Calor2000

October 12, 2000

Review of Jet Clustering at Tevtron

Anwar Ahmad Bhatti

The Rockefeller University

Run I Experience:

- Snowmass Accord
- Cone Clustering algorithm
- k_T clustering
- Energy calibration/Jet Resolution
- Underlying Event Subtraction

Run II Workshop Proposal:

- Problems with cone algorithms
- Requirements for good clustering algorithms
- Algorithms
  - Seedless Cone Algorithm
  - Mid-Point Cone
  - k_T Clustering
- Some preliminary comparison

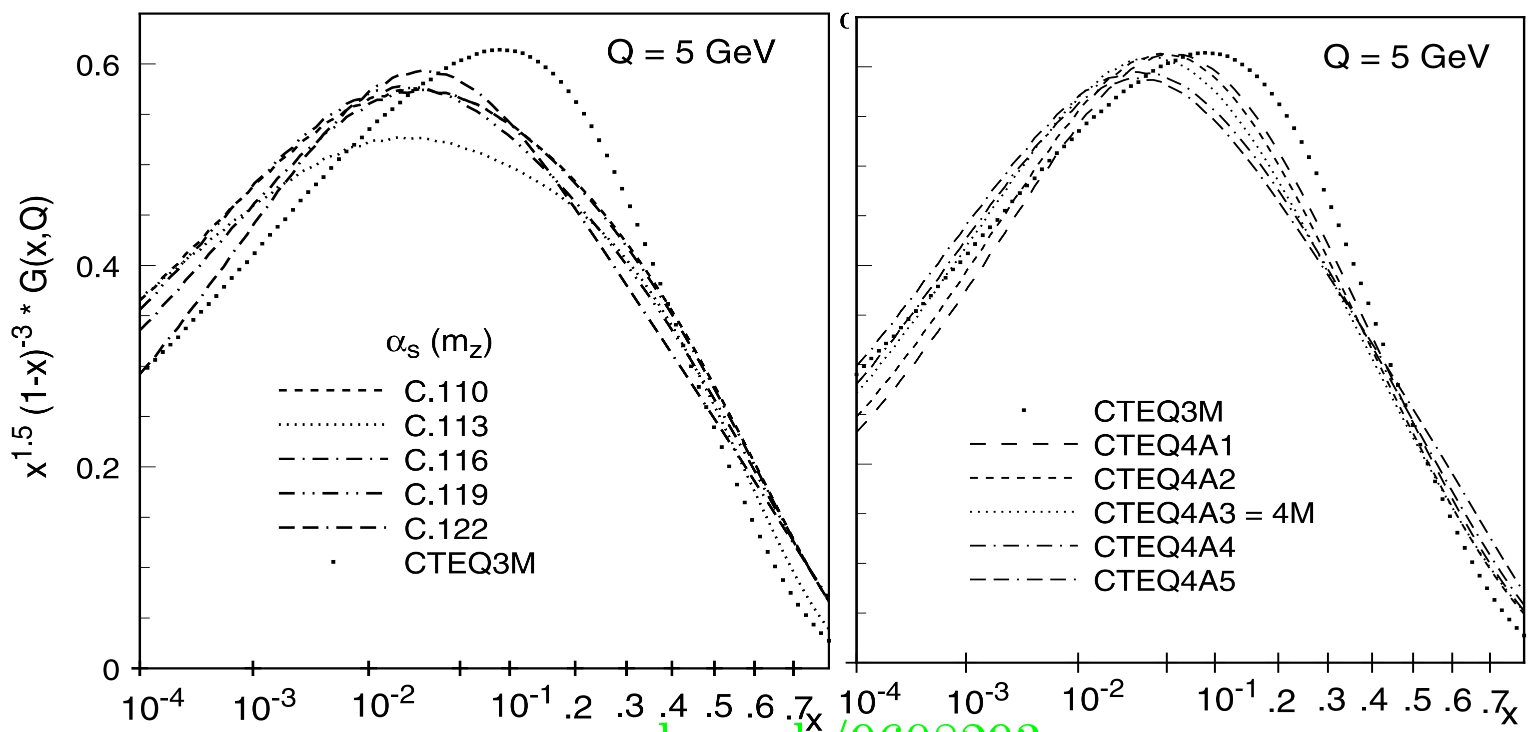
Why do we need better clustering?

QCD is well established.

- Jet Physics
  - Measure parton distribution functions
  - Test various calculation techniques (NNLO calculation, Resummation)
  - Estimate non-perturbative effects/background to new physics
  - Angular distributions
  - Multi-Jet States
  - Jet Fragmentation Studies
- Mass Reconstruction
  - Hadronic W/Z reconstruction
  - Top Mass measurement
  - Higgs Search

Question is how accurately can we do it?

Parton Distributions from $p\bar{p}$ Jet Measurements.

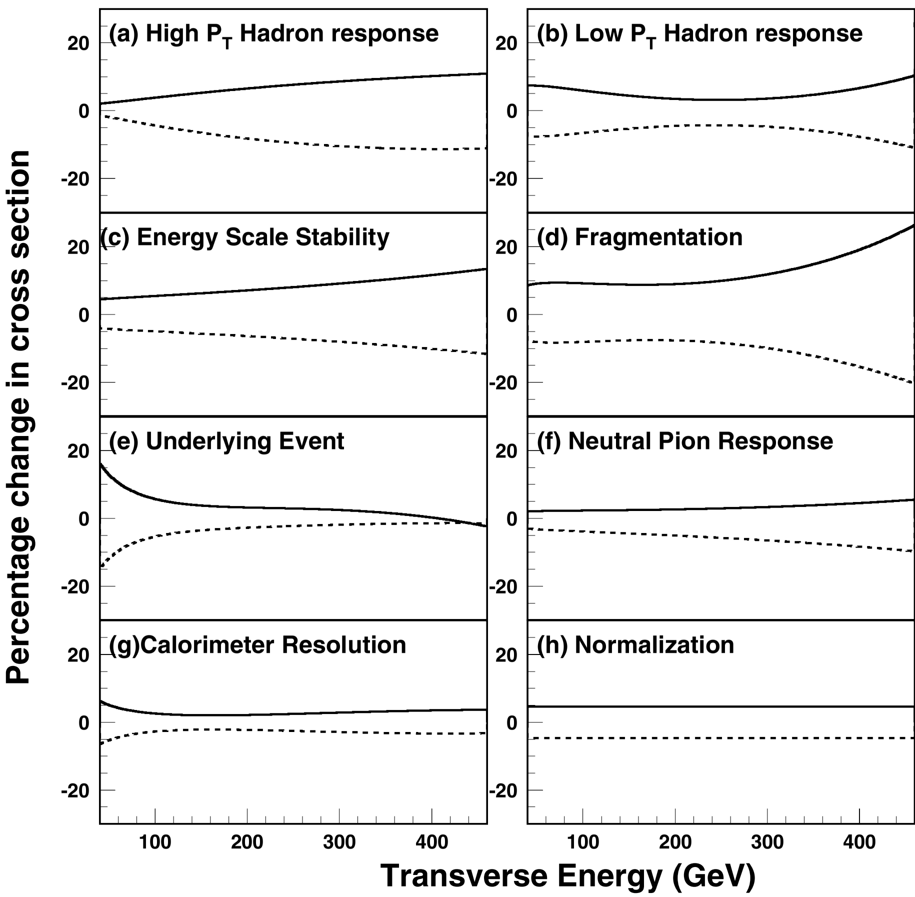


hep-ph/9608293

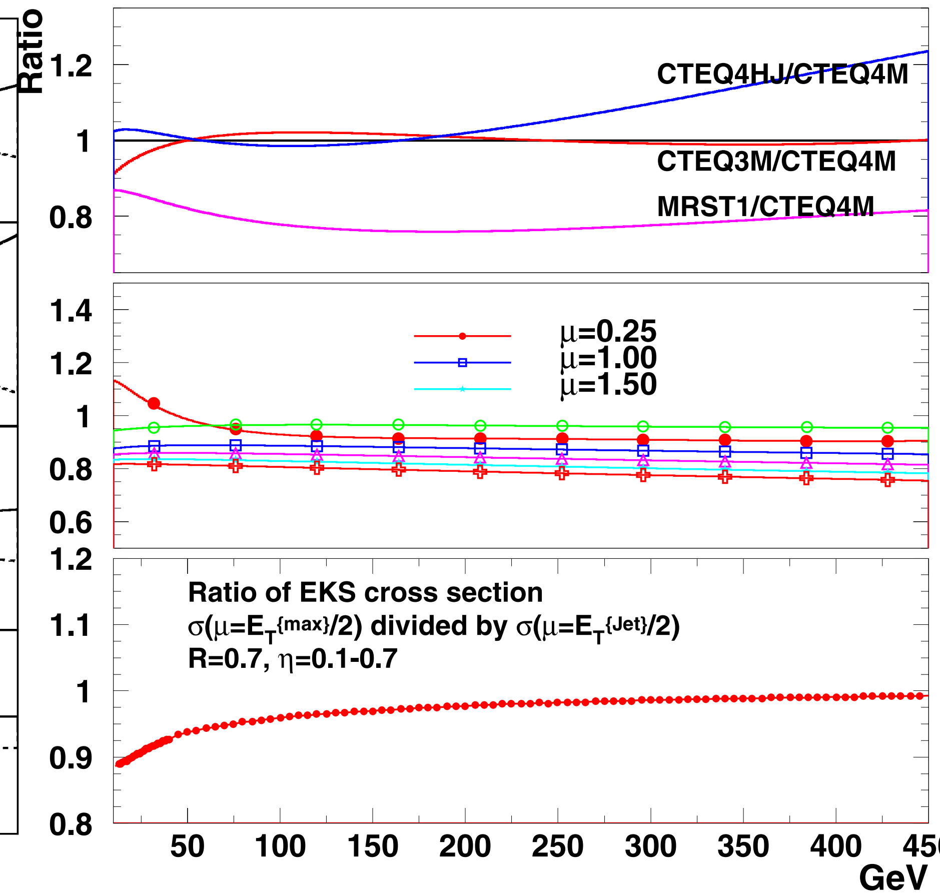
- The inclusive jet data improves the stability of gluon distributions.
- High statistics jet data from run II will be used to measure PDF at medium and large x.
- We must use the "same" jet definition in both theory and data.

Uncertainties in Inclusive Jet Production

Experimental



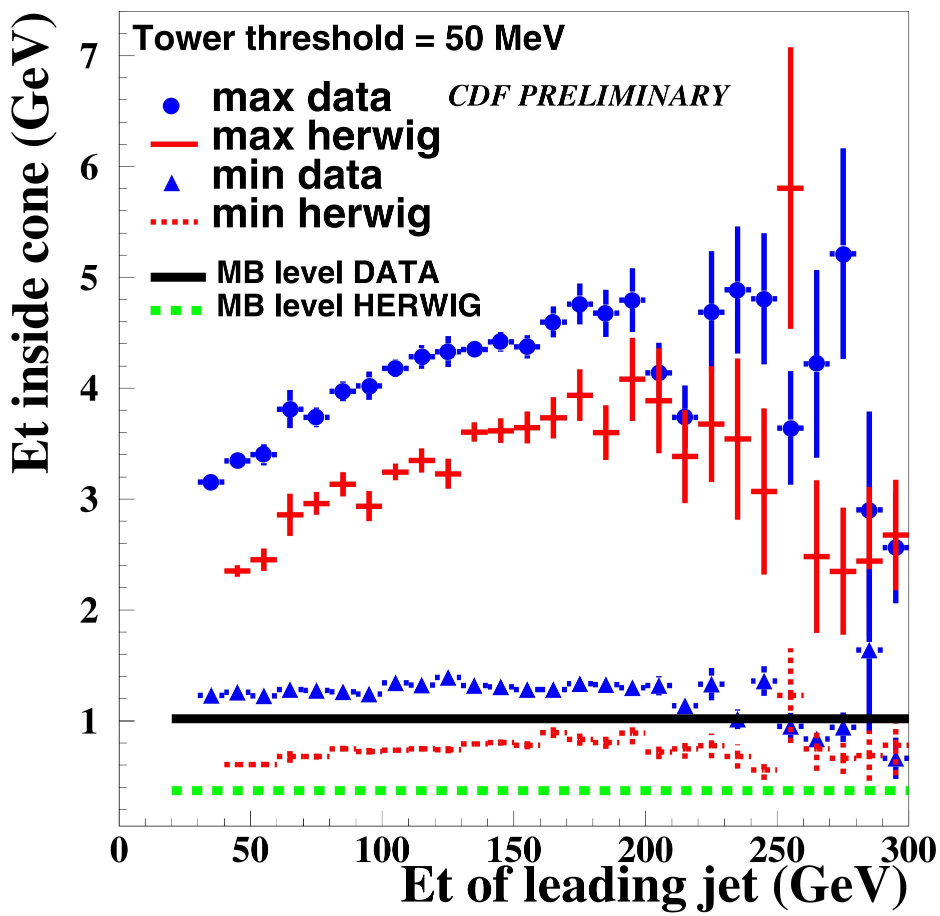
Theoretical



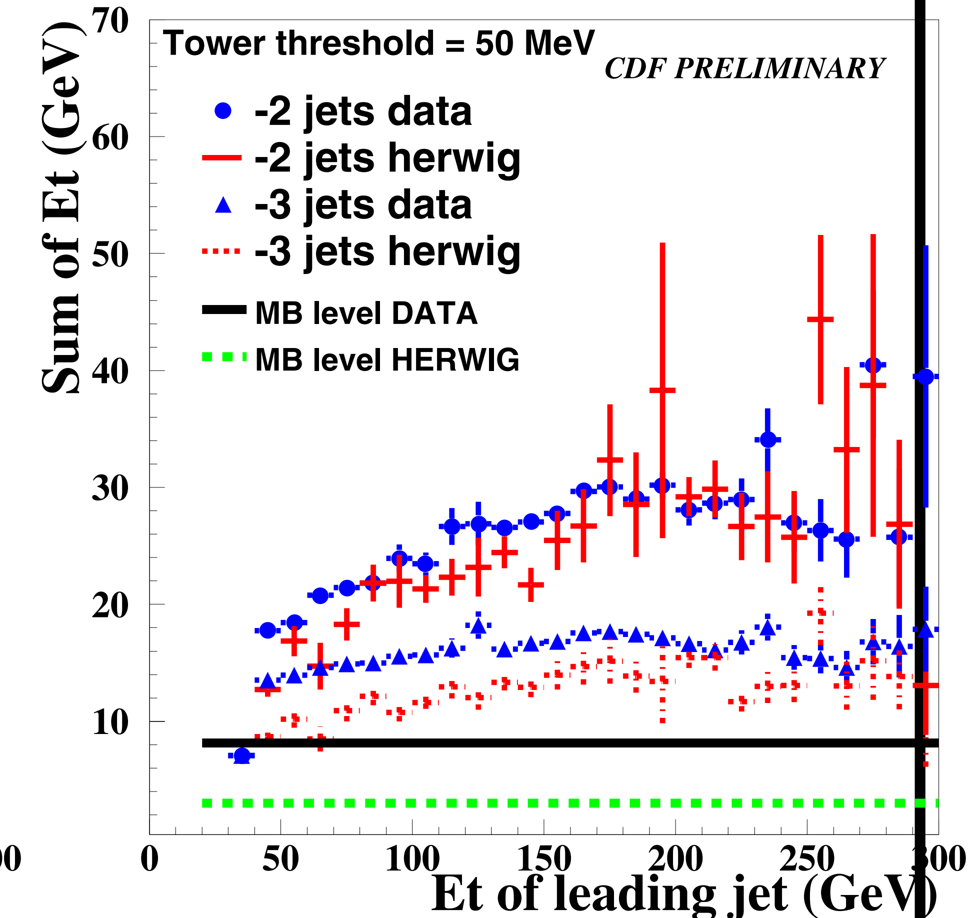
We hope to reduce these uncertainties in near future.

Ambient Energy in Jet Events

Max and min cones



Swiss Cheese

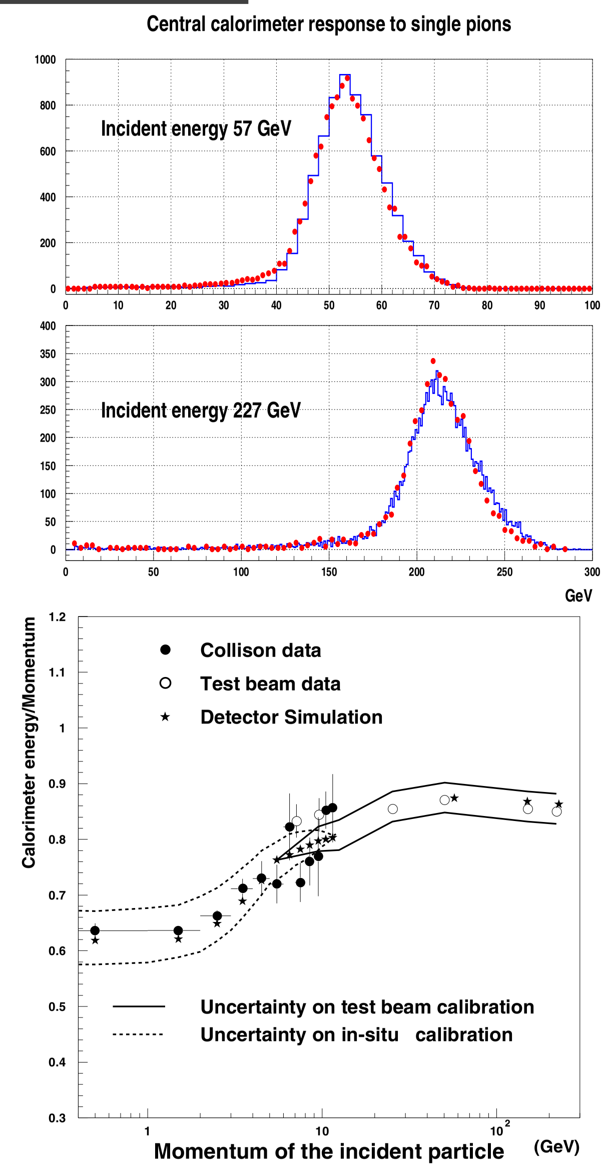


Herwig shows the same trends but magnitude might be lower.

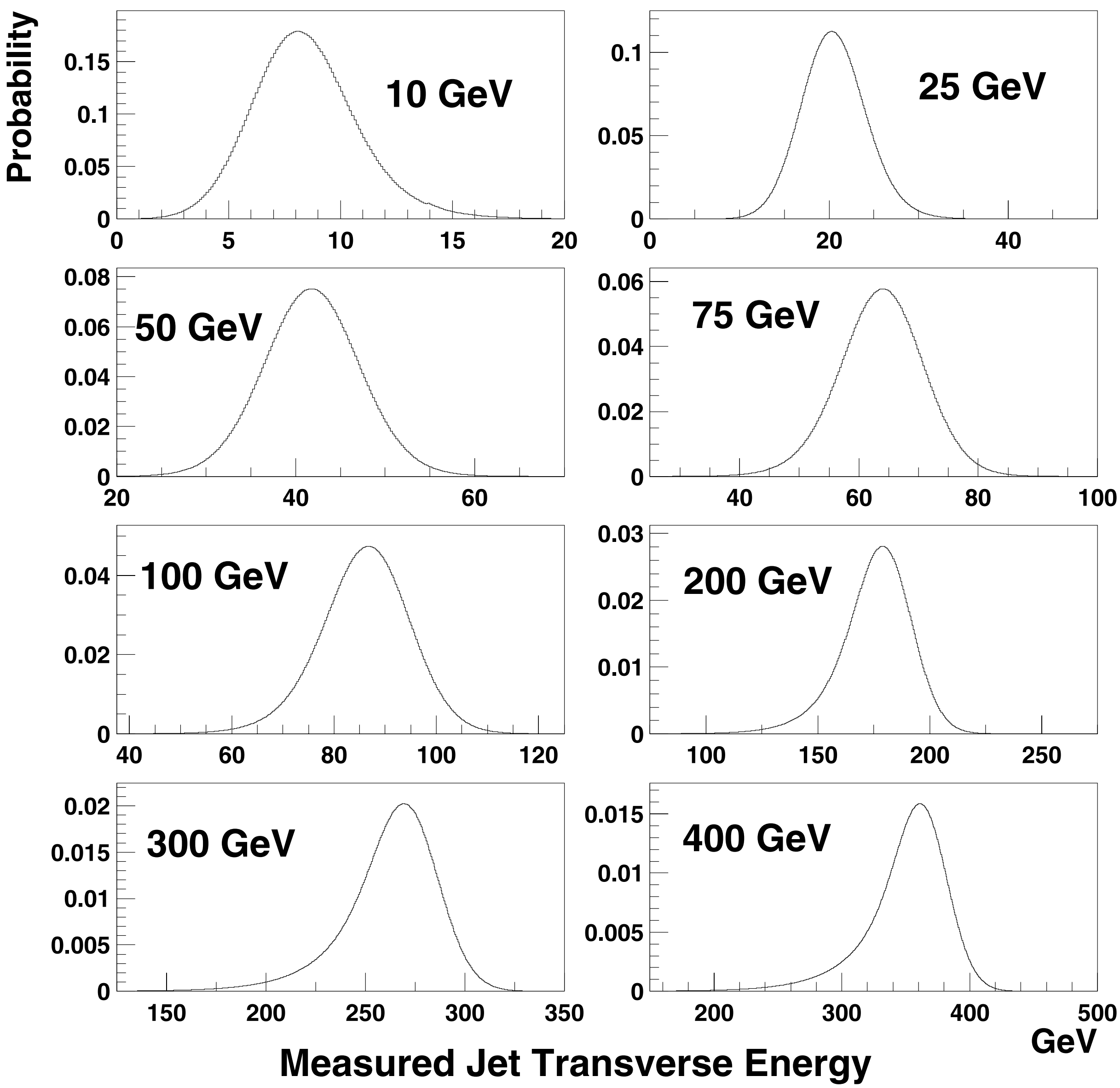
Better understanding will reduce the uncertainty.

CDF Energy Calibration

- Use test-beam data for single particle response for high P_T
- $p\bar{p}$ data for low momentum particles
- Charged Particle multiplicity and momentum distribution from CDF data
- Assume that EM calorimeter is linear.
- Convolute particles in a jet with single particle response to determine jet response.



CDF Jet Energy Response



Jet resolution is proportinal to Jet E_T

.

Snowmass Algorithm

- Associate a 4-vector with Each particle/tower
- Find a cone such that jet centroid is aligned with geometric center.
- Each particle statisfies

$$i \subset C \quad : \quad \sqrt{(\eta^i - \eta^C)^2 + (\phi^i - \phi^C)^2} \leq R.$$

In the Snowmass algorithm a "stable" cone (and potential jet) satisfies the constraints

$$\eta^C = \frac{\sum_{i \subset C} E_T^i \eta^i}{E_T^C}, \quad \phi^C = \frac{\sum_{i \subset C} E_T^i \phi^i}{E_T^C}$$

($i.e.$, the geometric center is identical to the E_T-weighted centroid with $E_T^C = \sum_{i \subset C} E_T^i$).

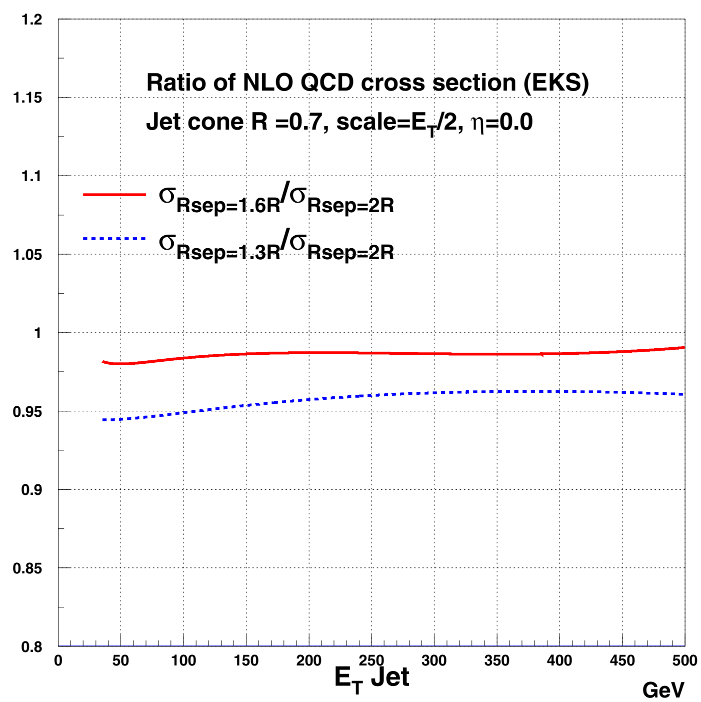
- $E_T^J = \sum E_T^i$
- $\eta^J = 1/E_T^J \sum E_T^i \eta^i$
- $\phi^J = 1/E_T^J \sum E_T^i \phi^i$

CDF Jet Clustering

- PreClustering
  - Join towers in forward/plug region to make 24 segments in ϕ
  - Make E_T ordered tower list with $E_T > 1$ GeV
  - First tower $\equiv$ first PreCluster
  - Add a tower to a precluster if within 7x7 towers AND adjecent to existing tower in the precluster.
  - If not, start a new Precluster
  - Stop when all towers are assigned.
  - Restore full segmentation of Forward/Plug Calorimeter.
- All tower, $E_T > 1$ GeV are uniquely assigned to a precluster.
- Clustering
  - Make tower list with E_T>100 MeV.
  - Order PreCulster in E_T
  - Add towers to a precluster if within $\Delta R < 0.7$
  - Iterate such that tower list in cluster is stable.
- Merging and Splitting
  - E_T order clusters
  - do i=1,number of clusters
  - do j=1,i-1
  - Find the common towers.
  - If commom E_T>75%, merge.
  - Else assign common towers to closer cluster in iterative fashion.
- Calculate jet parameters from tower list.

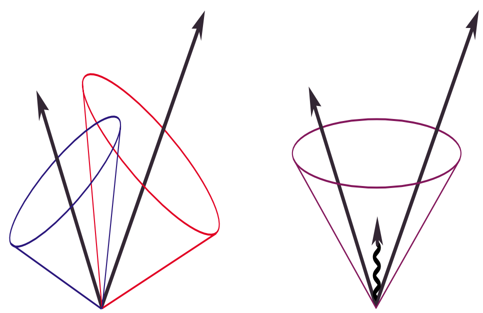
Ad-hoc Clustering Parameter in theory (R_{sep})

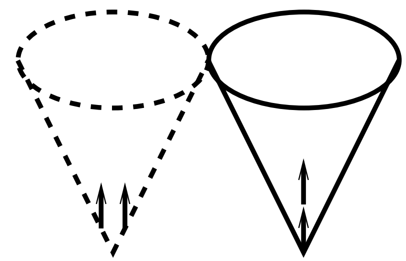
- Experimental clustering has overlapping cones and thus splitting merging.
- Introduce to mimic splitting/merging in data.

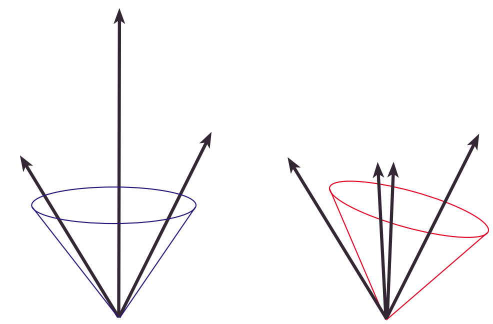


Successful in describing energy distribution within a jet at the cost of another parameter in theory which can be tuned.

Problems with current cone algorithms







Infra-red singularity

Two jets are merged into one of there is a soft radiation above seed threshold in between them.

Collinear Sensitivity

Jet is not reconstructed if the seed splits into two towers such that it falls below the threshold.

Collinear Sensitivity

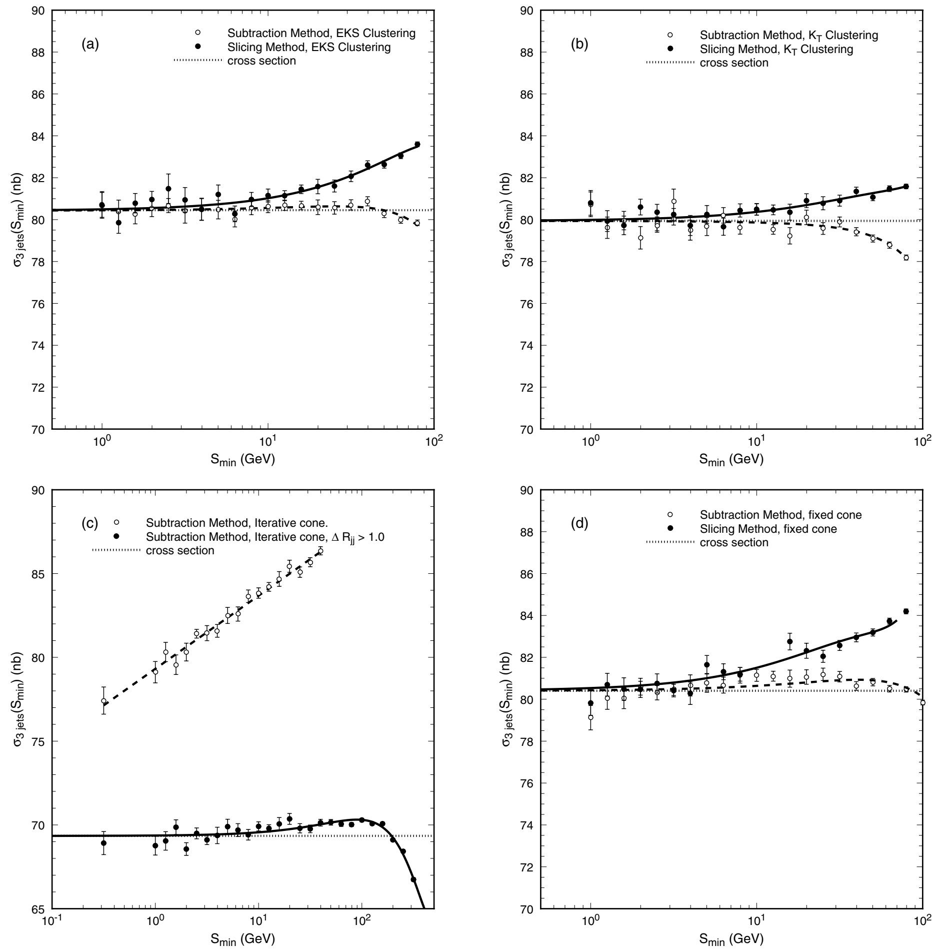
One (two) jets will be constructed if the middle parton splits into two in E_T ordered clustering.

Get rid of seeds/Add more seeds.

Seedless Algorithm — MidPoint Algorithm

Add seed even if there is no energy deposited.

NLO Three Jet Production hep-ph/p610433



- Jet cross section depends on the jet definition.
- Upto four partons in final state.
- Need a resolution parameter to define a parton.
  - Mass of two parton $< s_{min}$, Unresolvabe, treat as single parton
- Cross Section should not depend on s_{min}.

Desirable Features of clustering algorithm

- Infrared Safe
- Collinear Safe
- Invariance under boosts
- Boundary Stability
- Order Independence
- Straightforward Implementation
- Detector Independence
- Minimization of resolution smearing and angle biases
- Stability with luminosity
- Efficient use of computing resources
- Maximum Reconstruction Efficiency
- Ease of calibration
- Ease of use
- Fully specified

Clustering Algorithms for Run II

Recommendations

- Seedless Cone Clustering
- Midpoint Cone Clustering
- k_T algorithm

Implementation

CDF

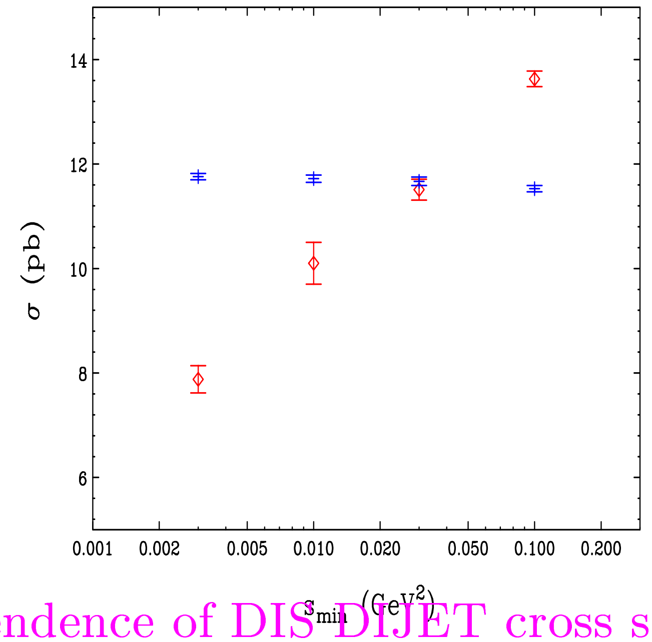
- JetClu (Run I cone algorithm)
- k_T Clustering
- Seedless Cone Clustering
- Midpoint Cone Clustering

DØ

- k_T algorithm
- Seedless Cone Clustering
- Midpoint Cone Clustering

MidPoint Cone Clustering

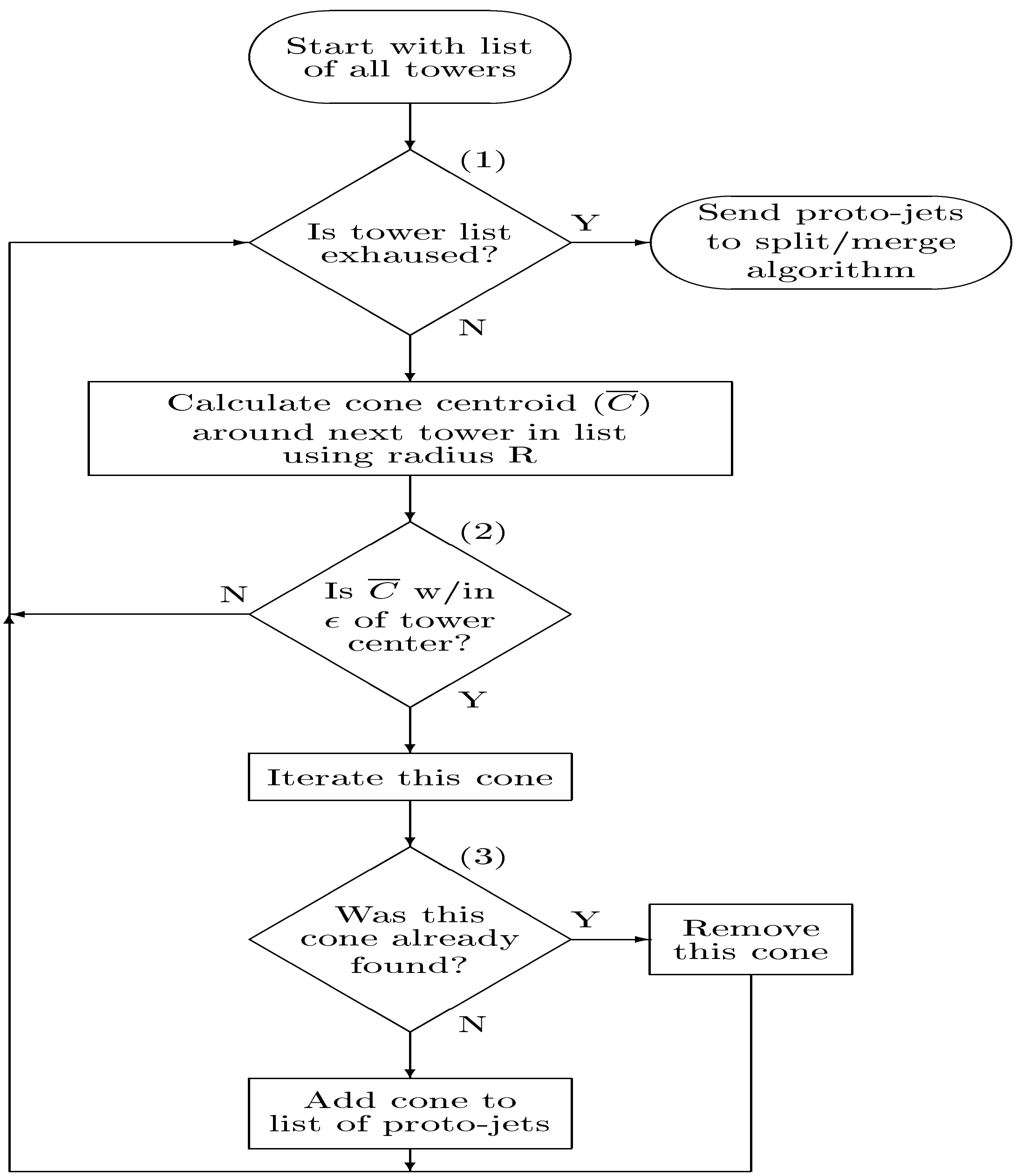
To remove the sensitivity to soft radiation, add additional seeds at positions given by $p_i + p_j$, $p_i + p_j + p_k$ etc.



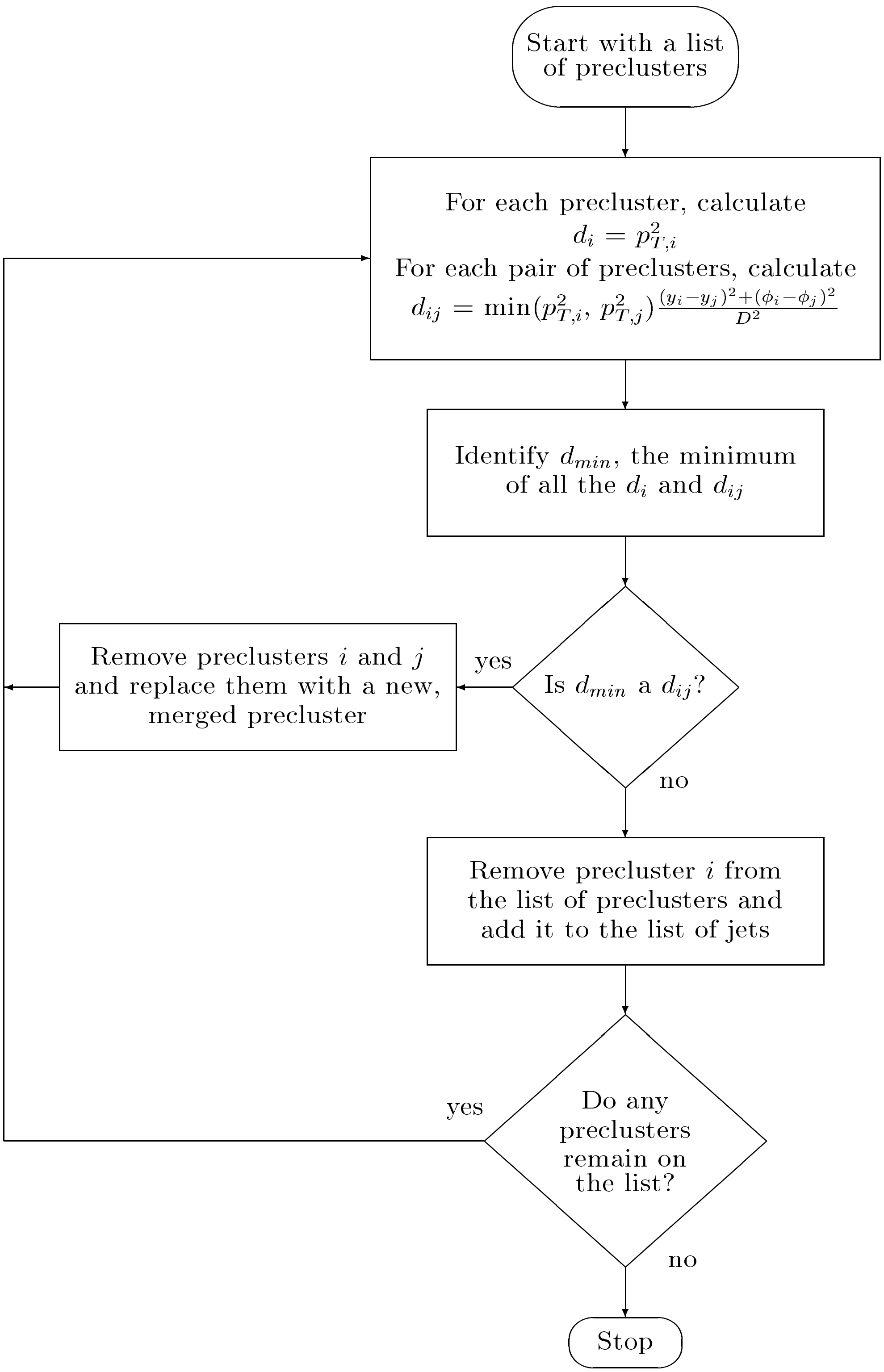
s_{min} dependence of DIS DIJET cross section at NLO.

- NLO $p\bar{p}$ 3-jet production is too slow.
- The cross section diverges as $\log(s_{min}/Q^2)$ for "no center seed" clustering.
- In data
  - Generate E_T order tower
  - Find protojets around towers with $E_T >$ threshold
  - Generate midpoints from list of protojets
  - Find protojets around midpoints
  - Goto split/merge stage

Seedless Cone Clustering Algorithm



k_T Algorithm



Jet Parameter Definitions

The cone algorithm starts with a cone defined in E-scheme variables as

$$i \subset C \quad : \quad \sqrt{\left(y^i - y^C\right)^2 + \left(\phi^i - \phi^C\right)^2} \leq R. \tag{1}$$

where for massless towers, particles, or partons $y^i = \eta^i$. The E-scheme centroid corresponding to this cone is given by

$$p^C = (E^C, \mathbf{p}^C) = \sum_{i \subset C} (E^i, p_x^i, p_y^i, p_z^i) \;, \tag{2}$$

$$\bar{y}^C = \frac{1}{2} \ln \frac{E^C + p_z^C}{E^C - p_z^C} \;, \quad \bar{\phi}^C = \tan^{-1} \frac{p_y^C}{p_x^C} \;. \tag{3}$$

A jet arises from a "stable" cone, for which $\bar{y}^C = y^C = y^J$ and $\bar{\phi}^C = \phi^C = \phi^J$, and the jet has kinematic properties

$$p^J = (E^J, \mathbf{p}^J) = \sum_{i \subset J = C} (E^i, p_x^i, p_y^i, p_z^i) \;, \tag{4}$$

$$p_T^J = \sqrt{(p_x^J)^2 + (p_y^J)^2} \;, \tag{5}$$

$$y^J = \frac{1}{2} \ln \frac{E^J + p_z^J}{E^J - p_z^J} \;, \quad \phi^J = \tan^{-1} \frac{p_y^J}{p_x^J} \;. \tag{6}$$

PreClustering

- Order independence (same results for partons, particles, towers)
- Detector independence
- CPU

CDF Preclustering

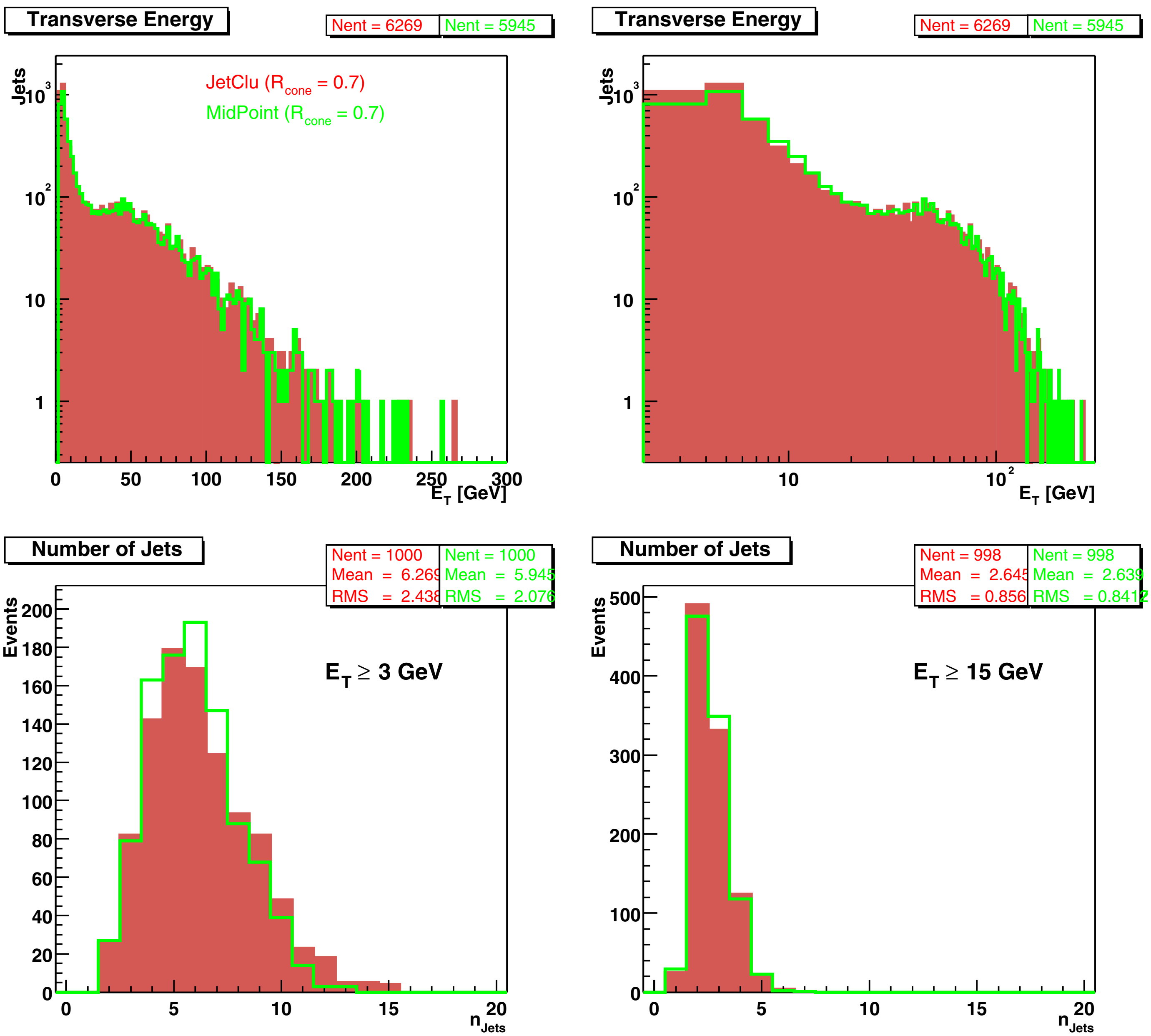
- CDF Run II calorimeter has 1536 towers.
- Each tower with $E_T > 100$ MeV is precluster.

DØPreclustering

- $\sim$ 45000 cells, $\sim$ 6000 towers.
  1. Identify each cell with a 4-vector
  2. Removes celles with $p_T < -500$ MeV.
  3. Add cell within a tower using 4-vector addition
  4. PreCluster towers with $\Delta R < 0.2$
  5. Redistribute negative energy preclusters to neighbours
  6. Redistribute preclusters with p_T <200 MeV to neighbours

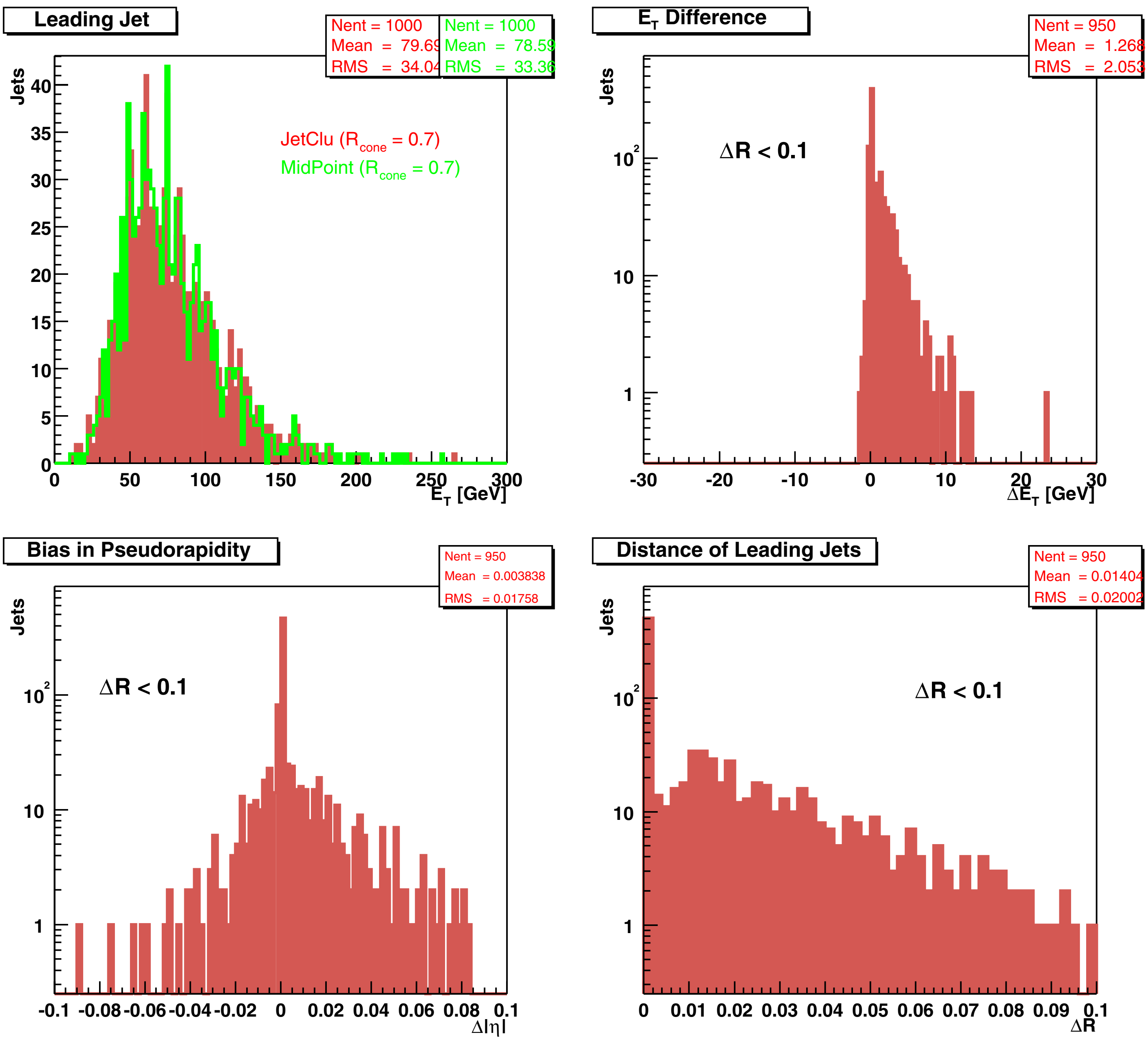
Comparison of MidPoint with JetClu

Jet50: JetClu vs. MidPoint (R_{cone}=0.7): E_T and n_{jets}



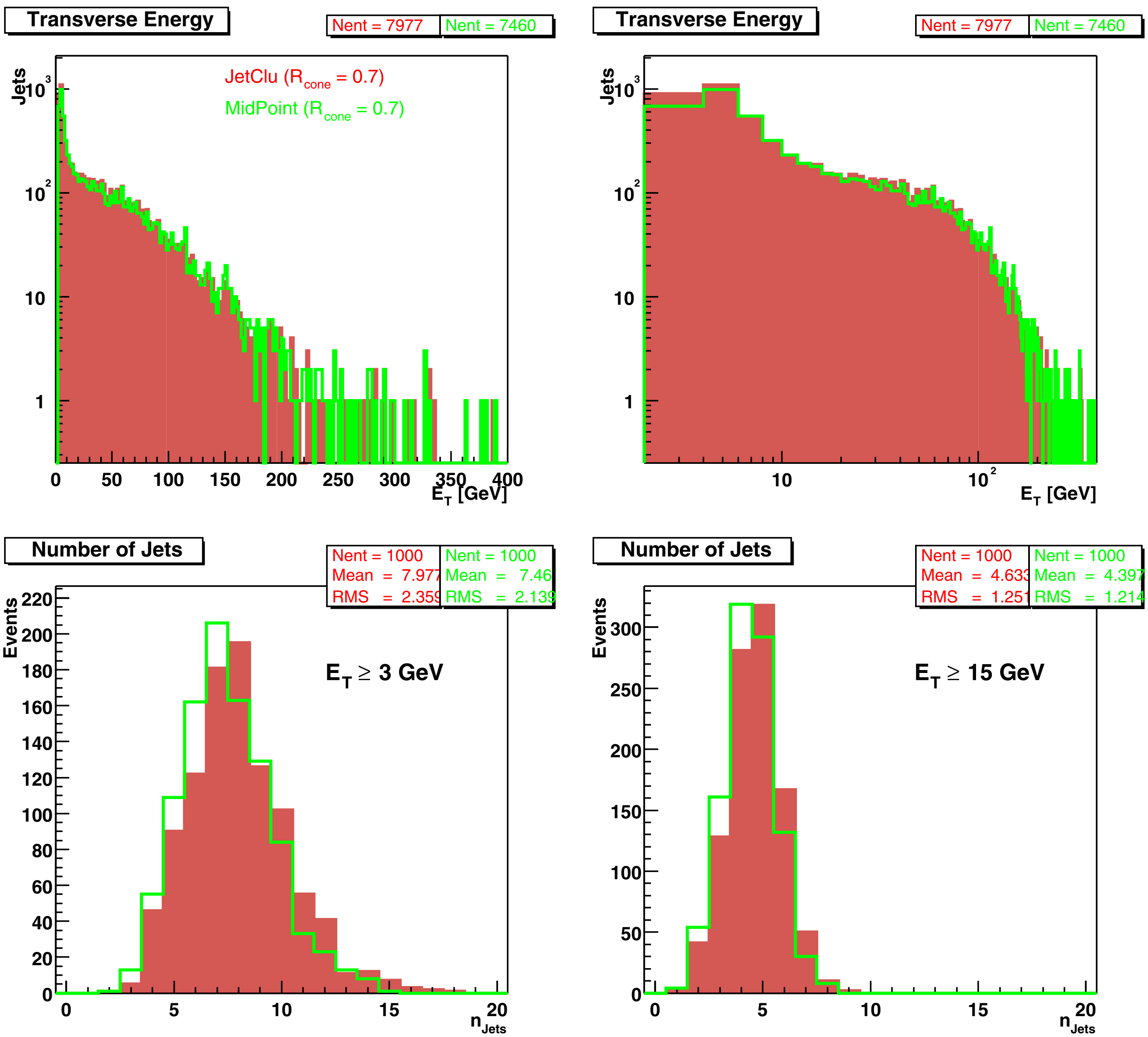
Comparison of MidPoint with JetClu

Jet50: JetClu vs. MidPoint (R_{cone}=0.7): Leading Jets



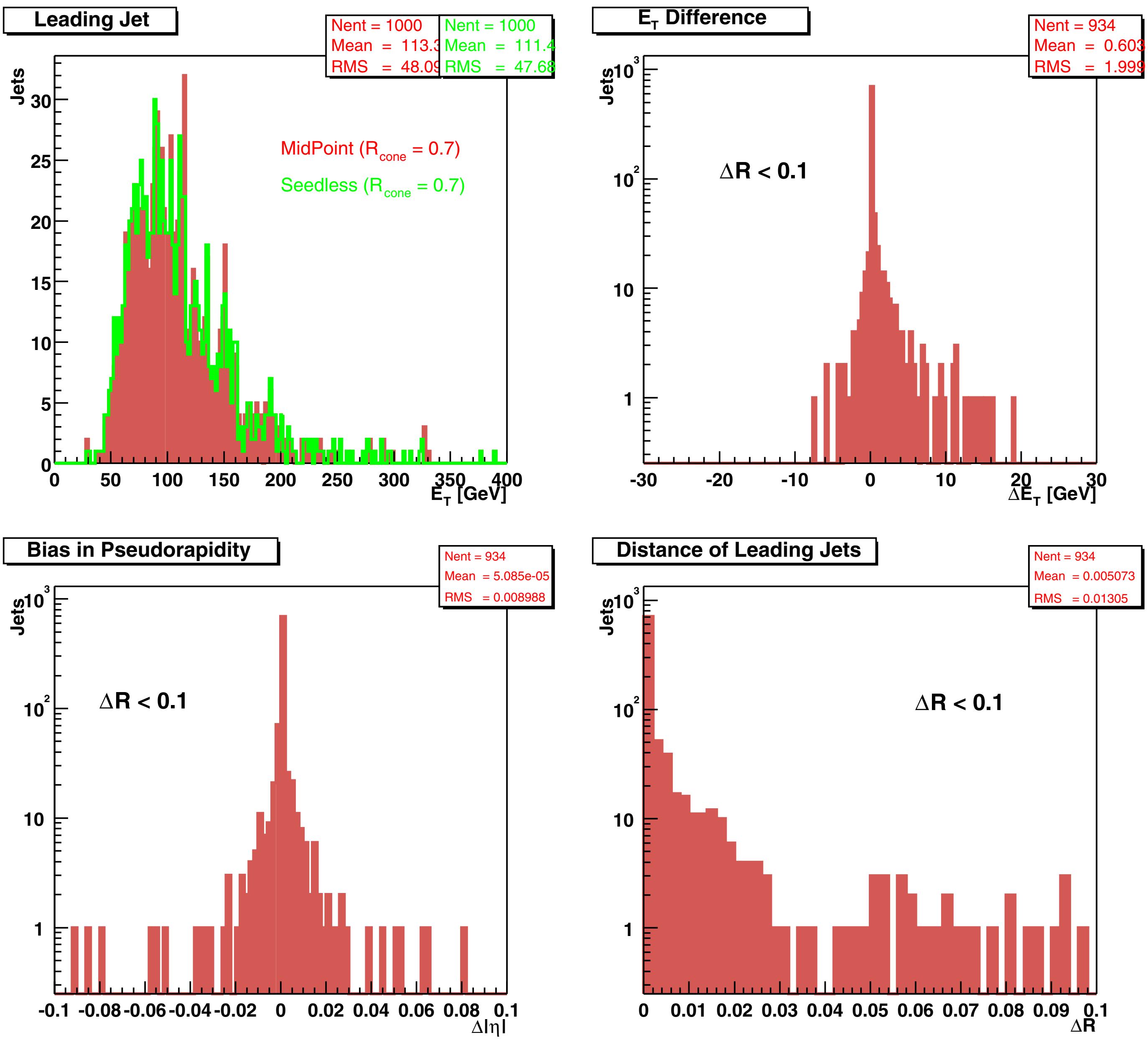
Comparison of MidPoint with JetClu

$t\bar{t}$: JetClu vs. MidPoint (R_{cone}=0.7): E_T and n_{jets}



Comparison of MidPoint with JetClu

$t\bar{t}$: MidPoint vs. Seedless (R_{cone}=0.7): Leading Jets



Conclusions

- Lot of progress in understanding jet clustering issues.
- Detailed clustering algorithms have been specified.
- Progress in energy calibration of k_T jets.
- The seedless algorithm is not two slow. Preclustering is needed for DZero.
- Ready to compare data with NNLO calculations
- Can we reach 1% accuracy in PDF? Still open.