

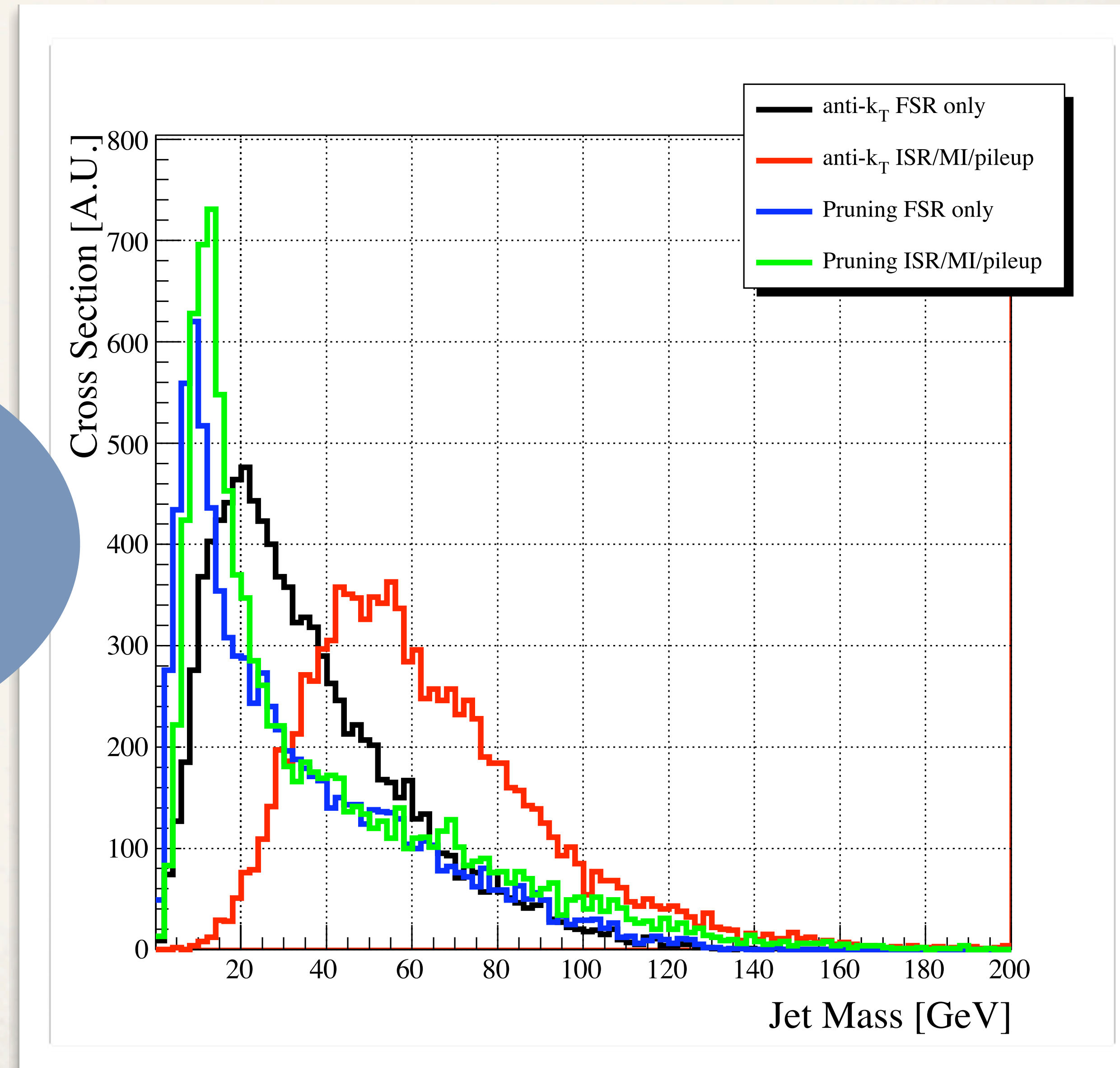
Quantify the difference by looking at jet shape variables

Jet Mass

Trimming is a smaller correction to the underlying FSR



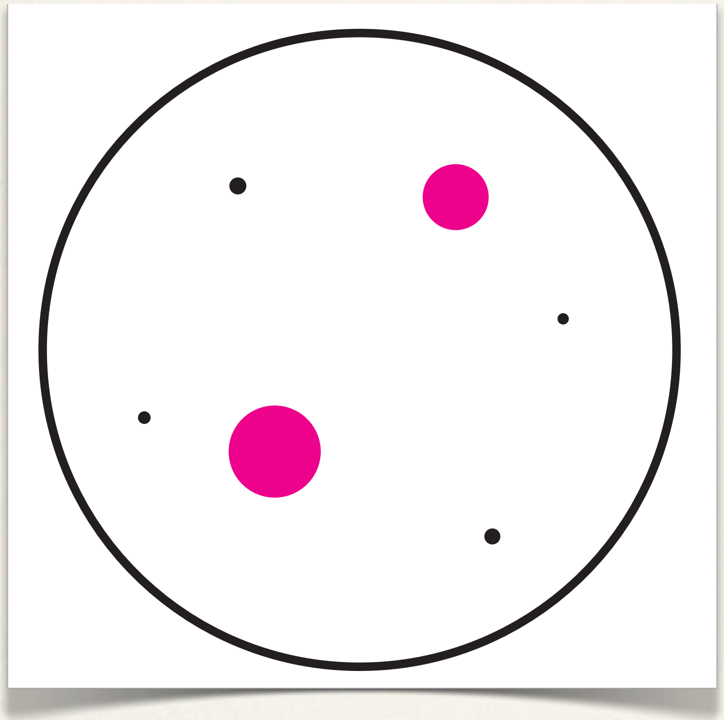
Trimming seems to restore the shape of the distribution to the pre-contamination state



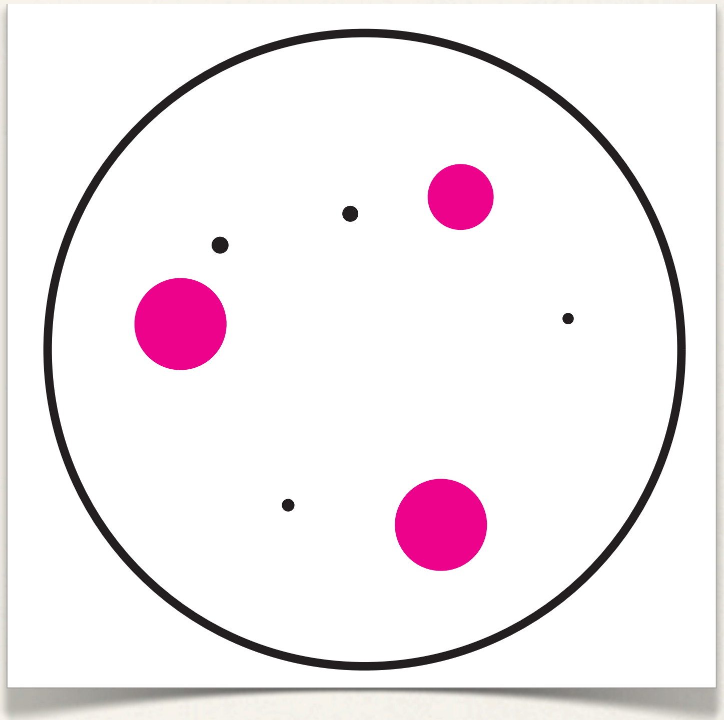
$\Delta R = 1.0$

David Krohn

Planar Flow



Pf ≈ 0



Pf ≈ 1

- Planar flow is another jet shape designed to measure how plane/pencil like a jet is.

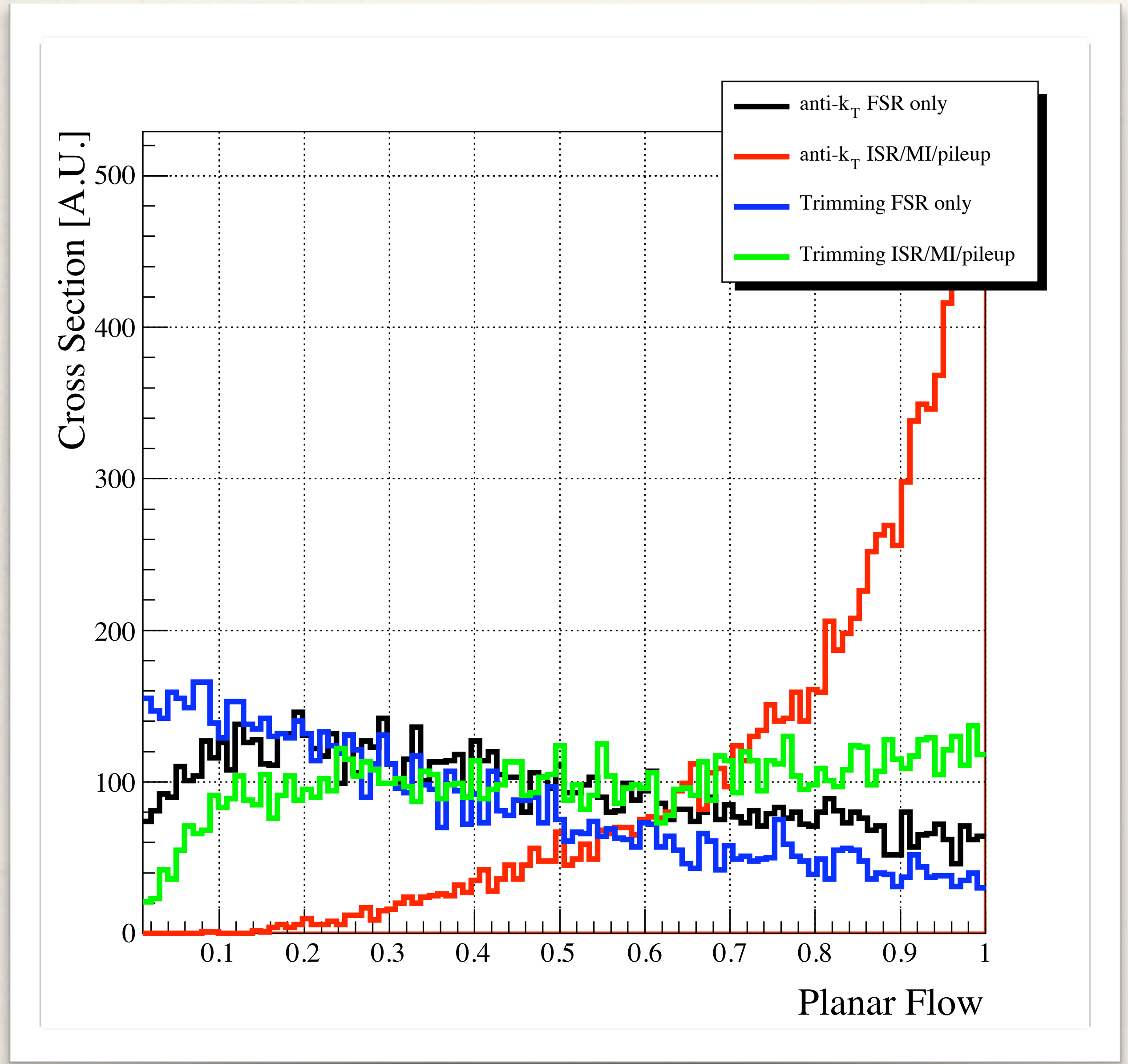
- Defined as normalized product of two jet moments:

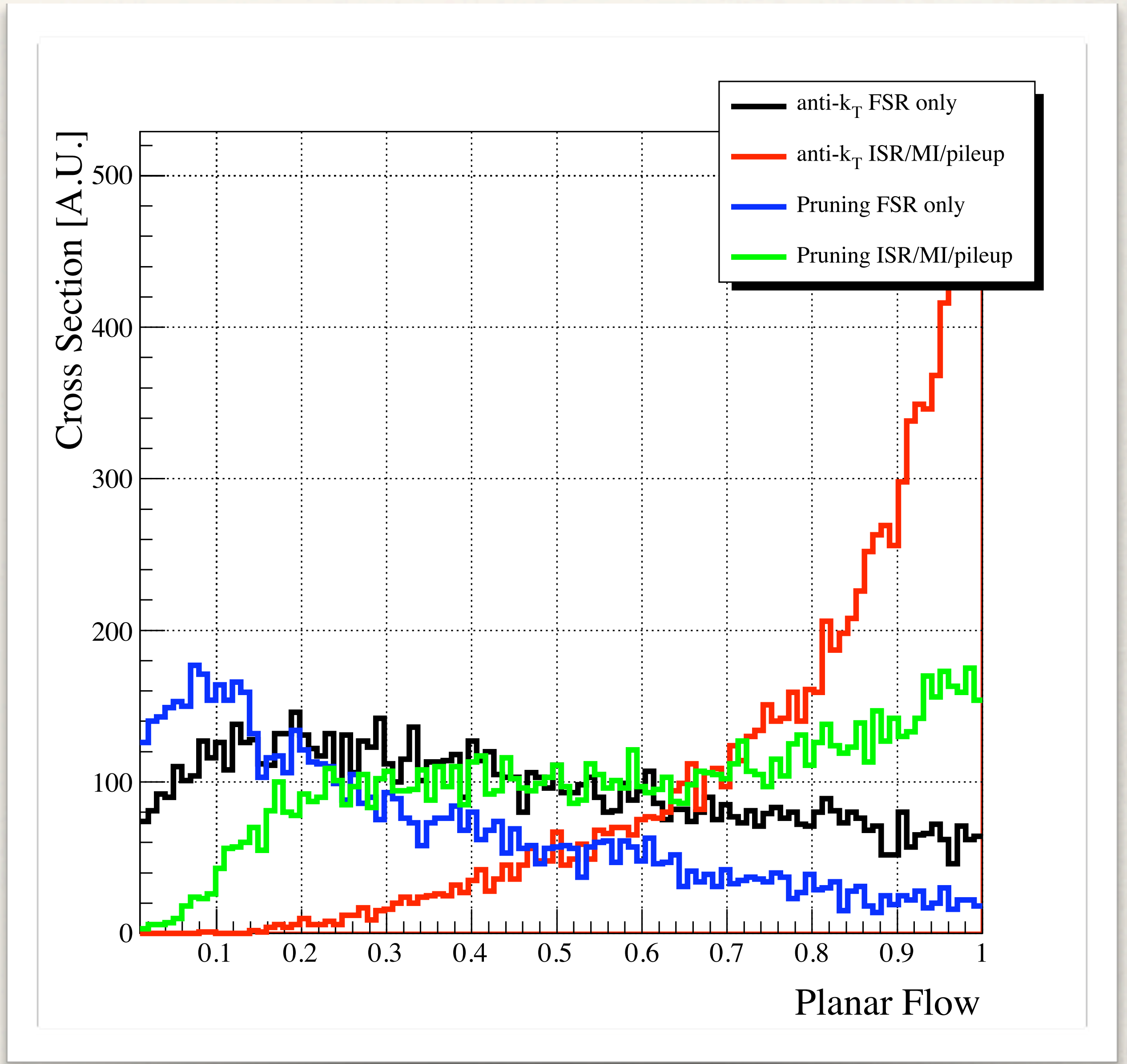
$$\mathrm{Pf} = \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} \qquad I_w^{kl} = \sum_i E_i \frac{p_{i,k}}{E_i}\frac{p_{i,l}}{E_i}$$

‣L. G. Almeida et al., *Substructure of high-pT Jets at the LHC*, Phys. Rev. D79 (2009) 074017, [0807.0234].
‣J. Thaler and L.-T. Wang, *Strategies to Identify Boosted Tops*, JHEP 07 (2008) 092, [0806.0023]

David Krohn

Planar flow: more sensitive to jet substructure details non-FSR skews PF to high values - most cut back

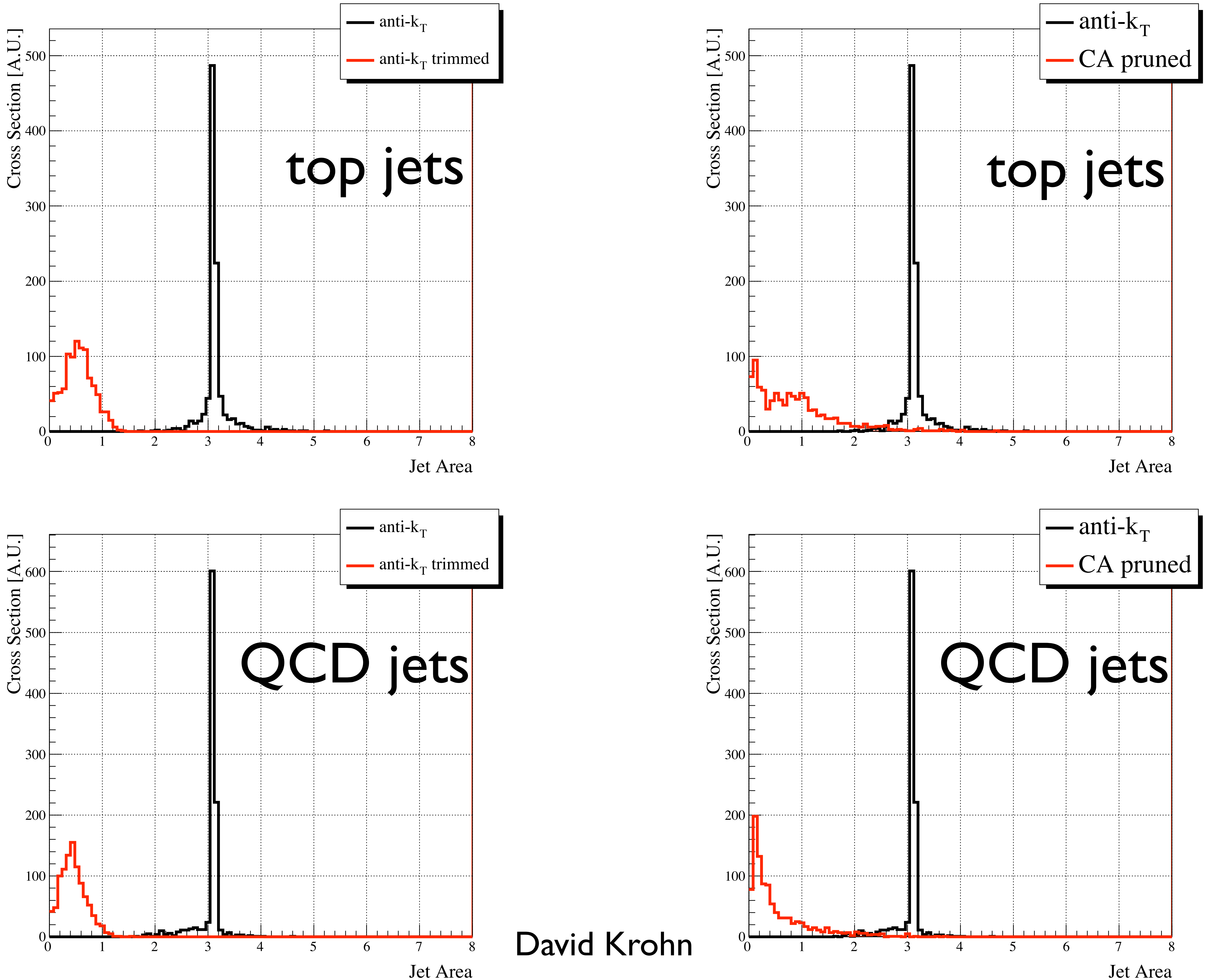




Not as good at restoring the distribution (compared to jet mass), but still progress

David Krohn

Jet areas: sensitivity to UE/pileup



David Krohn

Comparison conclusions

- Trimming corrects back to the FSR-only shape better
- Pruning over-corrects somewhat
  - Larger jet areas indicate it is not so simple - pruning is not removing "more" of the substructure than trimming, it is removing different parts
- Trimming is a local operation - objects clustered into subjets and trimmed or kept - similar to operations with topological clusters
- Pruning operates over the whole jet, and uses the algorithm to determine what to keep - seems better for reconstructing decays
- Can parameters of the algorithms be tuned to give similar behavior? Or merged into an algorithm good for both types of uses?
  - Current comparison work by groups at Princeton, UW, Oregon

Software Development

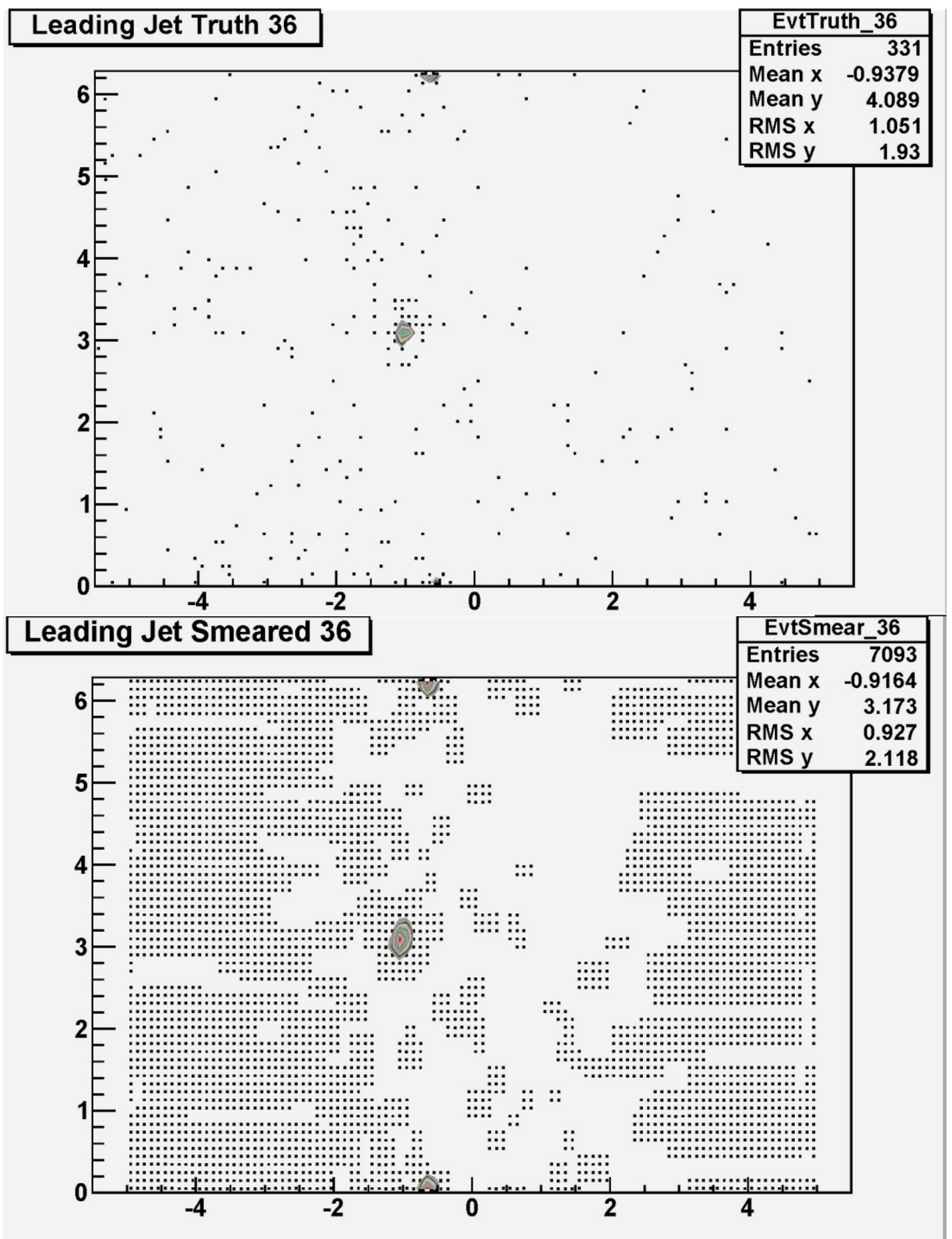
- Pruning/Trimming comparison a good context for development of better jet analysis software
  - FastJet plugins exist for both algorithms - easy to make jets - but how to compare?
  - SpartyJet framework discussed, used extensively at workshop
    - Works as a wrapper for FastJet, allows user to run jet-related analysis
- Jet substructure classes being developed in SpartyJet
  - Can work with different versions - original, pruned, trimmed - of a jet
  - Access to a wide variety of substructure observables without writing the framework yourself
  - At UW, pruning and trimming comparison being done with SpartyJet

Software Development

- Pruning/Trimming comparison a good context for development of better jet analysis software
  - FastJet plugins exist for both algorithms - easy to make jets - but how to compare?
  - SpartyJet framework discussed, used extensively at workshop
    - Works as a wrapper for FastJet, allows user to run jet-related analysis
- Jet substructure classes being developed in SpartyJet
  - Can work with different versions - original, pruned, trimmed - of a jet
  - Access to a wide variety of substructure observables without writing the framework yourself
  - At UW, pruning and trimming comparison being done with SpartyJet

Advanced theorist's detector - from Peter Loch

Only relies on a FastJet header file and STL! Simple, portable



Simple radial energy distribution in tower grids

- Ignore longitudinal development
  - Particle energy distributed transverse to direction of flight of particle in a plane through the particle impact point into the calorimeter
  - Shape of distribution from experiment/full simulation
- Integrated energy in profile is the same as particle energy
  - No calibration/acceptance/smearing of energy
- Distributed energies projected into regular eta/phi grid within modeled detector acceptance
  - Fakes calorimeter tower signal definition, including high eta losses
  - Different grid and eta acceptance for EM particles

Detector defaults

- EM acceptance -2.5 < eta < 2.5
  - Photons/electrons outside are mapped onto HAD towers
- HAD acceptance -5.0 < eta < 5.0
  - Particles outside are ignored completely
- Cylindrical calorimeter
  - R = 1200 mm, -2500 mm < z < 2500 mm
- High granularity
  - 0.025 x 0.025 (EM)
  - 0.1 x 0.1 (HAD)

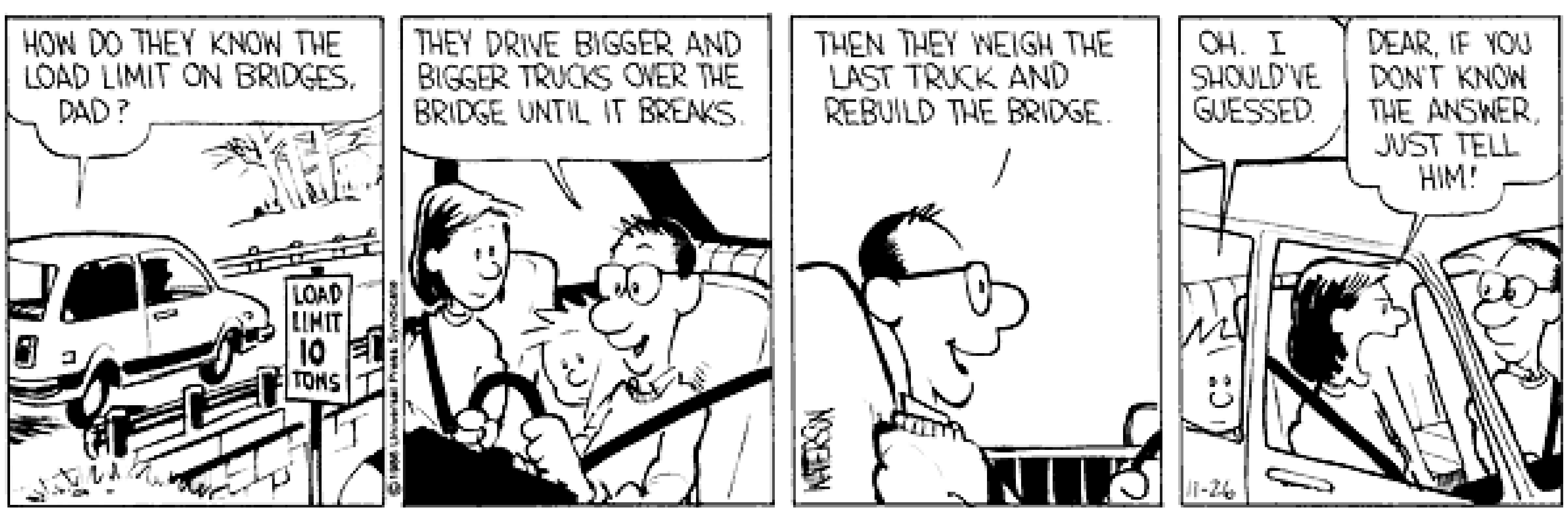
Shower shapes

- Presently Gaussian within cylinder
  - Lateral extend /cylinder radius 80 mm for EM particles (also in HAD grid!)
  - 160 mm for all others
  - Gaussian showers are too wide, this is just a simplification!
- Energy distributed in small "spots"

Peter Loch

Dangers of Jet Substructure from Monte Carlo

- Words of caution from Steve Mrenna:
    - Jet substructure will depend on physics that the MC doesn't get right (subleading effects)
    - Different MCs can give widely different results
    - Not well understood what physics details substructure techniques depend on
- Parton shower development can be fueled by theory calculations and more detailed substructure studies



Has the bridge already broken? It's hard to tell...

Seeing in color: Matt Schwartz

Jet Superstructure:

- Combines radiation pattern inside jets with global event structure to distinguish between processes
- Defines observables to tag jets as color connected to the beams or another central jet
- QCD radiation occurs inside color dipoles

Boosted color singlet decays are largely color-disconnected from the rest of the event

QCD background events are more color-connected

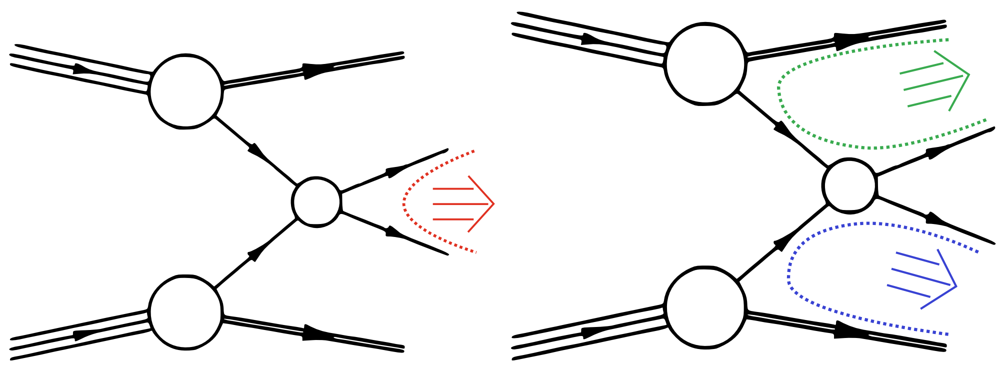
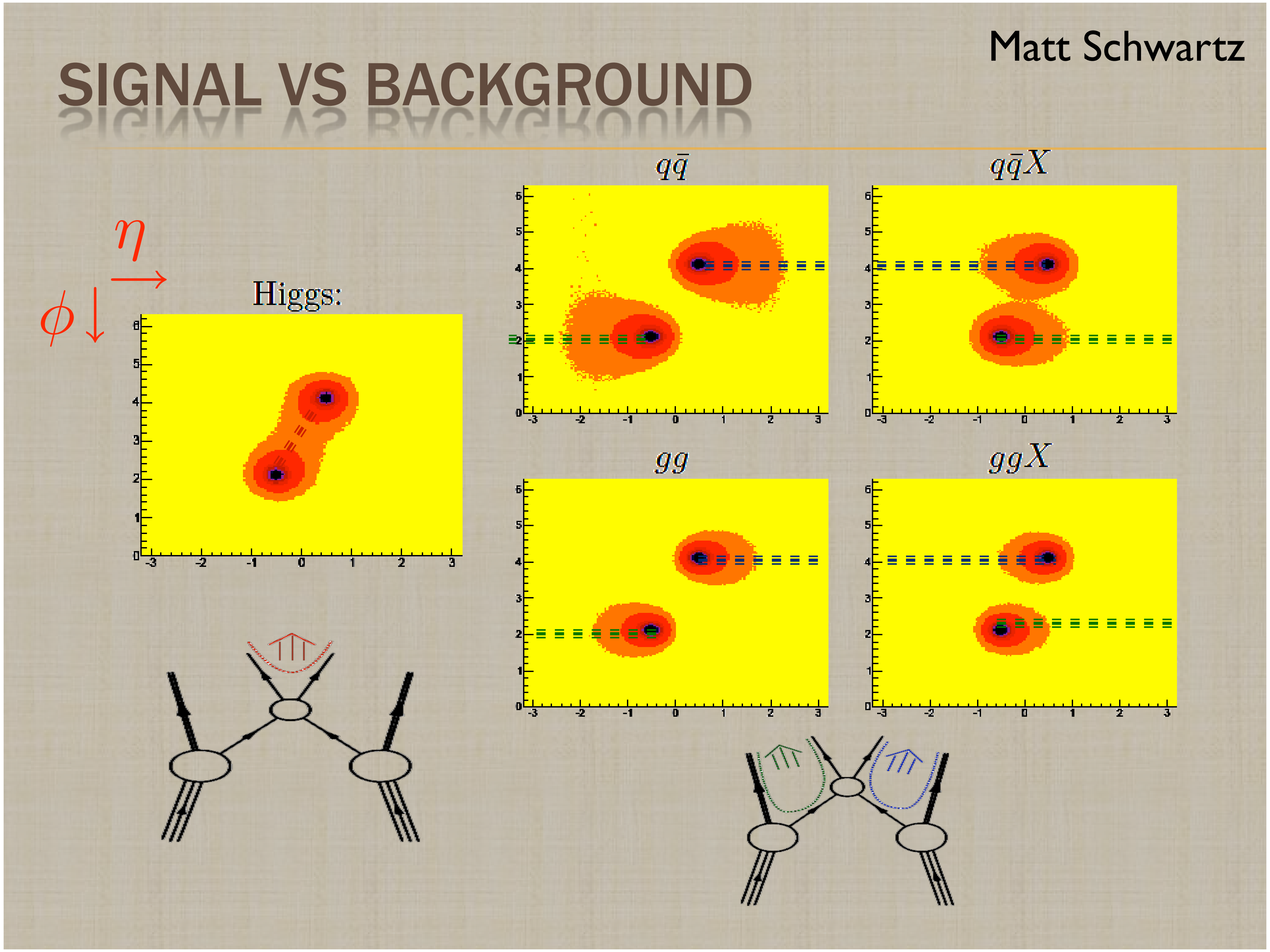


FIG. 1: Possible color connections for signal ($pp \to H \to b\bar{b}$) and for background ($pp \to b\bar{b}$).

Matthew Schwartz and Jason Gallicchio (arXiv:1001.5027)

Study color connections using Monte Carlo

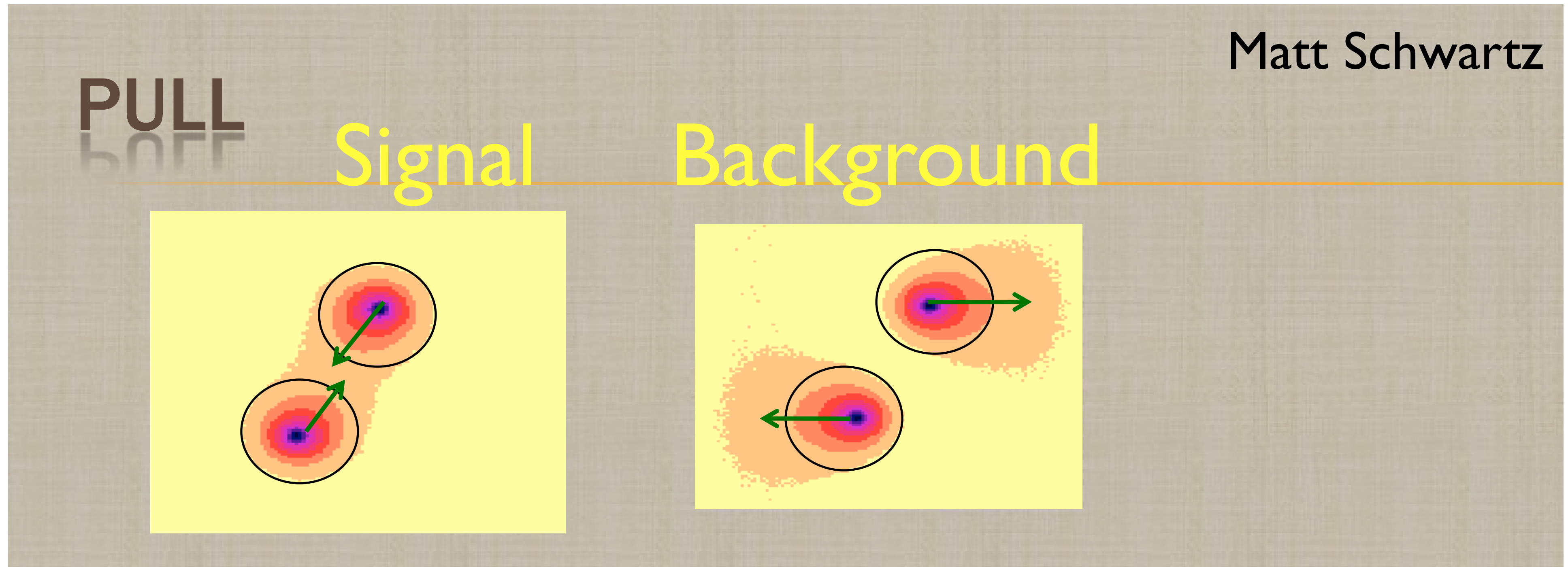
Parton shower single parton configurations millions of times, map radiation pattern



Signal (Higgs) events have most of their radiation between the central jets, background (dijet) events have radiation towards the beam (higher rapidity)

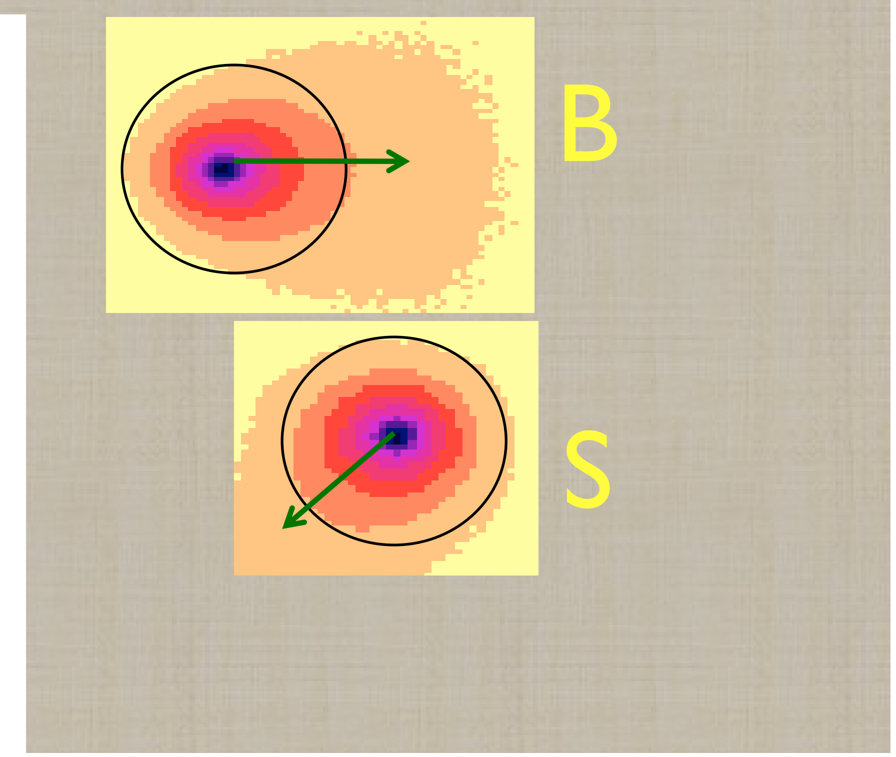
Quantify with Observables: Pull

Pull is a pT-weighted vector describing the direction of radiation in a jet wrt the jet centroid (~axis)



$$\vec{p} = \sum_{i \in J} \frac{p_T^i |r_i|}{p_T^J} \vec{r}_i$$

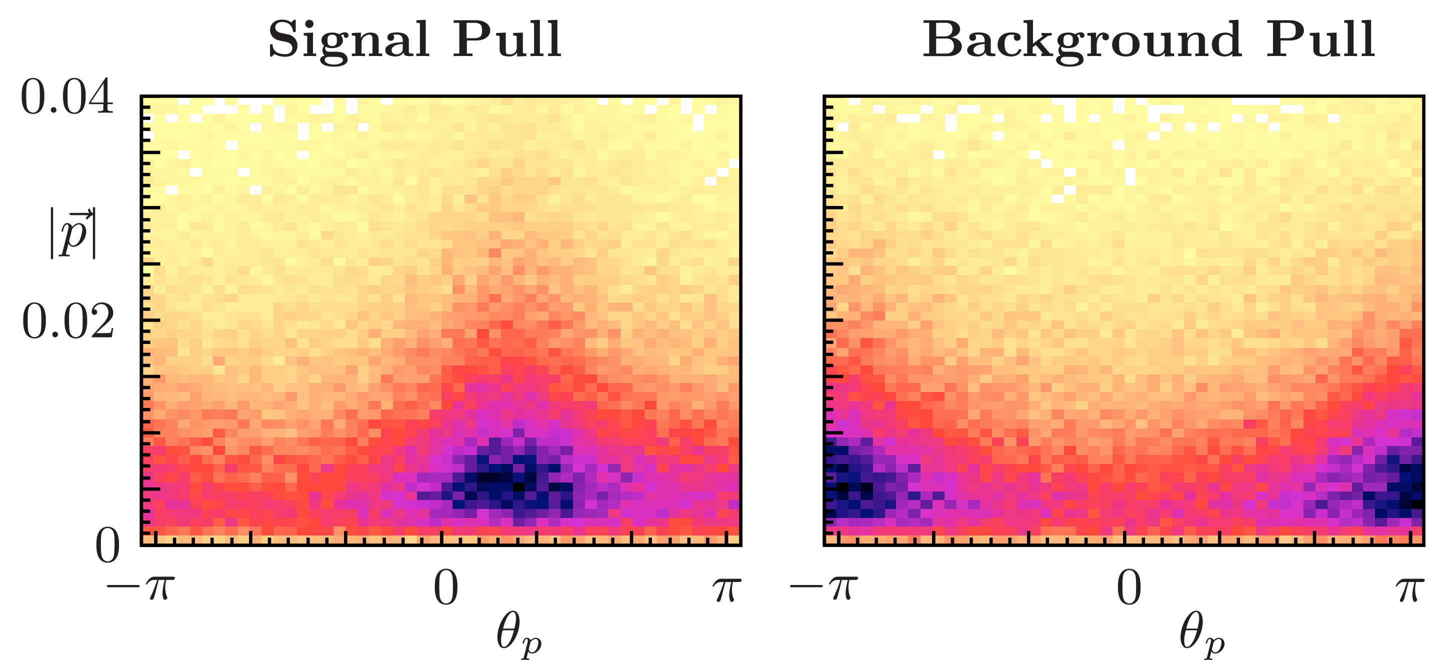
$$\vec{r}_i = (y_i, \phi_i) - (y_J, \phi_J)$$



Jets color-connected to the beam have pulls towards large rapidity
Centrally color-connected jets have "central" pulls

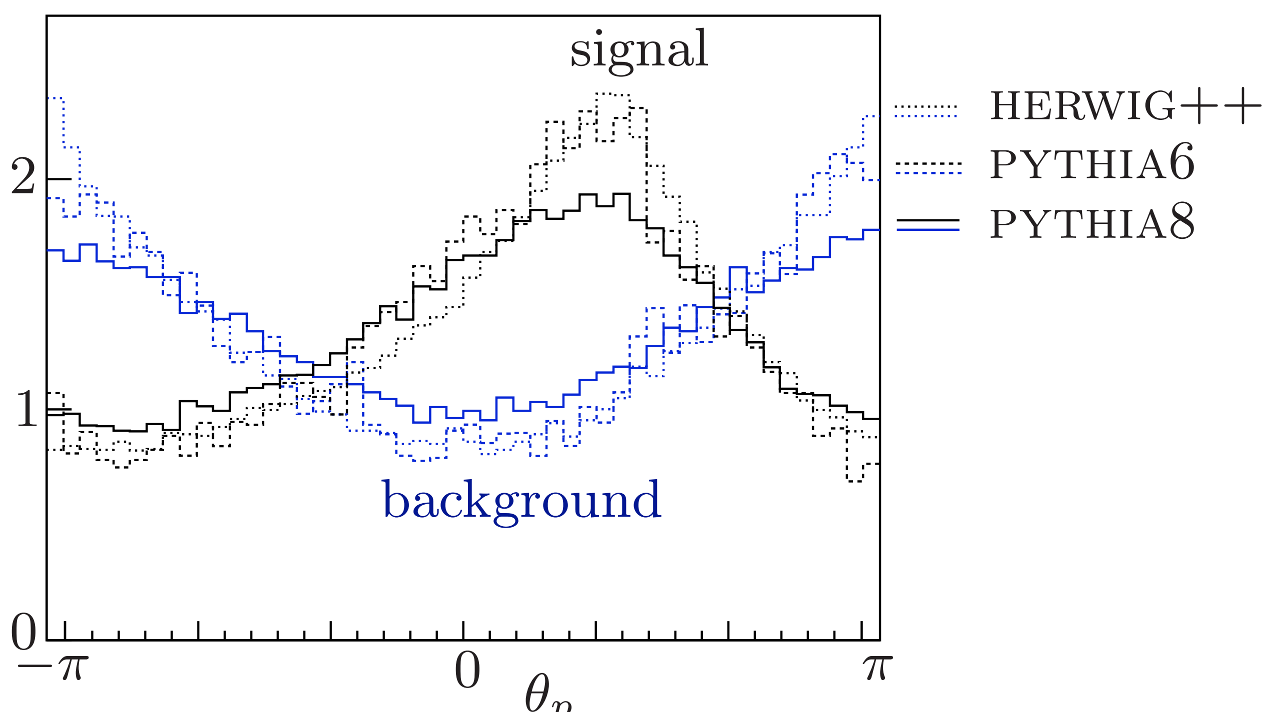
Pull Angle

The magnitude of pull is not very informative - but the angle is



Beam Central Beam

pull angle distribution for fixed parton configuration

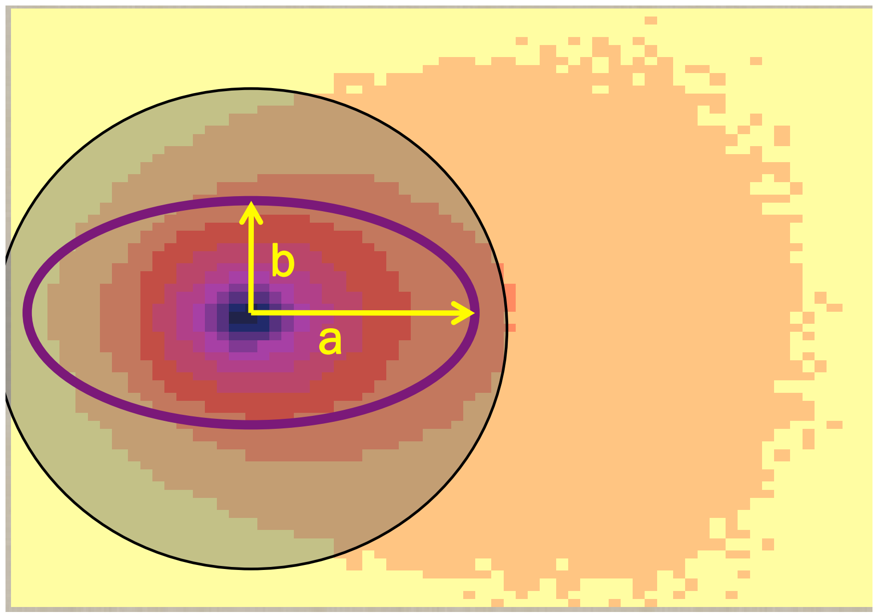


$$\vec{p} = \sum_{i \in J} \frac{p_T^i |r_i|}{p_T^J} \vec{r}_i$$

$$\vec{r}_i = (y_i, \phi_i) - (y_J, \phi_J)$$

Jet Superstructure paper

Higher Moment Observables



Matt Schwartz

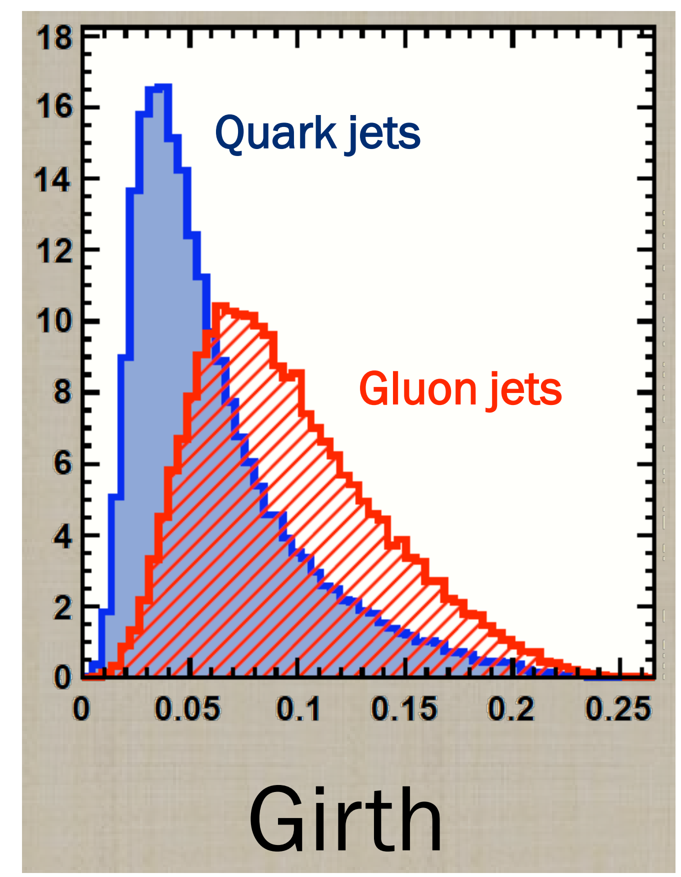
The radiation pattern will depend on the color charge of the jet and the event topology
- observables can sort this out

$$I = \sum_i \frac{p_T^i}{p_T^J} |r_i| \begin{pmatrix} \Delta\phi_i^2 & -\Delta\phi_i \Delta y_i \\ -\Delta\phi_i \Delta y_i & \Delta\eta_i^2 \end{pmatrix}$$

$\rightarrow$ yields eigenvalues a, b

girth: $\sqrt{a^2 + b^2}$ - size of jet

eccentricity: $\sqrt{1 - \frac{b^2}{a^2}}$ - no clear use



Use girth for quark vs. gluon jet separation

Advantages of Jet Superstructure

- Largely orthogonal to other observables
  - Uses inter-jet measures for discrimination
  - Could be combined with fat-jet substructure (e.g., the Higgs filter) to identify color singlet decays
- Very straightforward application in clean events
  - Easy to calibrate from Z production
  - Would learn something about shower reconstruction in a detector in applying to real events
- The "pull" observable may be calculable for simple event topologies
  - Calculable in SCET? - the technology is developing rapidly

Homework (for Boost 2010)

- Continue comparisons of jet substructure techniques
  - Public samples exist, public code exists for several substructure techniques
  - David Miller exploring using trimming/pruning for pileup reduction, other tests (e.g., Z reconstruction) started at workshop
  - Groups are currently comparing pruning and trimming - David Krohn, Chris Vermilion, Michael Spannowsky and others
- Continue theory-experiment interaction on applications of jet substructure
  - Uncertainty on how to calibrate jet substructure
  - New applications can benefit from discussion
    - e.g., pileup reduction using trimming/pruning

Homework (for Boost 2010)

- Continue software development to integrate jet substructure and analysis tools
  - SpartyJet being developed into a more full-fledged analysis suite
  - FastJet plugins part of a universal software set
  - Peter Loch's basic detector code useful for theorists to study basic detector effects
- Theory working group
  - Compare formulations for resummation in QCD and SCET of event shapes
    - Current Stony Brook/Berkeley/Washington collaboration
  - Find observables to better understand jet substructure, involves understanding role of jet algorithms

Summary

- Workshop was successful in bringing together theorists and experimentalists to discuss important aspects of jet substructure
  - Successful implementation of jet substructure requires collaboration between both groups
  - Issues raised for both experimentalists (calibration, performance after detector simulation) and theorists (how do tools compare, ensure techniques robust)
- Positive discussions, progress
- Wiki forum for continued collaboration:

  http://librarian.phys.washington.edu/lhc-jets/

- Boost 2010 should bring more developments!

  http://www.physics.ox.ac.uk/boost2010/index.asp