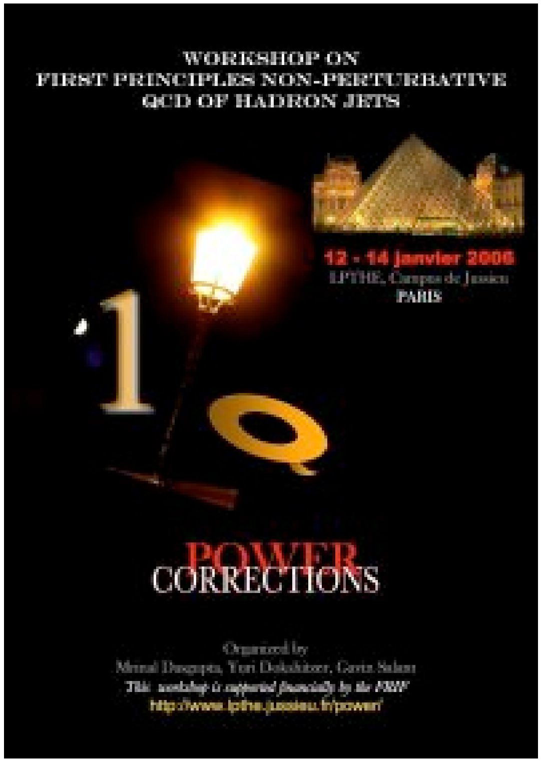
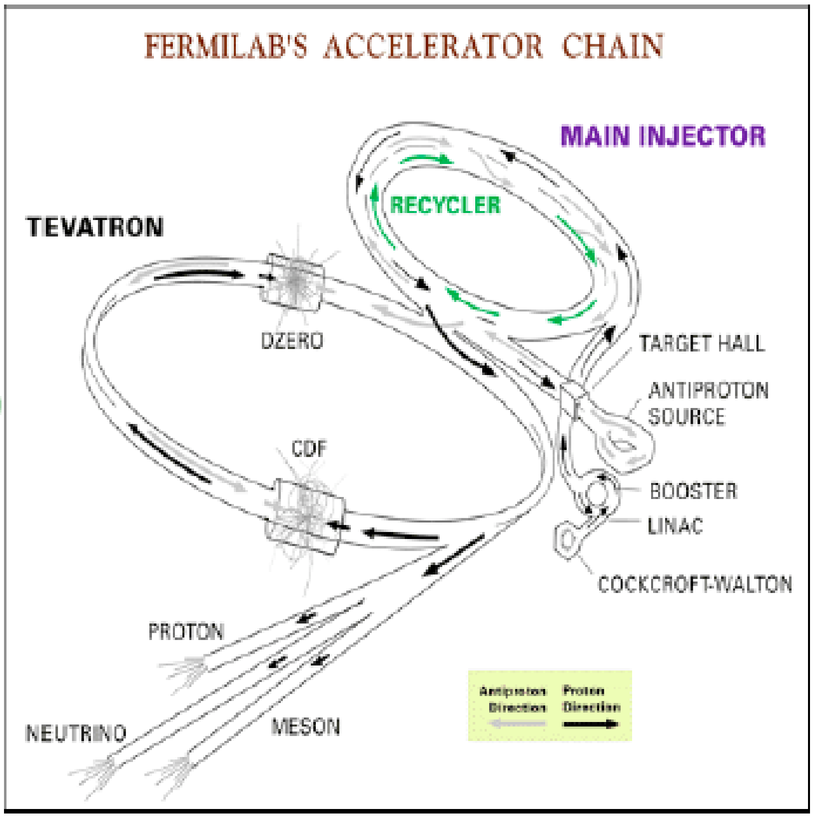
The inclusive jet cross section, jet algorithms, underlying event and fragmentation corrections

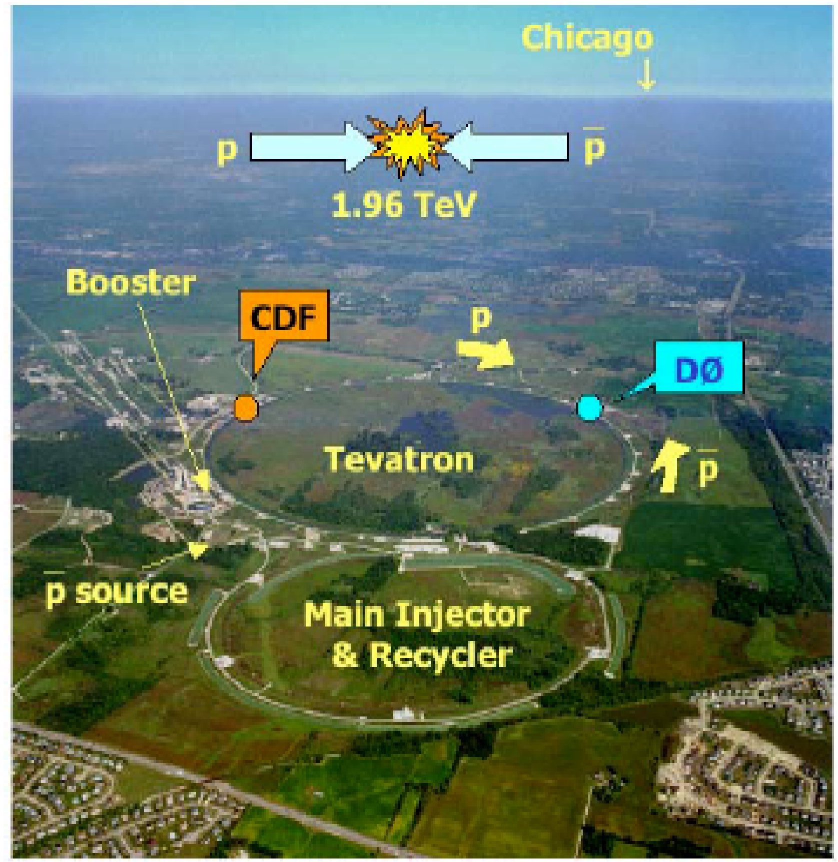
J. Huston
Michigan State University



Tevatron in Run II

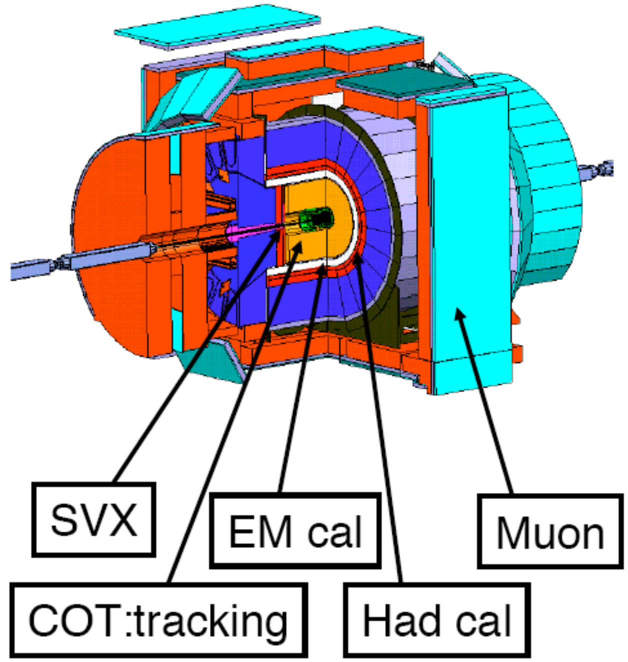
36 bunches (396 ns crossing time)

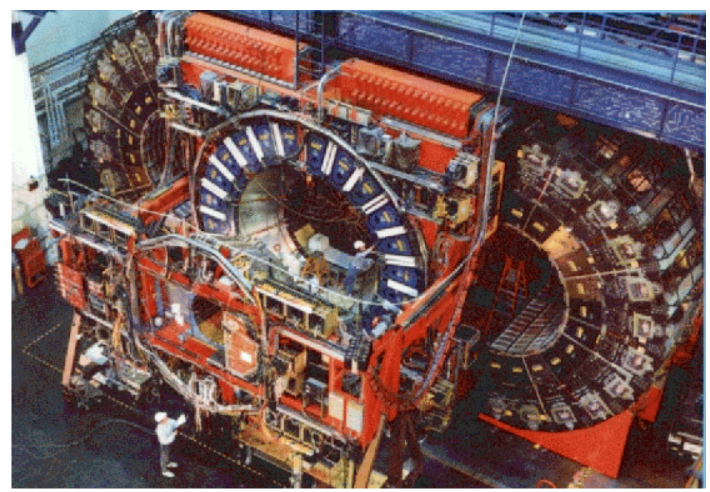




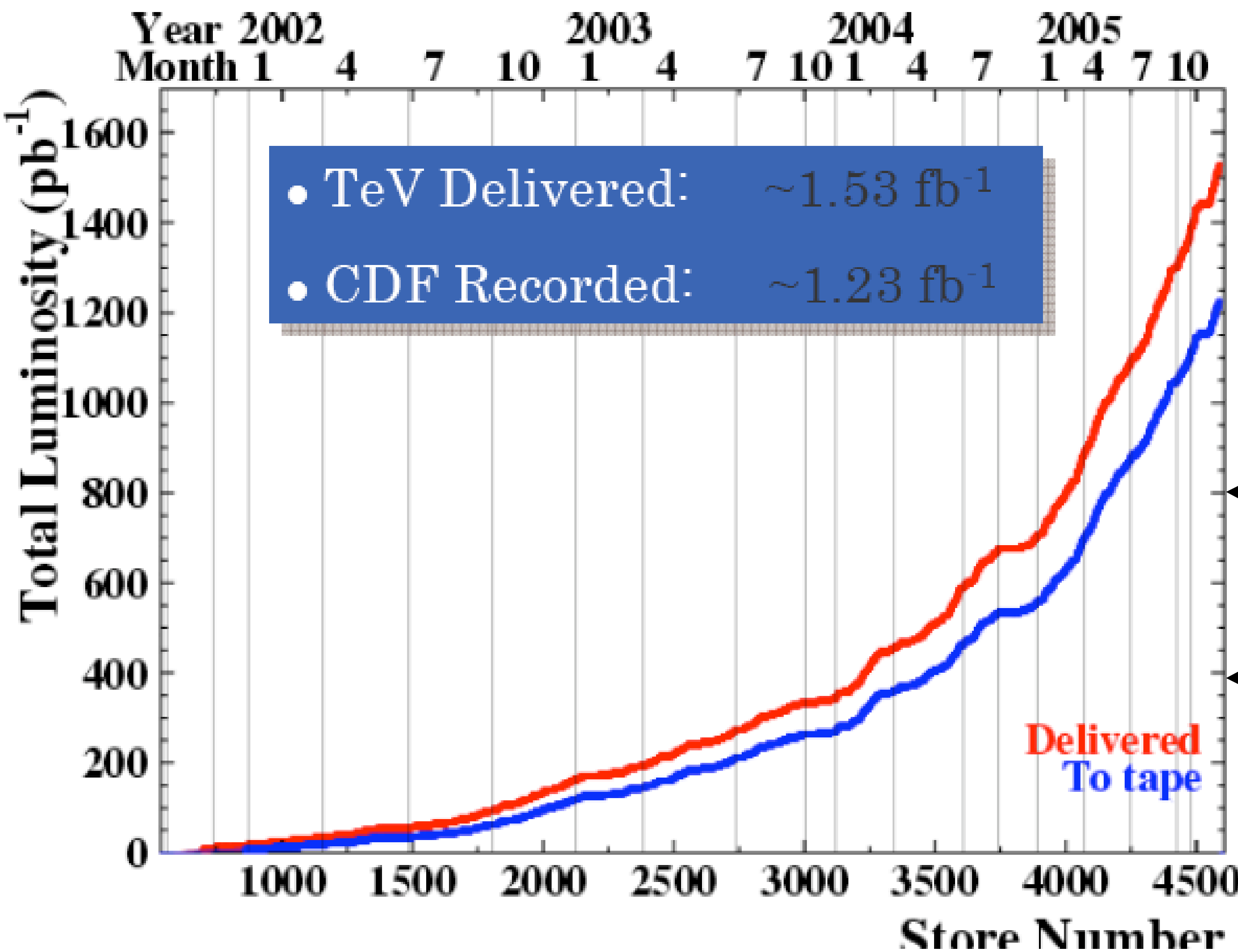
CDF in Run II

- Silicon detector (SVX):
  - top event b-tag: ~ 55%
- COT: drift chamber
  - Coverage: |η|<1
  - $\sigma_{Pt} / Pt \sim 0.15\% \, P_T$
- Calorimeters:
  - Central, wall, plug
  - Coverage: |η|<3.6
  - EM: $\sigma_E / E \sim 14\% / \sqrt{E}$
  - HAD: $\sigma_E / E \sim 80\% / \sqrt{E}$
- Muon: scintillator+chamber
  - muon ID up-to |η|=1.5





Run 2 luminosity



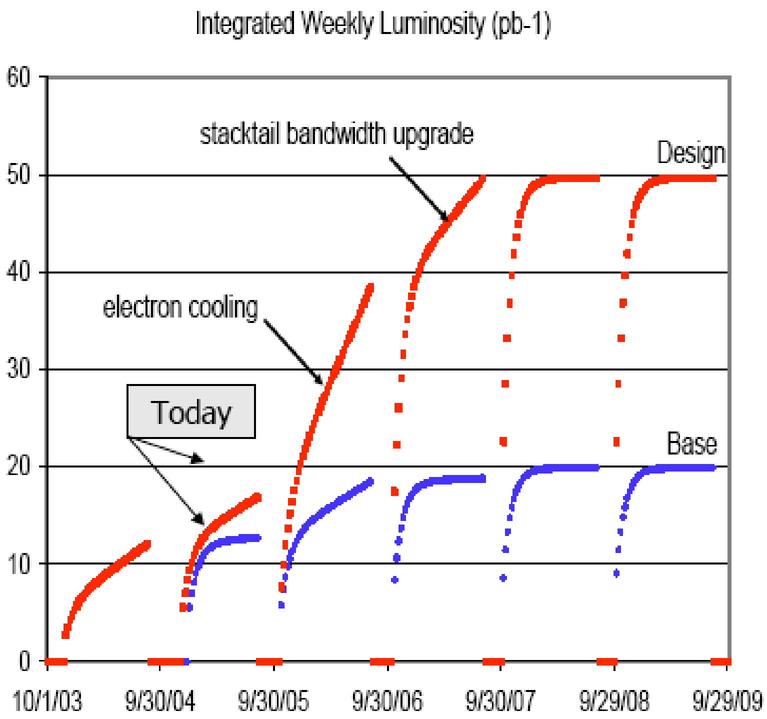
Moriond

most published analyses

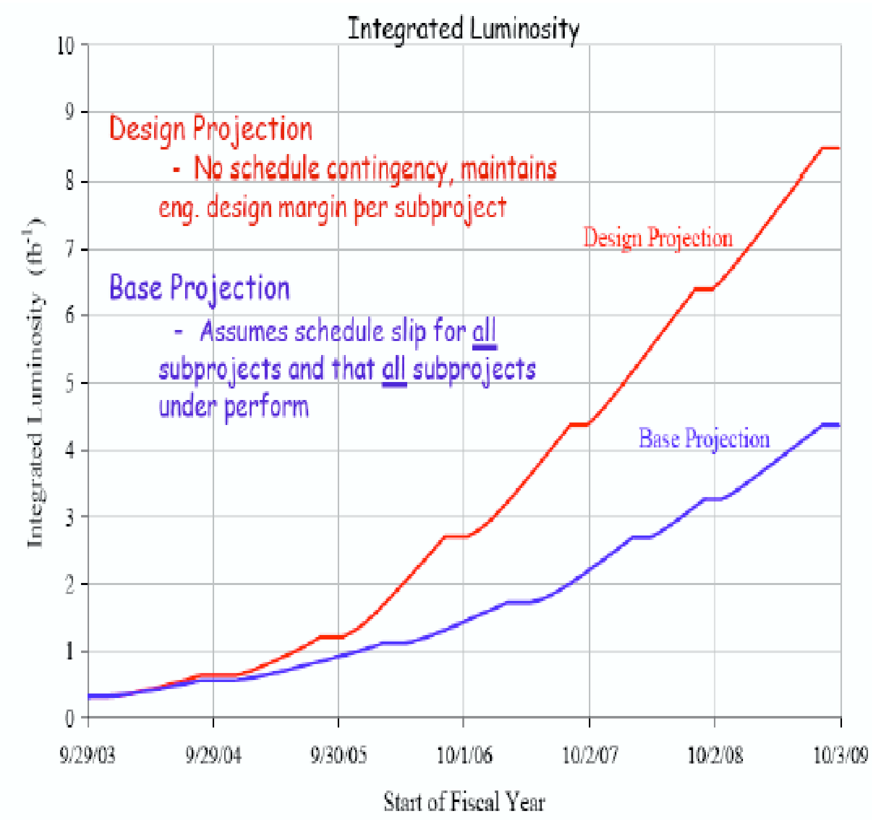
Reminder: ultimate goals

Increase in number of antiprotons
→ key for higher luminosity

Expected peak luminosity
→ $3 \cdot 10^{32}$ cm^{-2}sec^{-1} by 2007



ultimately 4-9 fb^{-1}

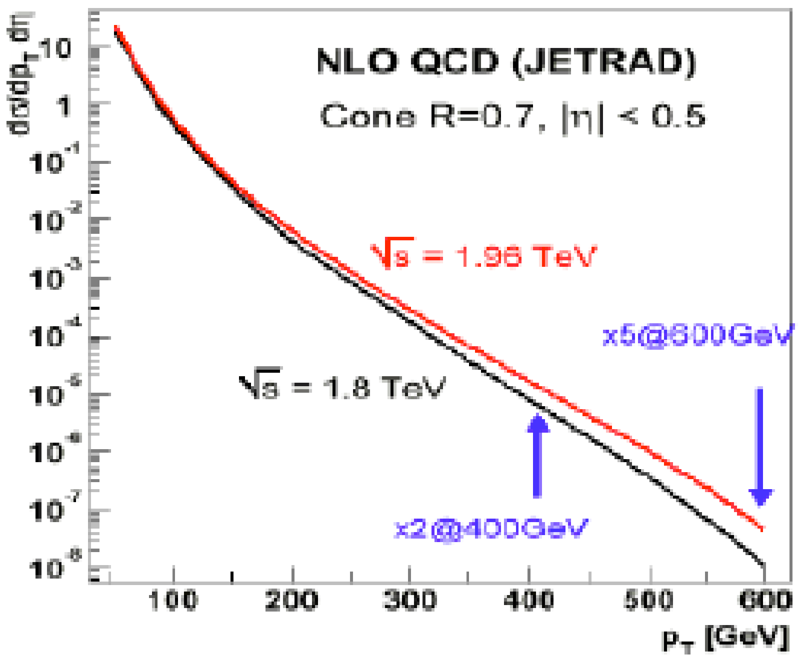


Why is the inclusive jet cross section interesting?

- Theoretically straightforward and hence a powerful test of pQCD
- Covers a wide range of transverse momentum
  - sensitivity to the running of α_s
- Probes small distance scales $\lambda \sim 1/p_T \sim 1/600$ GeV $\sim 10^{-18}$ m
  - sensitive to new physics such as quark compositeness
- Probes large x->helps to constrain gluon distribution at large x

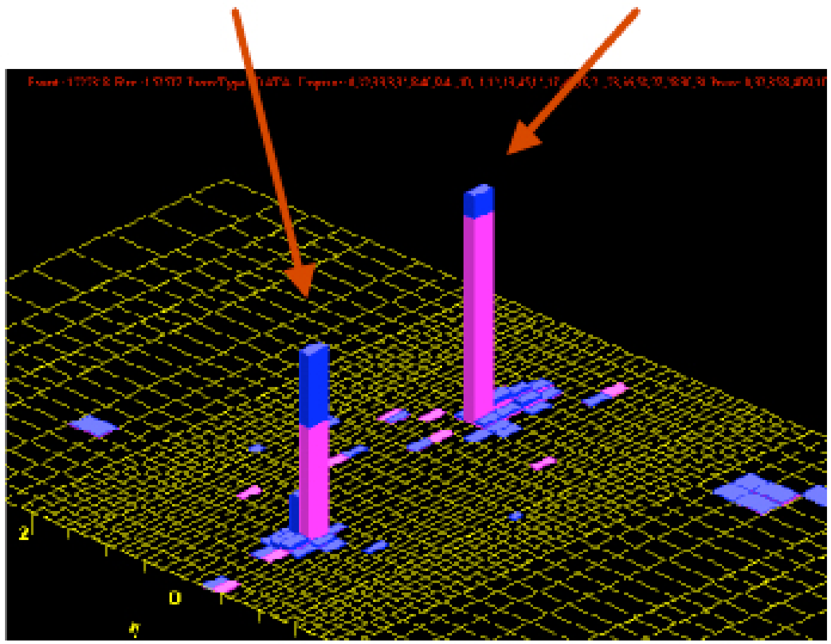
Jet Production at the Tevatron

- Nowhere is the increase in center of mass energy more appreciated



J2 E_T = 633 GeV (corr)
546 GeV (raw)
J2 η = -0.30 (detector)
= -0.19 (correct z)

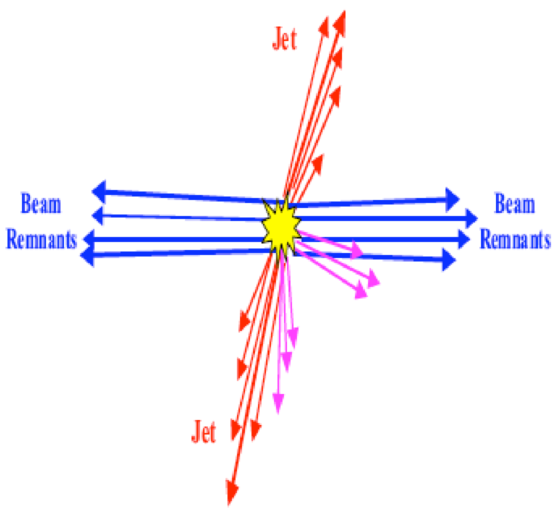
J1 E_T = 666 GeV (corr)
583 GeV (raw)
J1 η = 0.31 (detector)
= 0.43 (correct z)

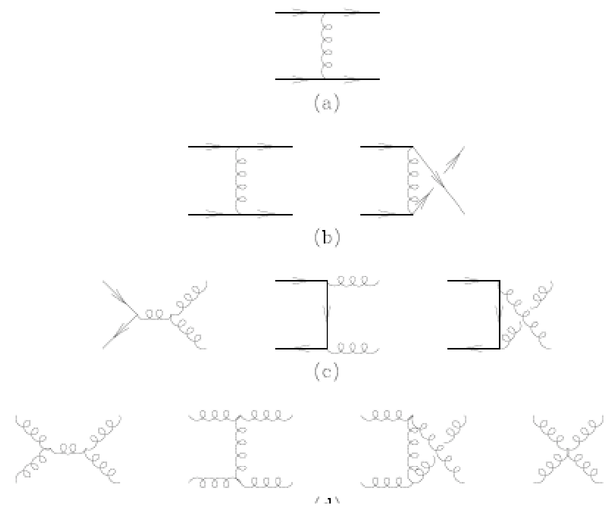


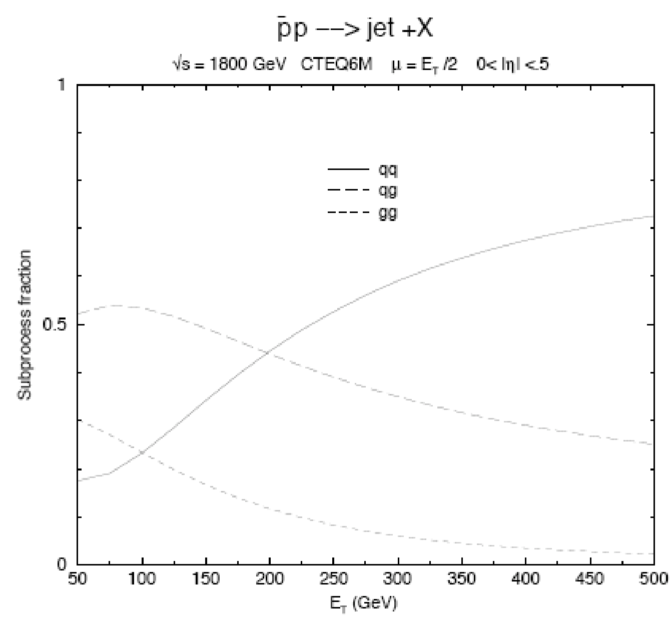
CDF Run 2 Preliminary

Jet Production

- 2->2 hard-scattering with gg dominant at low p_T and qq at high p_T
- Additional radiation from initial and final states
- Interaction of beam remnants including semi-hard DPS
- Measure either charged particle transverse momenta

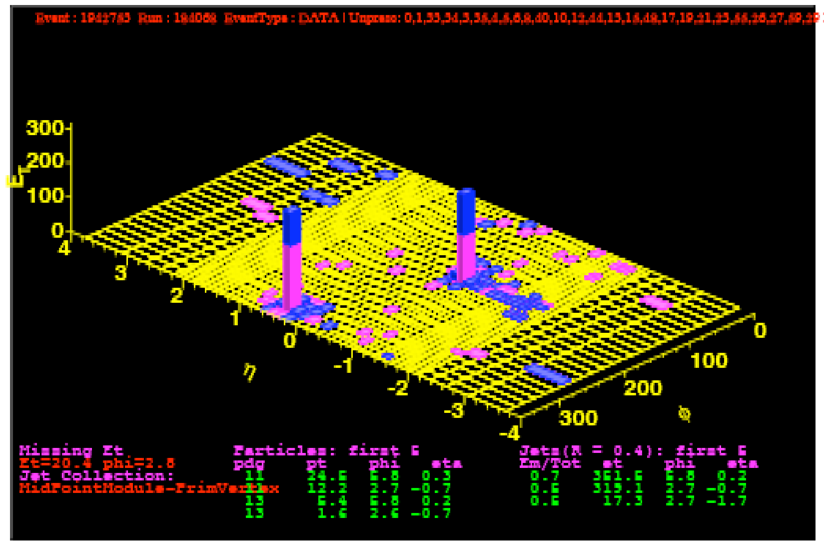


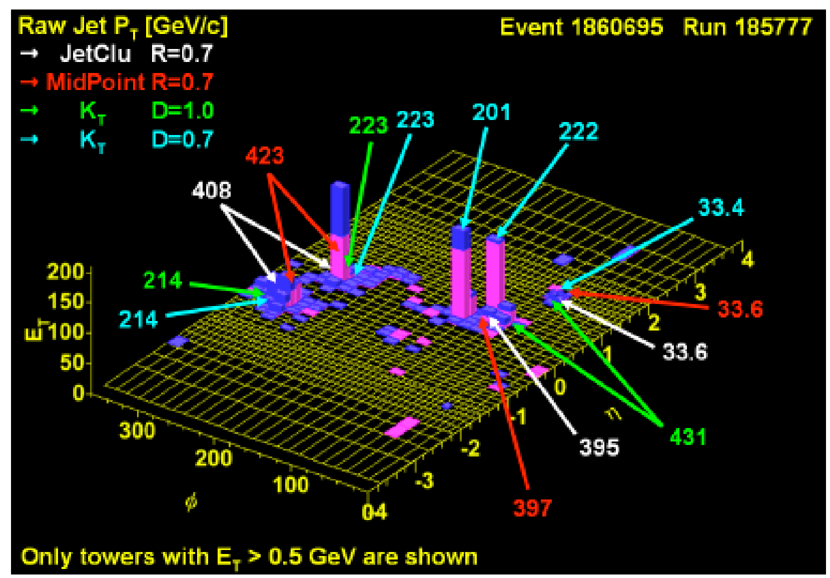




Jet Algorithms

- Need an algorithm(s) to define jet at trigger level and offline
- Sometimes the structure is very straightforward (2 cleanly separated jets) and sometimes it is not
- Algorithm must be able to handle all cases and as much as possible have similar behavior at the parton, particle and calorimeter tower level



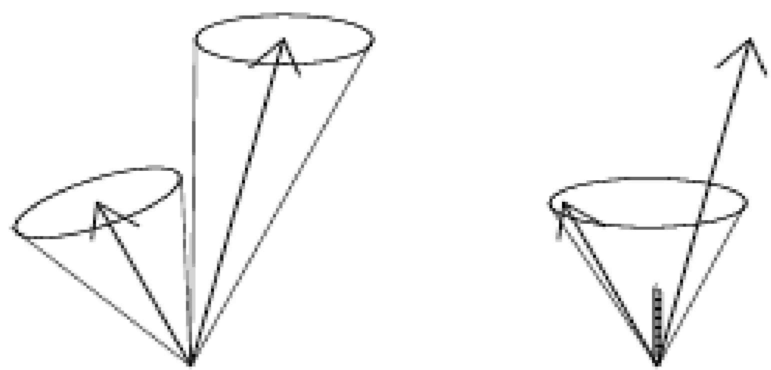


Midpoint cone algorithm

- Begin with 1 GeV seed towers
- Cluster towers with p_T>100 MeV/c into jet if

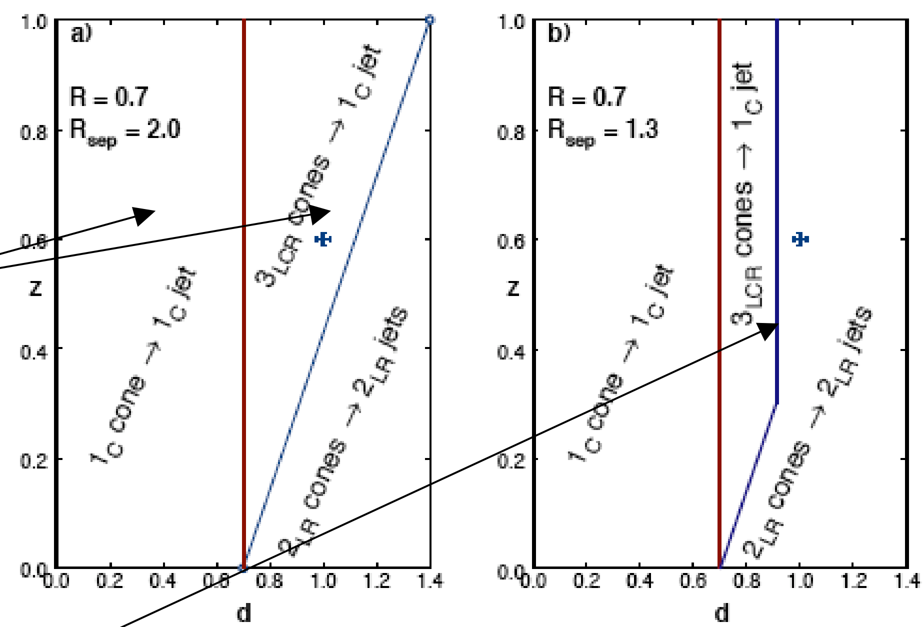
  $$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.7$$
- Start new search cones at the midpoints of stable cones
- Merge jets if overlapping energy is > 0.75 times the energy of the smaller jet
- Calculate jet quantities from stable cones
  - p_T, y, ϕ

- Clustering begins around seeds (bad thing)
- Presence of soft radiation can cause merging of jets (bad thing)
- Addition of seeds at midpoints lessens the sensitivity (good thing)



Aside: some issues regarding cone algorithms

- see hep-ph/0111434, S. Ellis, J. Huston, M. Tonnesmann, *On Building Better Cone Jet Algorithms*
- Consider the phase space for 2 partons, plotting $z(=E_{parton2}/E_{parton1})$ vs d (ΔR for the two partons)
- Naïve expectation was that all partons in these two areas would lead to the reconstruction of 1 jet in the data
- But it was found in the data that jets separated by $1.3{*}R_{cone}$ were almost always reconstructed as separate jets
- So an R_{sep} factor of 1.3 was added to the theory calculation



Example

- Consider 2 partons separated by ΔR=1.0 with z=0.6
- Are they reconstructed at 1 jet or 2 jets?
- At pure parton level, yes; at hadron level no
- No stable solution due to smearing by parton showering/hadronization
- 2nd parton is lost

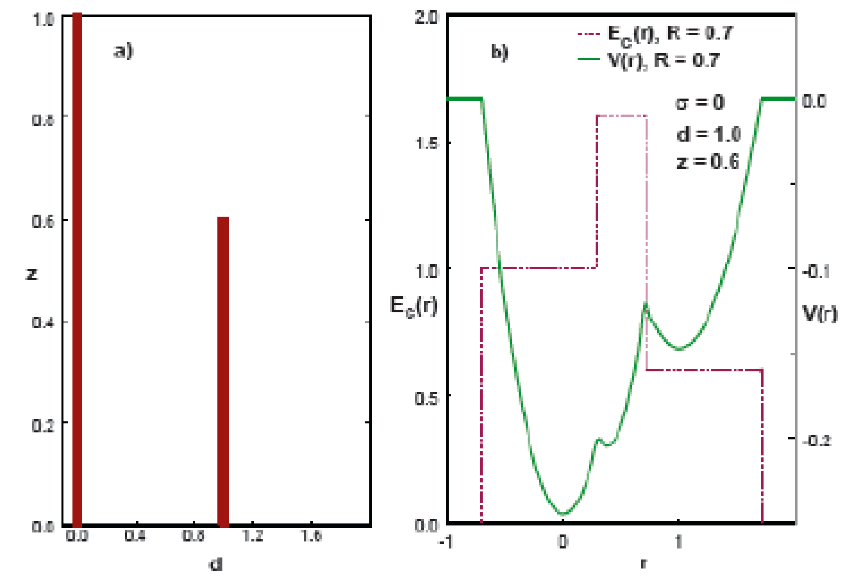


FIG. 2: 2-Parton distribution: a) transverse energy distribution; b) distributions $V(r)$ and $E_C(r)$ in the perturbative limit of no smearing.

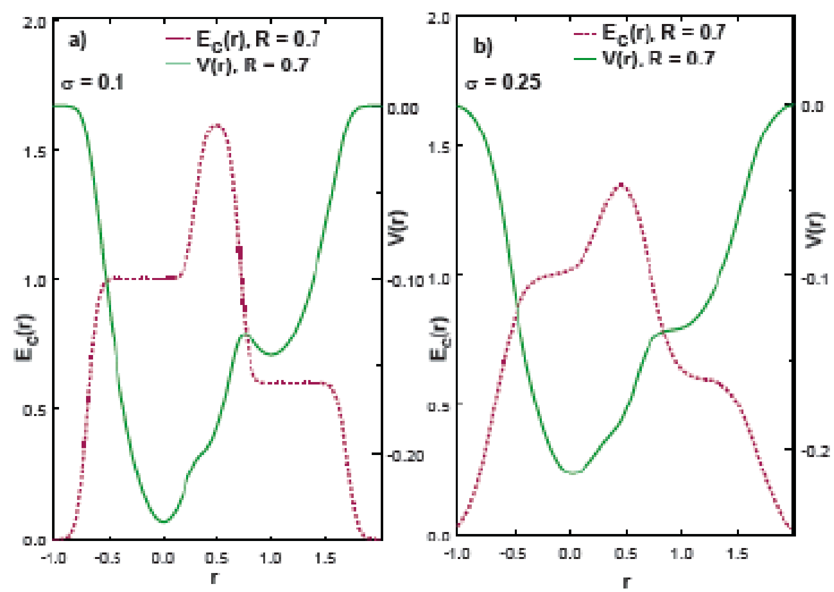
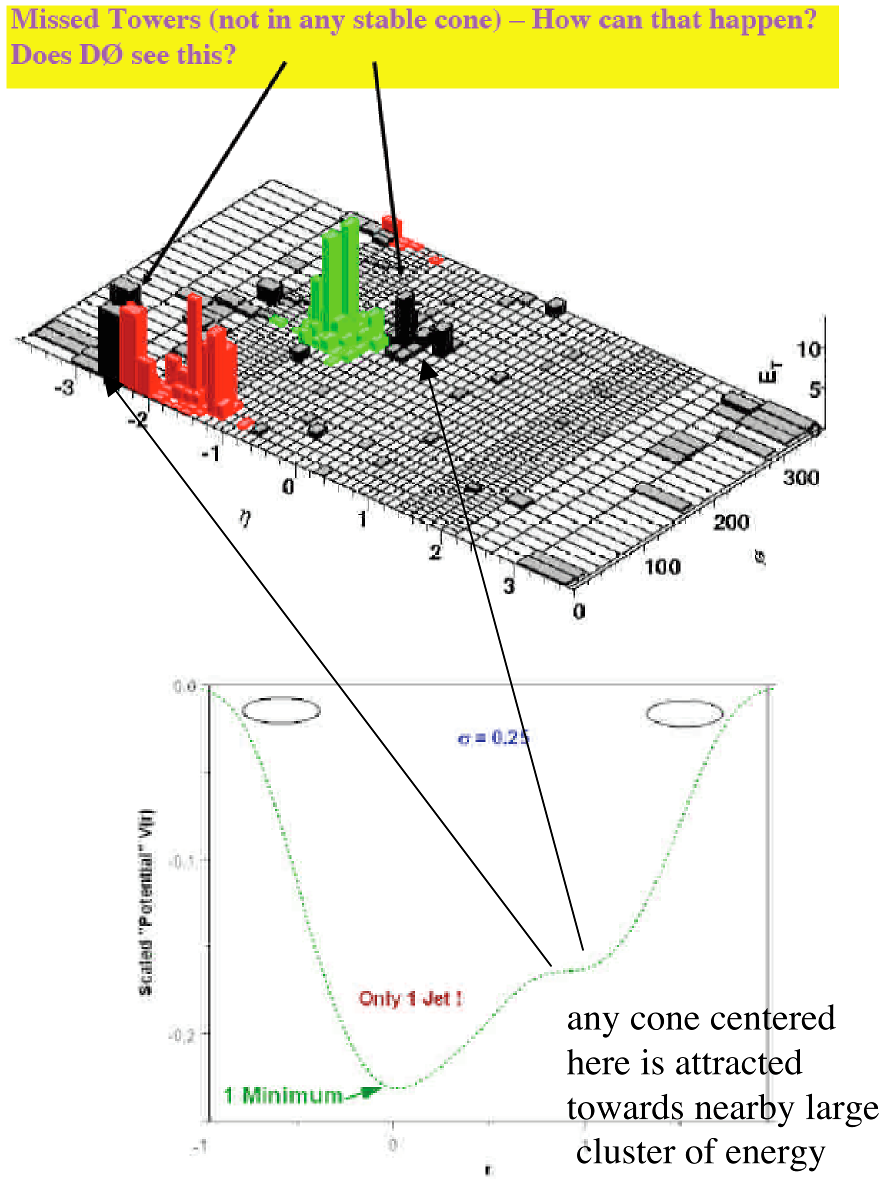


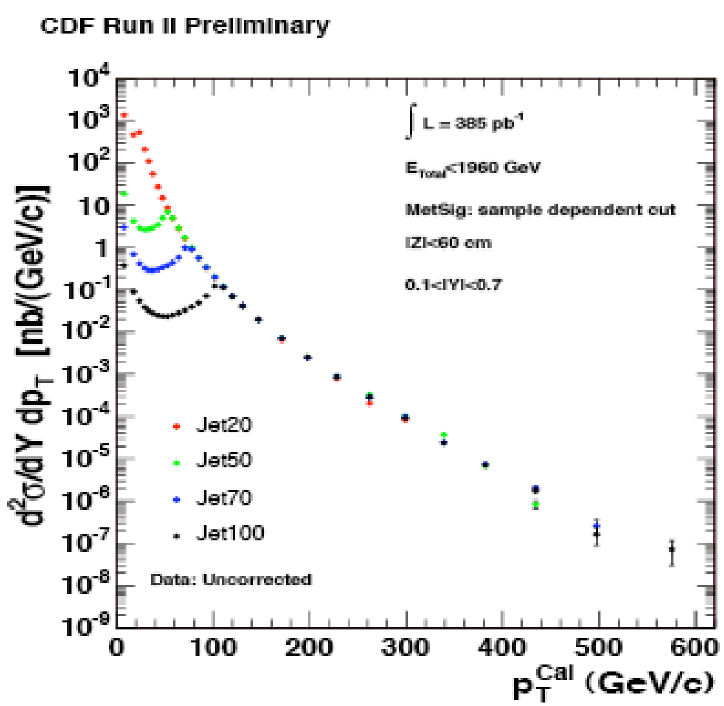
FIG. 3: The distributions $V(r)$ and $E_C(r)$ with $d = 1.0$ and $z = 0.6$ for smearing width $\sigma =$ a) 0.1; b) 0.25.

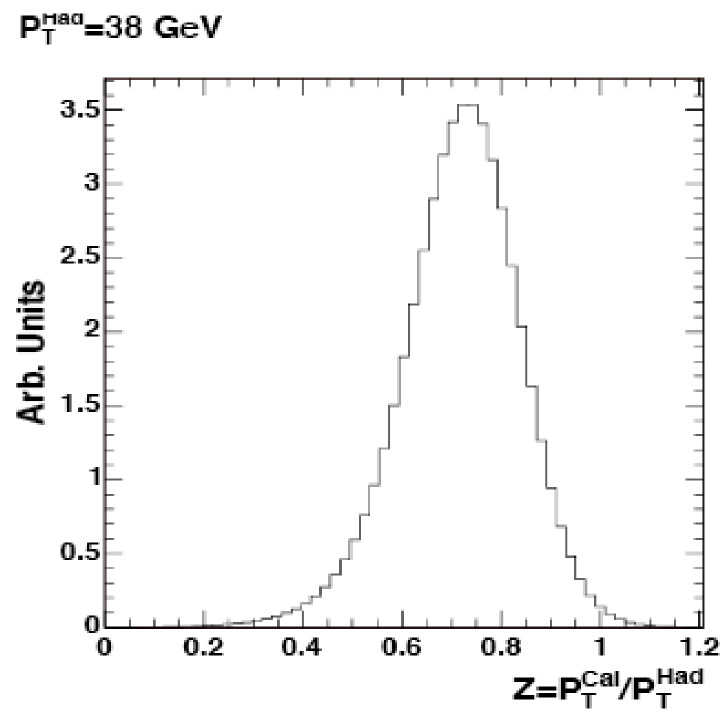
Solution to dark tower problem

- Implement an initial search cone step with the search cone size = $R_{cone}/2$
- Less sensitive to effects of *great attractors* far away
- After stable cones are formed, expand jet cones to full size and decide whether to split/merge overlapping jets according to the standard criteria
- ~5% effect on the cross section independent of p_T



Raw cross section

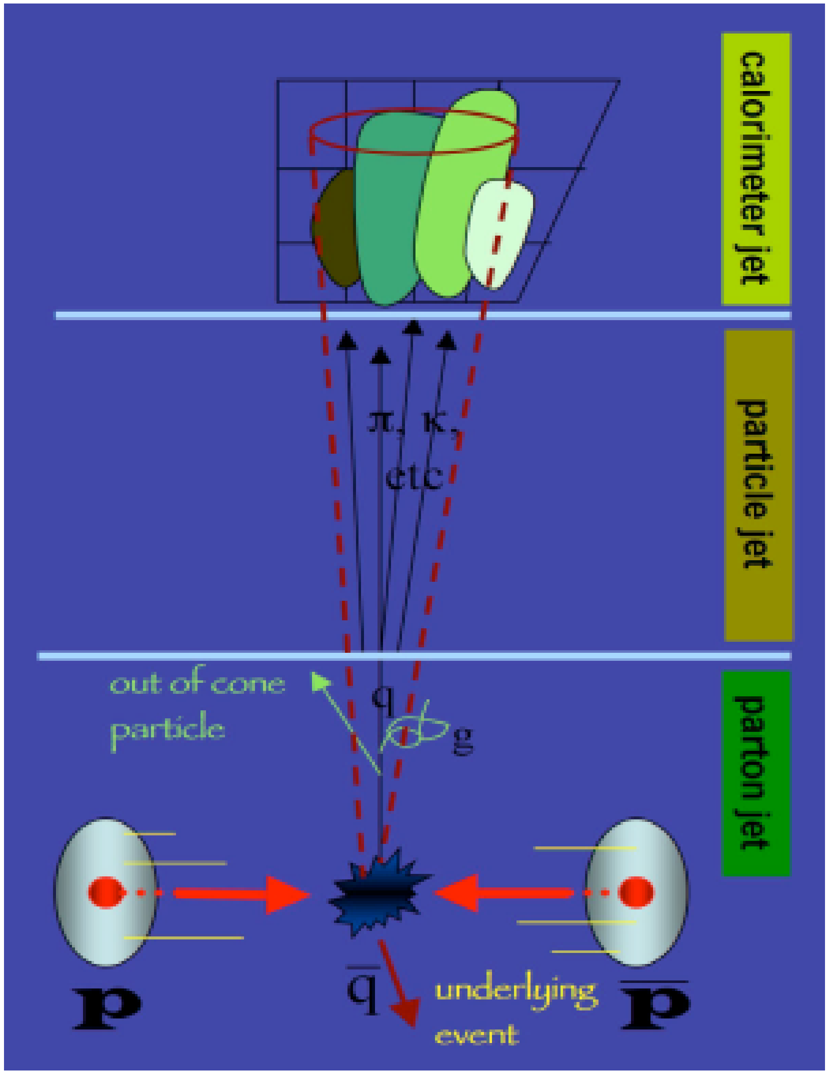




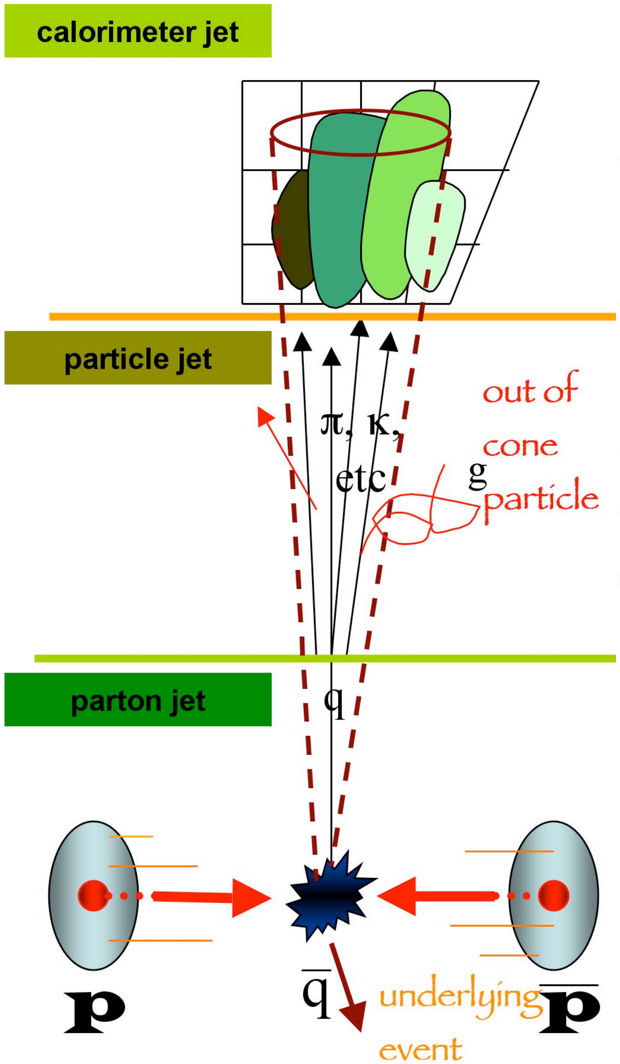
* Need to understand and correct for detector effects and additional contributions to the jet cone that are not present in the theory.

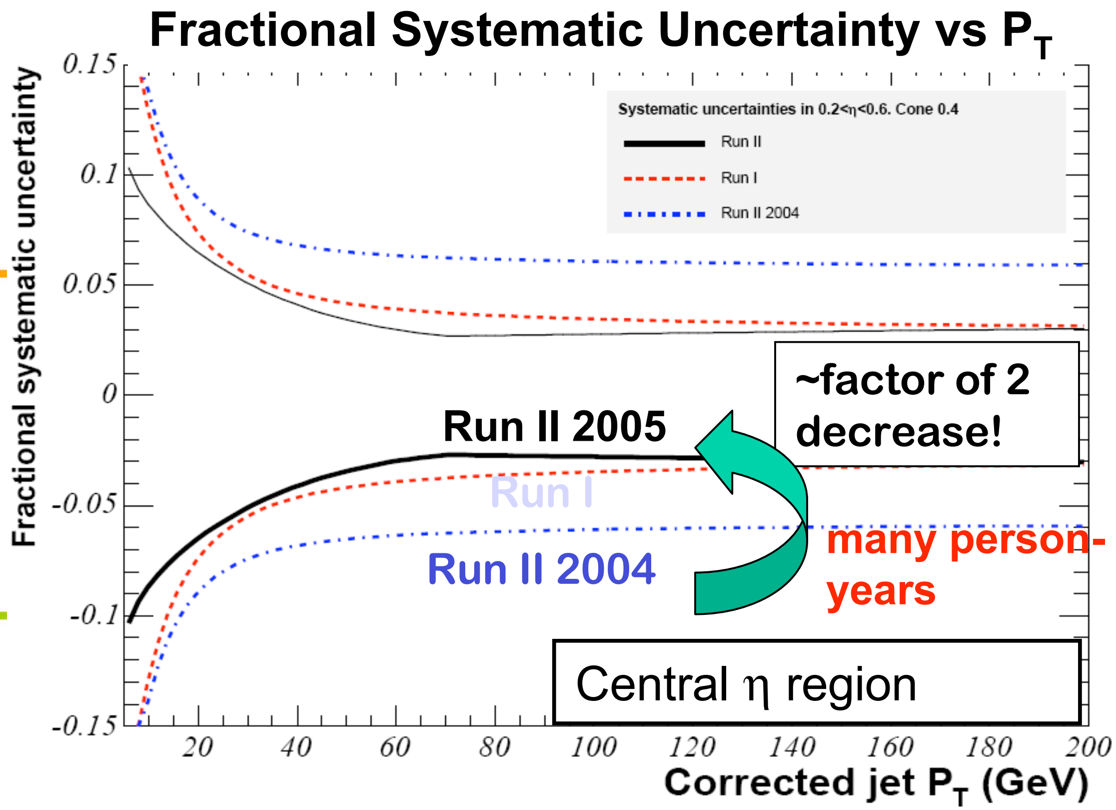
Jet Corrections

- Need to relate what we see at the calorimeter level to the hadron and parton level
- From data
  - correct for multiple interactions in a single crossing
- From Monte Carlo (Pythia)
  - correct for absolute scale and smearing
    - calorimeter to hadron level corrections
  - correct for underlying event and fragmentation
    - hadron to parton level corrections



CDF Jet Energy Scale: New

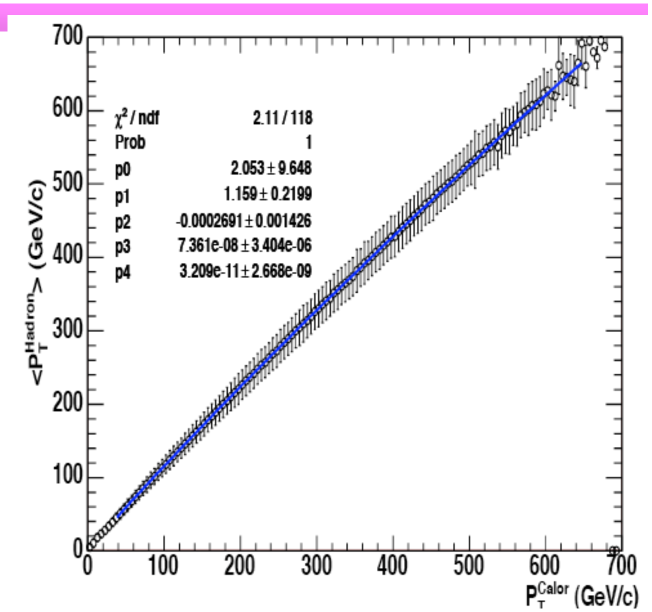


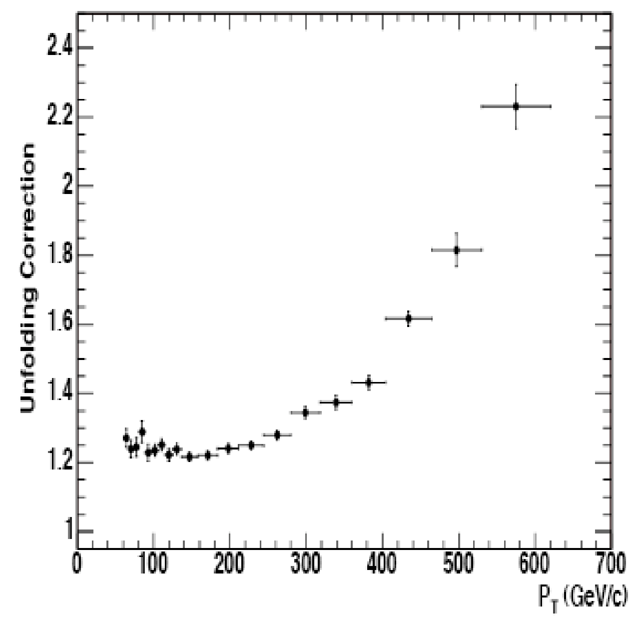


have to correct calorimeter energy depositions for detetector, algorithm and physics effects to obtain "true" jet energy

Calorimeter to hadron corrections

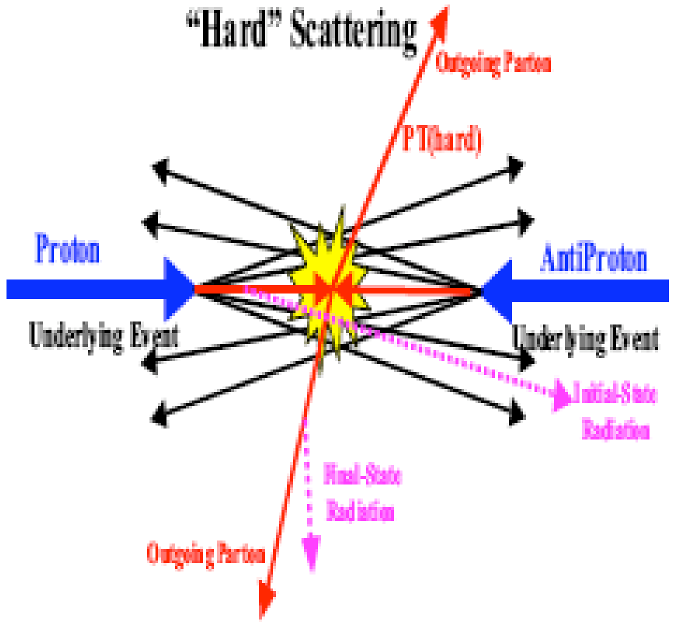
- Jet by jet correction is applied to correct for absolute scale
    - corrects for hadron energy not sampled by calorimeter
    - MC jets matched at hadron and calorimeter level
- Then spectrum is unfolded
    - accounting for smearing and any underestimate of the absolute correction of step 1
- Cross section is now at *hadron level*

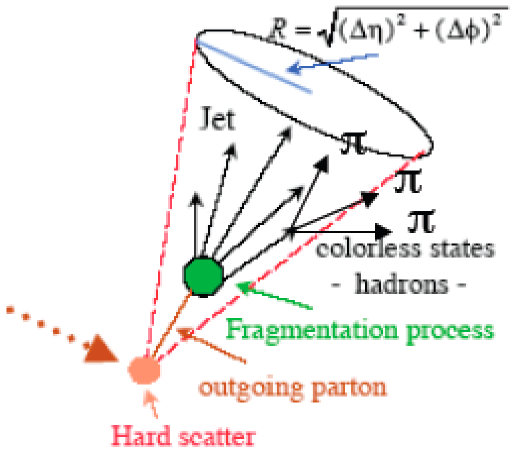




Hadron to parton corrections

- Now have to correct for underlying event and for fragmentation
- Underlying event
  - energy due to beam-beam remnants, soft spectator interactions, multiple parton scattering
  - but at least part of ISR/FSR included in NLO calculation
- Fragmentation
  - energy lost outside the cone during hadronization
    - hadrons ending up outside the cone from partons whose trajectories lie within the cone





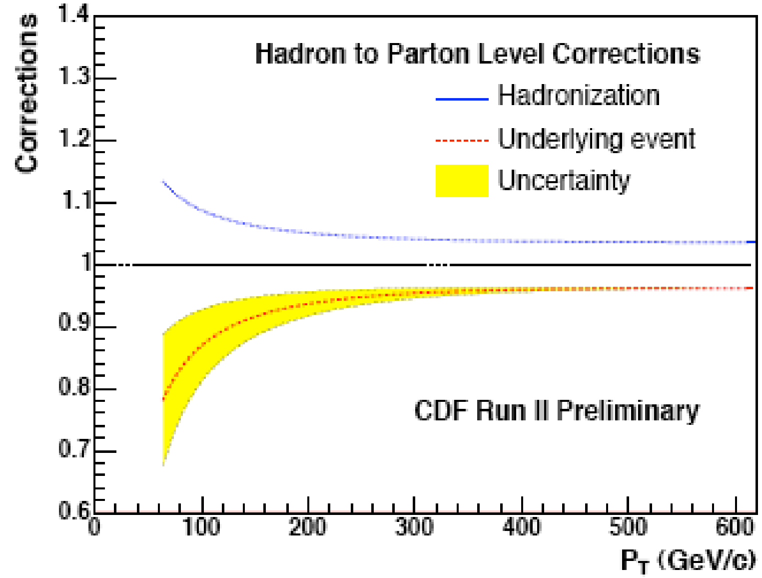
Hadron to parton correction

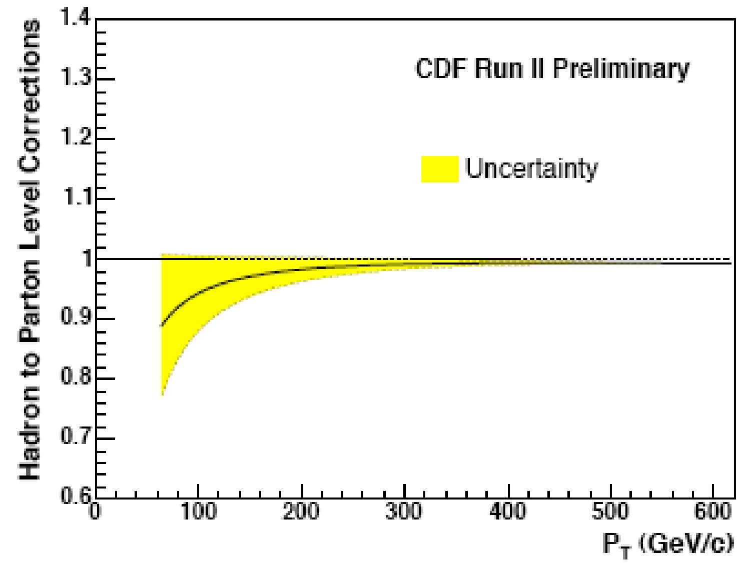
- Corrections determined by running Pythia (with Tune A) with/without UE and fragmentation
- The two corrections go in the opposite direction and (mostly) cancel out with a net decrease in the cross section at low p_T

$$C_i^{UE} = \frac{\sigma_i^{hadron(UE)}}{\sigma_i^{hadron(no-UE)}}$$

$$C_i^{frag} = \frac{\sigma_i^{hadron(no-UE)}}{\sigma_i^{parton(no-UE)}}$$

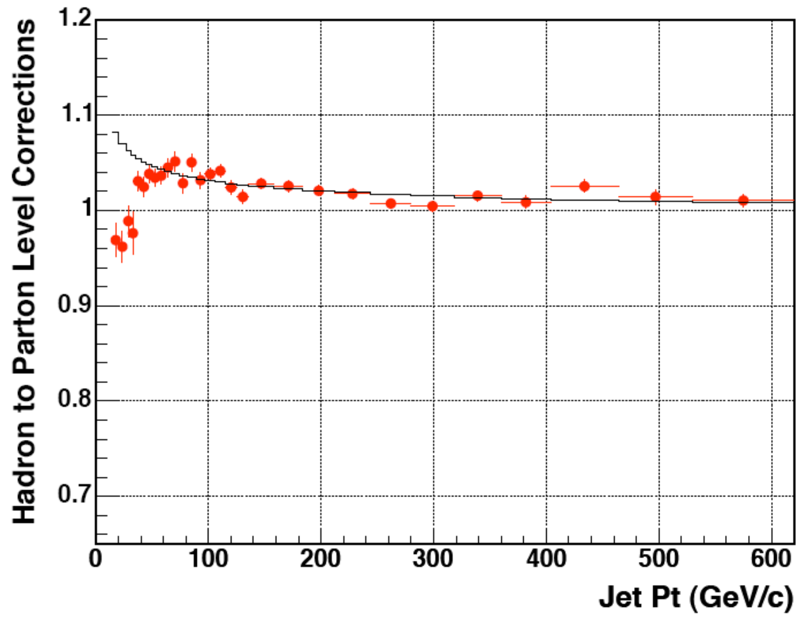
$$C_i^{UE}.C_i^{frag} \rightarrow C_i^{had\rightarrow parton} = \frac{\sigma_i^{hadron(UE)}}{\sigma_i^{parton(no-UE)}}$$





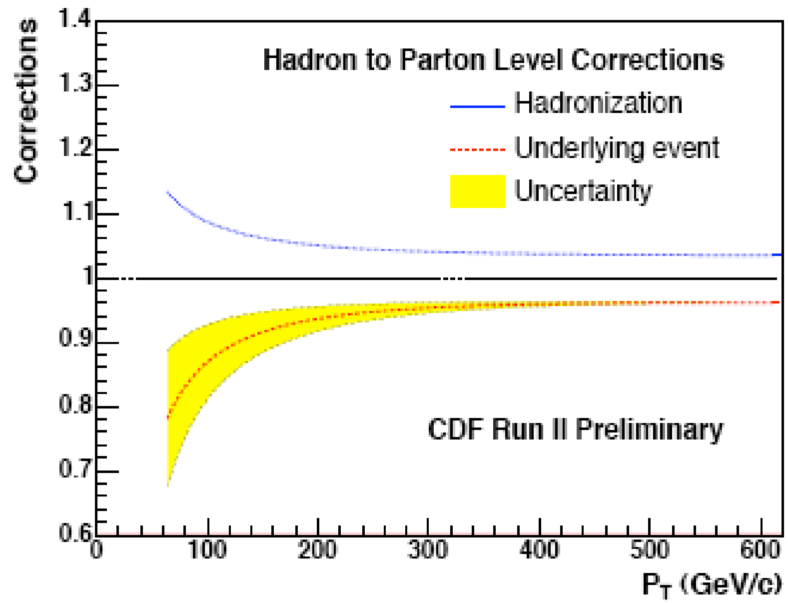
Jet cone of 0.4

- Interestingly enough, the combined effects of underlying event and fragmentation cancel to within a few percent for a cone of radius 0.4
  - underlying event subtraction is smaller



Hadron to parton corrections continued…

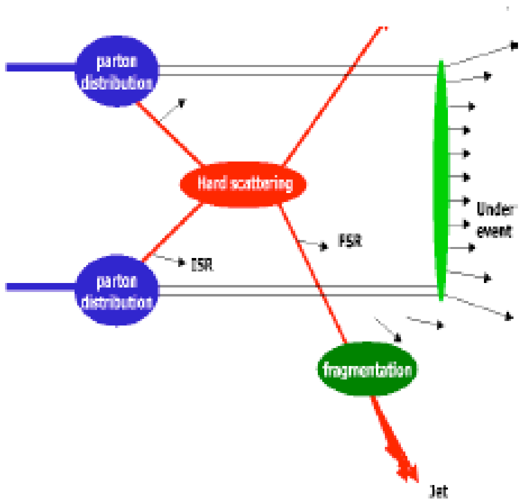
- Pythia with Tune A describes a wide variety of global event observables at the Tevatron
- Herwig, with its parameterization of UA5 data for the underlying event, tends to provide too soft a description
- Difference is taken as systematic error for underlying event subtraction
- Herwig and Pythia give essentially the same result for fragmentation correction so no error is assumed
    - fragmentation correction: a gluon produces an A_1 which then decays into 3 pions; one of the pions is kicked outside the cone
    - NLO theory knows nothing about this and a correction must be made to the data



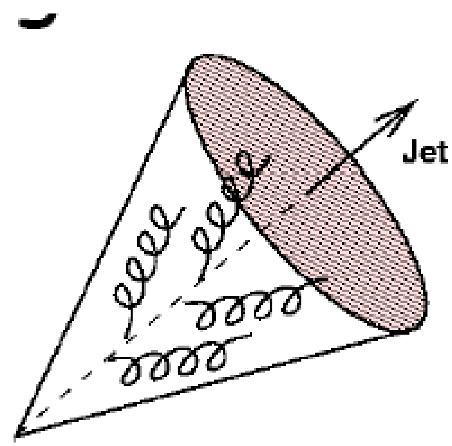
Question: is it fair to use a parton shower Monte Carlo (many partons) for the fragmentation corrections for a NLO calculation (1 parton)?

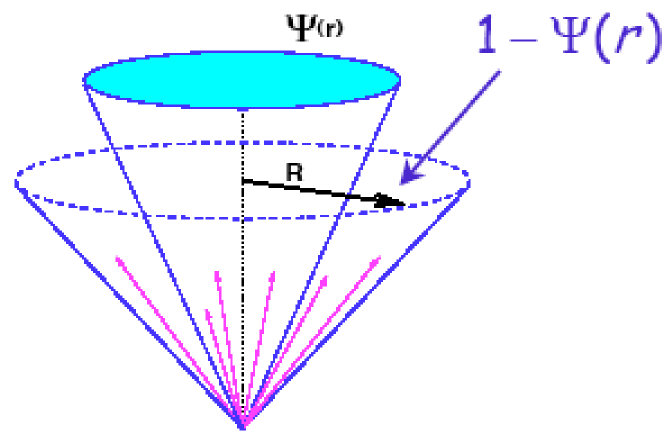
Answer: yes, to the extent that both describe the jet shape

CDF:Jet Fragmentation



- Jet shape dictated by multi-gluon emission form primary parton
- Test of parton shower models and their implementations
- Sensitive to quark/gluon final state mixture and run of strong coupling
- Sensitive to underlying event structure in the final state

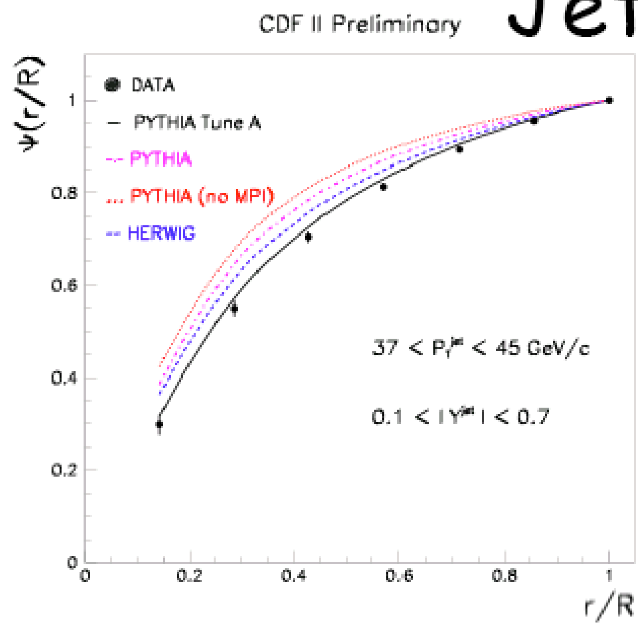


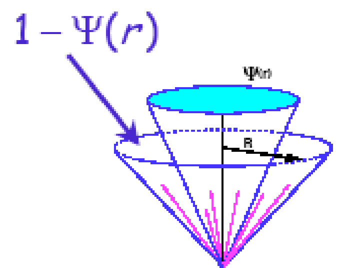


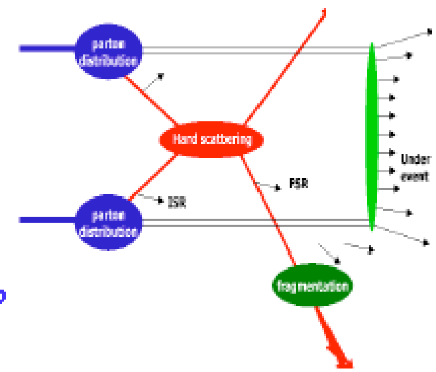
$$\Psi(r) = \frac{1}{N_{jets}} \Sigma_{jets} \frac{P_T(0,r)}{P_T^{jet}(0,R)}$$

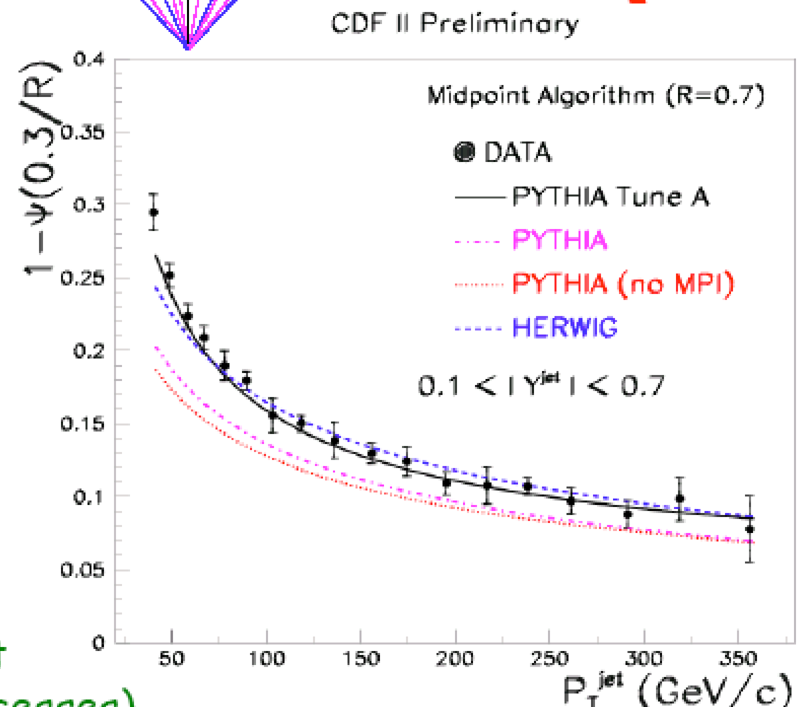
Jet Fragmentation

Jet shapes







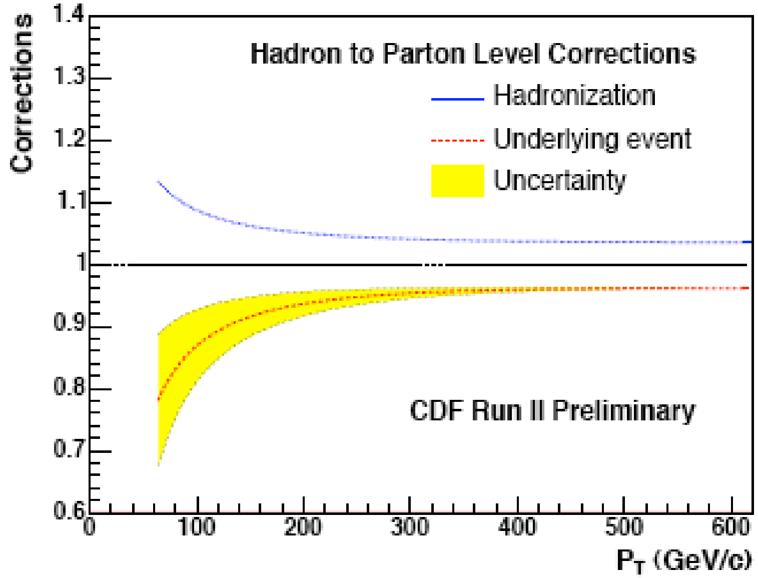


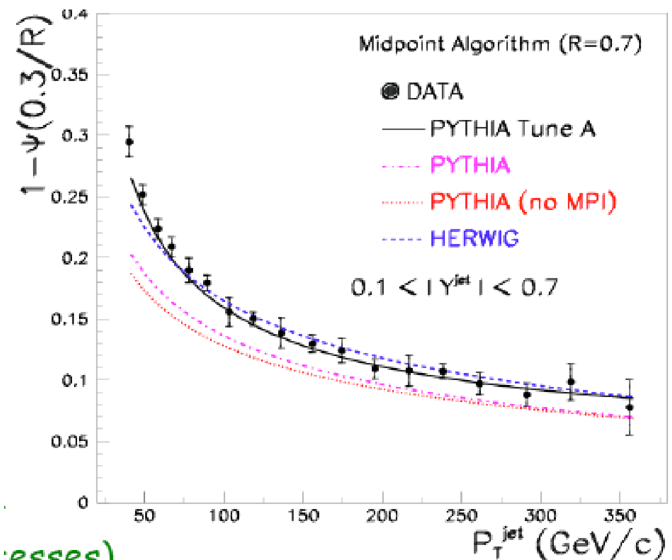
- PYTHIA Tune A describes the data (enhanced ISR + MPI tuning)
- PYTHIA default too narrow
- MPI are important at low Pt
- HERWIG too narrow at low Pt

We know how to model the UE at 2 TeV (at least for QCD jet processes)

Revisit fragmentation corrections

- Fragmentation corrections decrease in importance as p_T of jet increases (as expected)
- In fact, a closer look shows that the fragmentation correction is basically constant, on the order of 1 GeV/c, over the entire p_T range
- Even though the jet energy is increasing linearly, the p_T in the outermost annulus of the jet increases much more slowly as the jet becomes more collimated





:esses)

LEP study from 6-7 years ago

- Bill Gary did a very nice analysis of fragmentation/splashout at LEP (but never published as far as I know)
- Splashout is ~1 GeV
- Also get that value from a simple FF model

Consider the hadrons that represent the decay products of a high E_T parton. Let η be the rapidity of the hadrons relative to jet axis. Let $\vec{k}_T$ be the transverse momentum of the particles relative to jet axis. Let the distribution of hadrons be

$$\frac{dN}{d\eta d\vec{k}_T} = \frac{A}{\pi \langle k_T^2 \rangle} \exp\left\{-k_T^2/\langle k_T^2 \rangle\right\}, \tag{10}$$

where A is the number of hadrons per unit rapidity and $\langle k_T^2 \rangle$ is average k_T^2 of the hadrons. Then the E_T lost is approximately

$$E_T^{out} = \int_0^{\eta_1} d\eta \int d\vec{k}_T \, \frac{1}{2}|\vec{k}_T| e^{\eta} \frac{dN}{d\eta d\vec{k}_T}, \tag{11}$$

where $\eta_1 = -\ln(\tan(R/2))$. Performing the integral gives

$$E_T^{out} = \frac{\sqrt{\pi}}{4} A \sqrt{\langle k_T^2 \rangle} \left(e^{\eta_1} - 1\right). \tag{12}$$

Taking $\sqrt{\langle k_T^2 \rangle} = 0.3$ GeV and[10] $A = 5$, I find

$$E_T^{out} \approx 1.1 \text{ GeV}. \tag{13}$$

NEW LEP ANALYSIS

Measure All Energy Outside Leading Two Jets And Compare To Models

Cone 0.7:

Data (per jet) = 4.42 ± 0.27 GeV

Model	Hadron Level	Parton Level
Pythia 5.7	4.15 GeV	3.46 GeV
Herwig 5.8	4.28 GeV	3.60 GeV
NNLO	4.33 GeV	3.01 GeV

Cone 1.0:

Data (per jet) = 2.6 ± 0.18 GeV

Model	Hadron Level	Parton Level
Pythia 5.7	2.46 GeV	2.00 GeV
Herwig 5.8	2.55 GeV	2.14 GeV
NNLO	2.55 GeV	1.69 GeV

Conclusion:

NNLO Parton Level Is More Than 5σ From Data, About 0.9-1.4 GeV Extra Splashout Is Needed, Same As CDF Xt Scaling

Back to Run 1 (630 GeV)

- x_T scaling problem is primarily a low E_T problem (using jets down to 20 GeV/c at 630 GeV/c and ~3 times higher at 1800 GeV)
- Fragmentation corrections remove about half of the CDF discrepancy

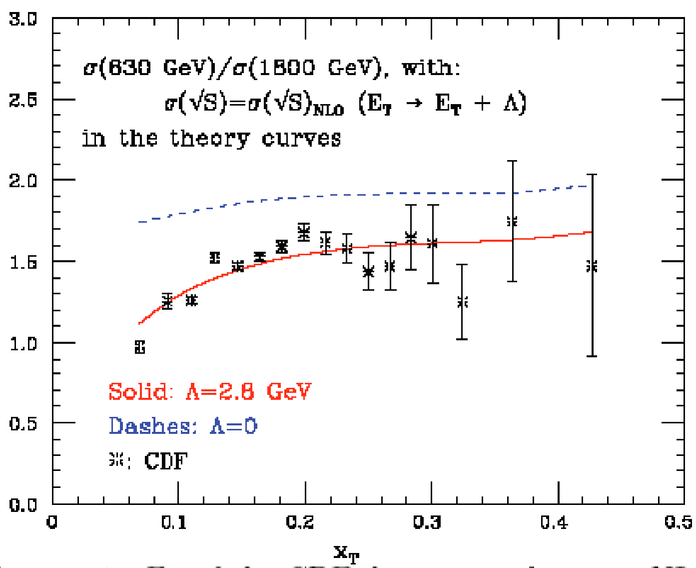
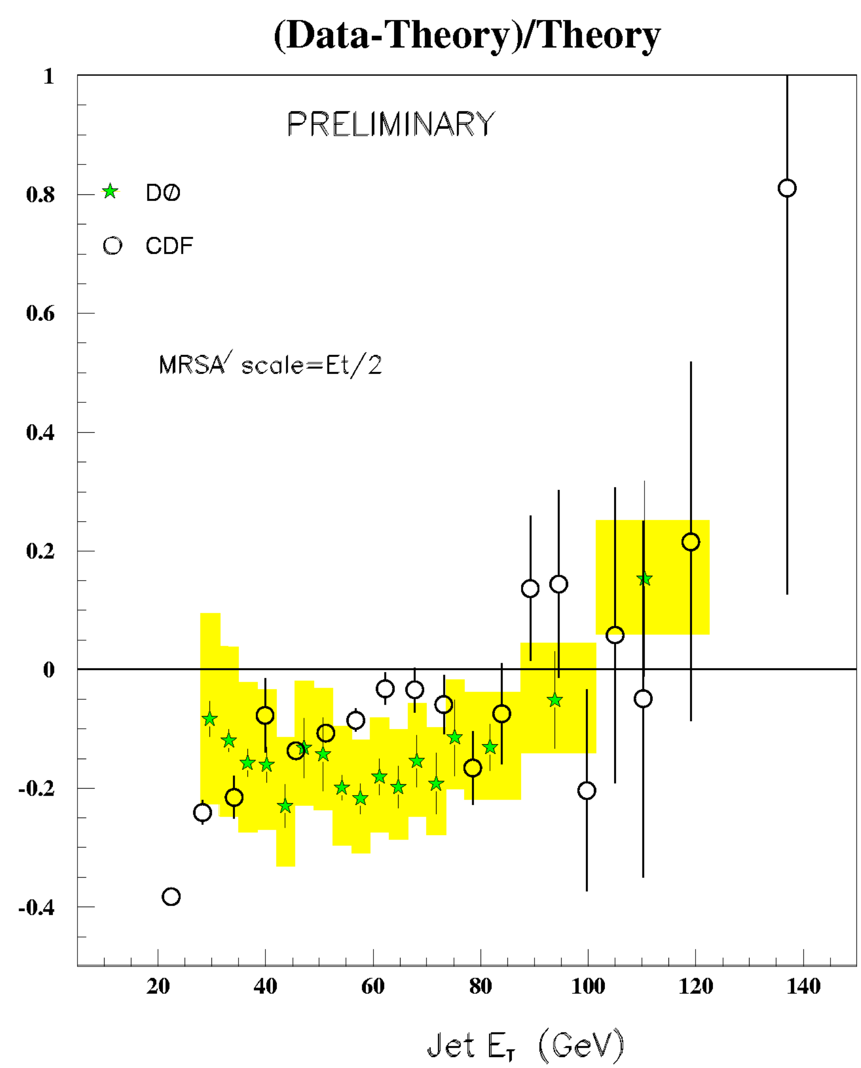


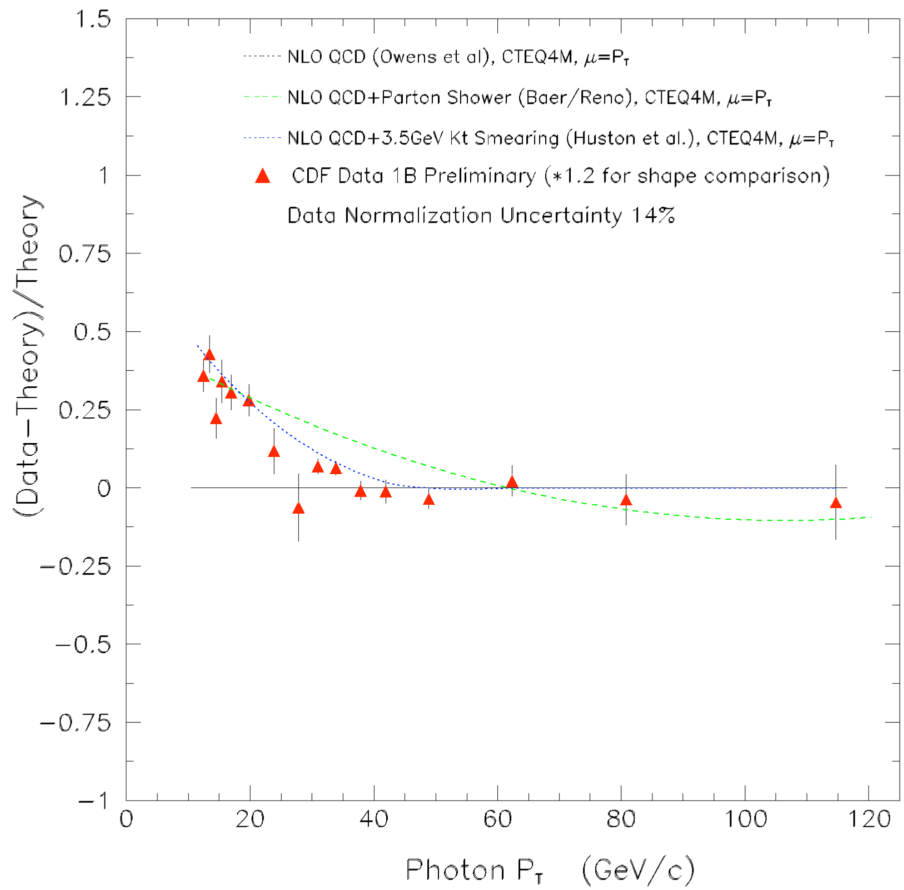
Figure 17. Fit of the CDF data using the exact NLO jet cross-section (CTEQ3M, $\mu = E_T/2$), assuming an E_T-independent shift Λ in the jet energy.

- Other power type corrections (jet algorithms…) for rest?



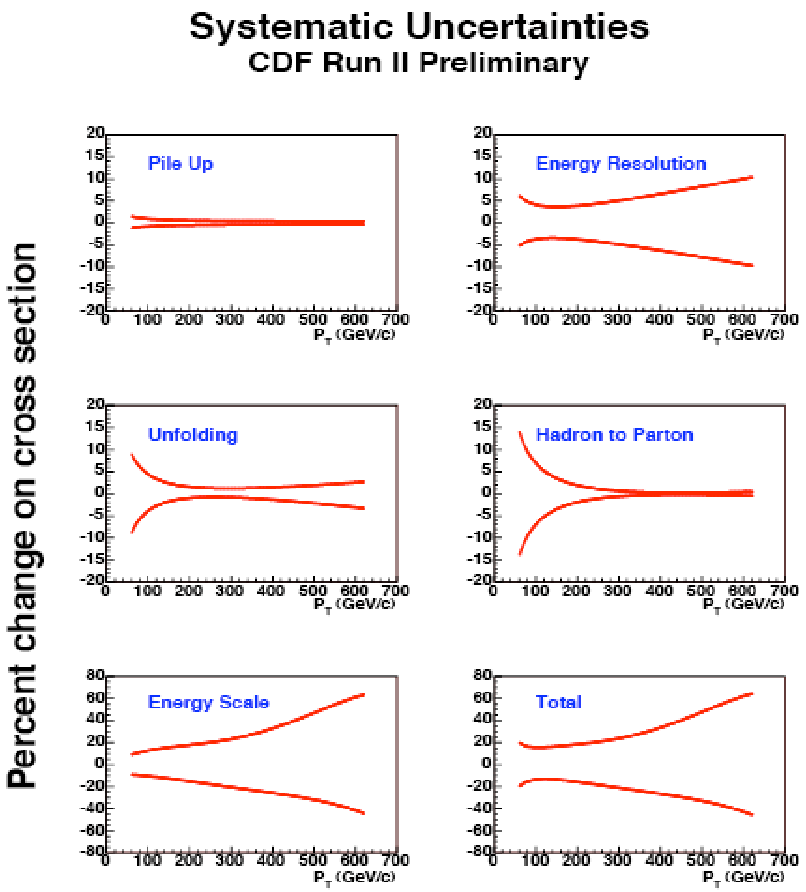
While I'm thinking about it, another power correction

- Soft gluon radiation from initial state partons gives a k_T kick to inclusive object (in this case photon, but could also be jet) that increases the cross section at low p_T from pure NLO pQCD result
- Power correction type behavior: falls off as $1/p_T^2$
- Promising results from joint threshold resummation approach
  - but haven't heard much lately from those guys…



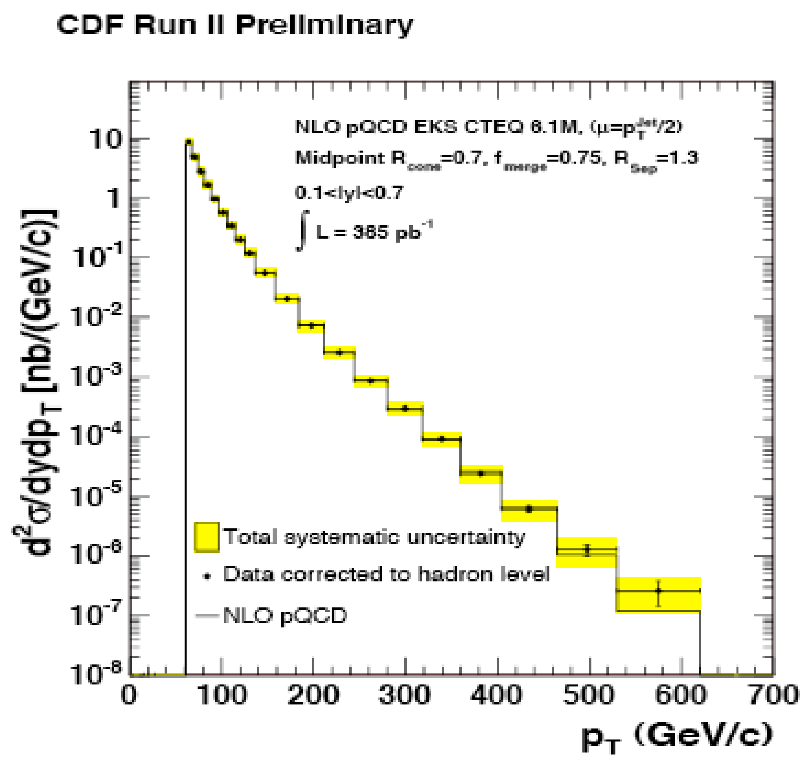
Systematic uncertainties

- Uncertainty on jet energy scale is dominant over most of range
- Hadron-parton contributes to low p_T uncertainty
  - probably an overestimate



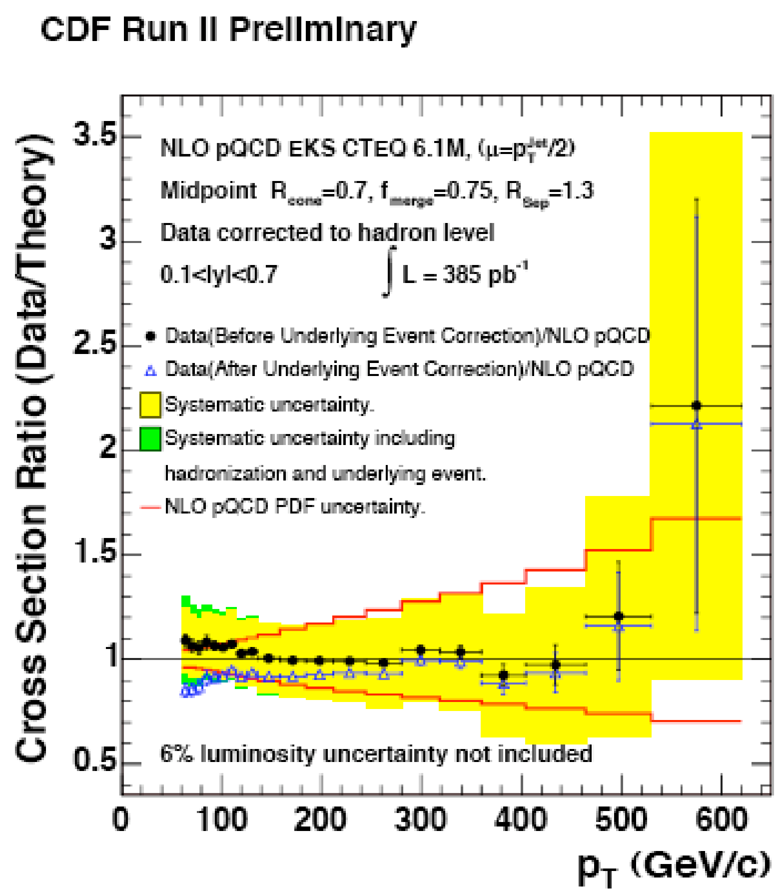
Results

- Results compared to EKS (Ellis-Kunzst-Soper) NLO calculation using an R_{sep} of 1.3 and scale of $p_T^{jet}/2$



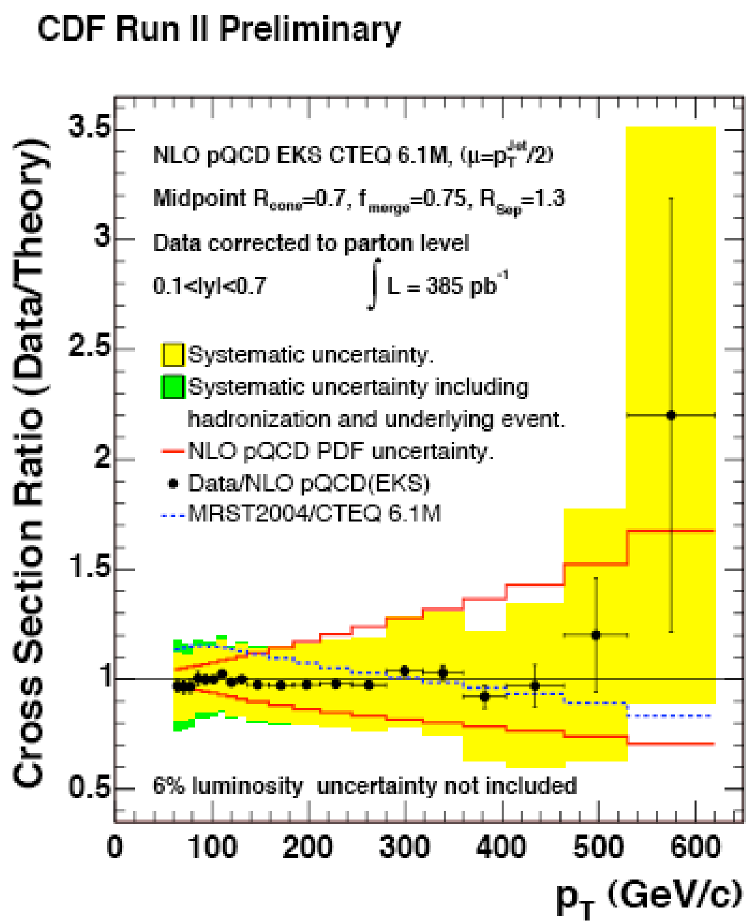
Results

- Solid dots
  - data is corrected to hadron level, i.e. not corrected for UE or fragmentation
- Open triangles
  - after applying the UE correction
  - this is the way in which the Run 1 jet cross sections from CDF and D0 were presented, i.e. with no fragmentation corrections
  - deficit obvious at low p_T



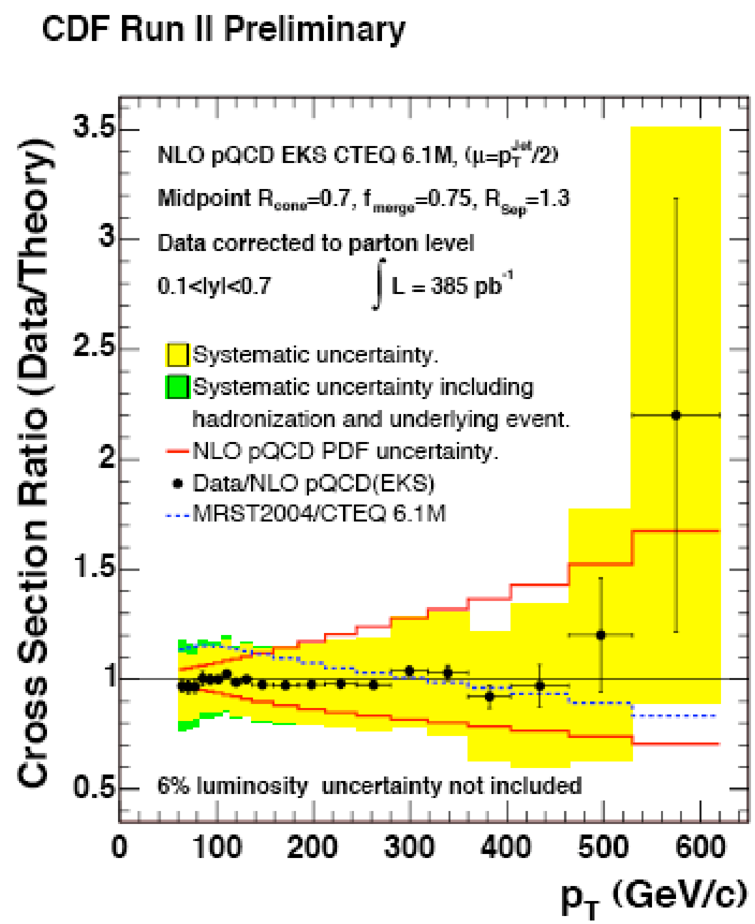
Fully corrected cross section

- Results agree well with NLO predictions using CTEQ6.1 pdf's
- CTEQ6.1 pdf's included jet data from both CDF and D0 in Run 1 and as a result already have an enhanced gluon distribution at high x
- MRST2004 (with its *physical gluon*) agrees better than previous generations of MRST pdf's



Fully corrected cross section

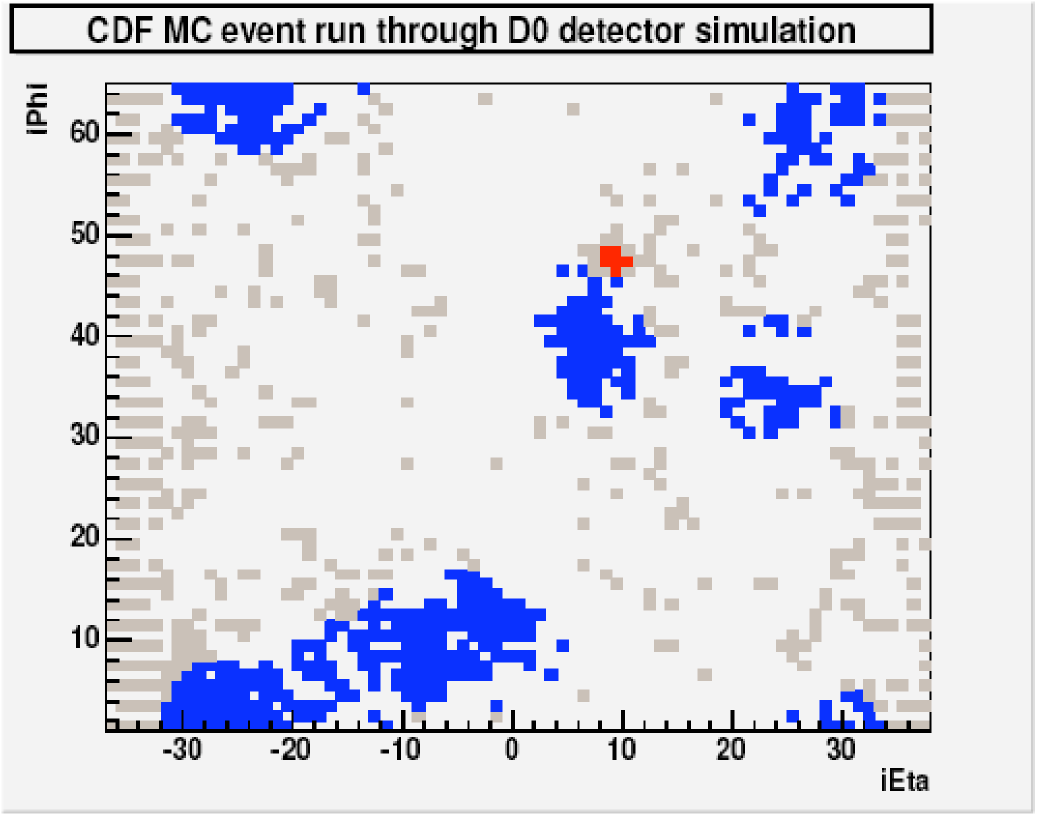
- Results agree well with NLO predictions using CTEQ6.1 pdf's



Return to Dark Towers

Zbenek Hubacek talk at TeV4LHC meeting at CERN

- To address CDF observation of unclustered E_T

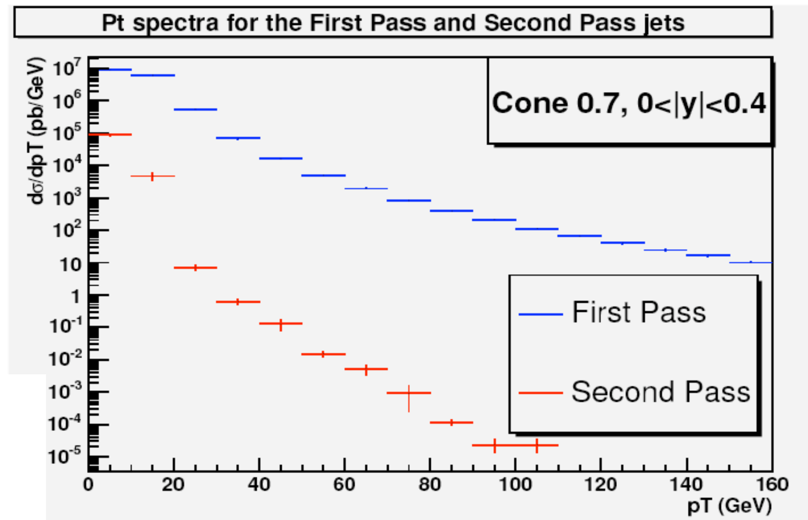


- RunII cone $R = 0.7$
- Jet towers
- Unclustered towers $pT < 2GeV$
- Unclustered towers $pT > 2GeV$

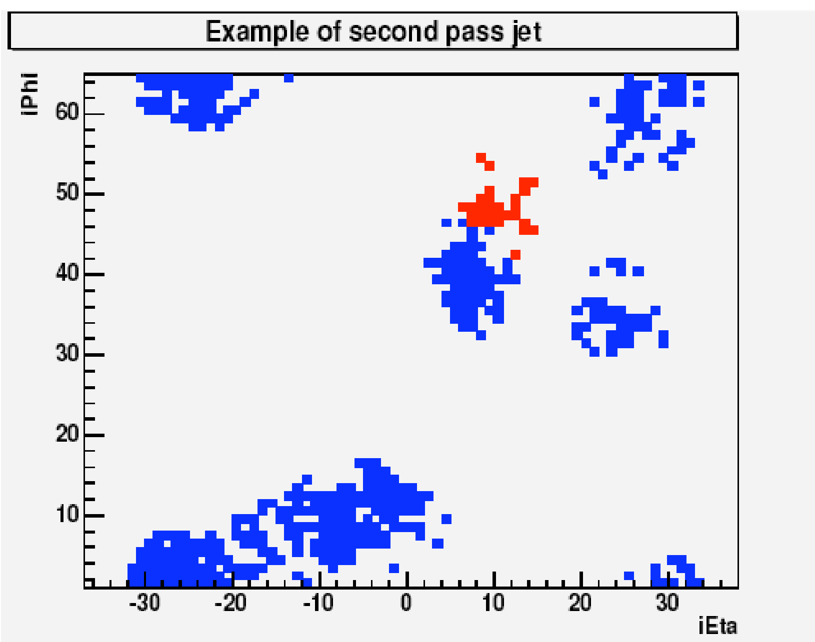
We see it too!

D0 study

- After first iteration of jet-finding algorithm, remove found-jet towers and re-run jet clustering algorithm



Contribution to the cross-section is negligible



The unclustered energy made a second pass jet!

…but if unclustered energy is added to first pass jets (as is done for the modified CDF midpoint algorithm), contribution is not negligible

•NLO theory is agnostic on this point

•MC@NLO (with inclusive jet production) is not

more study needed

Current situation

- CDF is using an initial seach cone of R/2 with their midpoint cone algorithm
- D0 is not
  - ~5% difference in cross section in data (CDF data > D0 data)
- Theory is agnostic on the matter
- CDF is using an R_{sep} parameter of 1.3 with NLO theory
  - because jets separated by >1.3*R_{cone} are not merged in data
- D0 is using an R_{sep} parameter of 2
  - ~5% difference in theory prediction (D0 theory > CDF theory)
- So if CDF and D0 were measuring exactly the same jets, the comparison to theory would differ by 10%!
- We don't want this to happen with ATLAS and CMS

k_T algorithm

- Inclusive k_T algorithm
  - merging pairs of nearby particles in order of increasing relative p_T

$$k_{T,ij} = \min(p_{T,i}^{2}, p_{T,j}^{2})\frac{\Delta R^2}{D^2}$$

$$k_{T,i} = p_{T,i}^{2}$$

  - D parameter controls merging termination and characterizes size of resulting jets
- p_T classification inspired by pQCD gluon emissions
  - infrared and Collinear safe to all orders in pQCD
  - no merging/splitting
    - no R_{SEP} issue comparing to pQCD

– Successfully used at LEP and HERA

– Relatively new in hadron collider

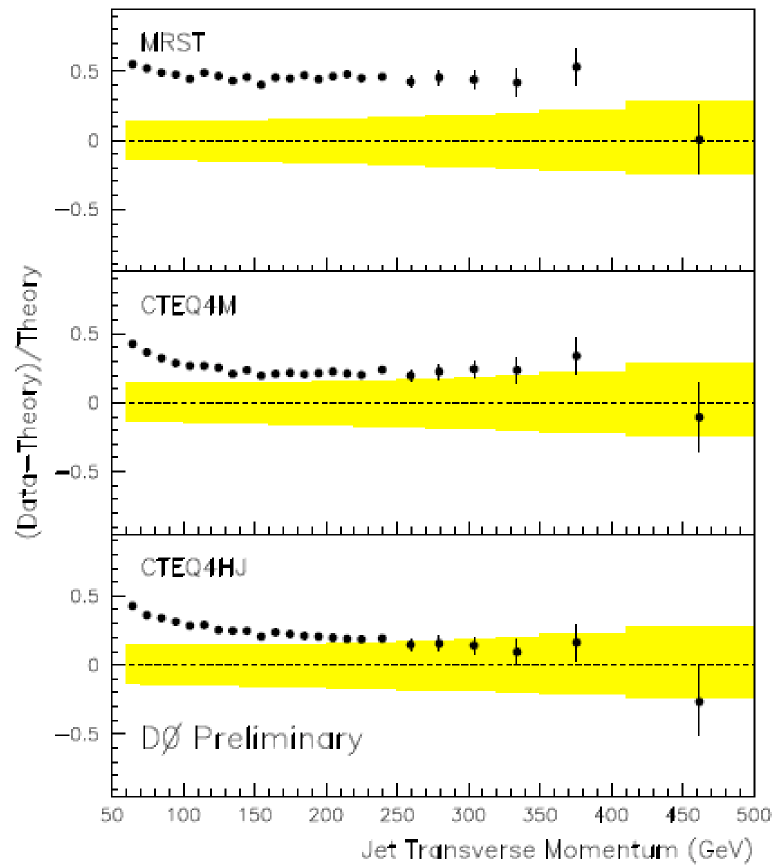
- More difficult environment

→ Underlying Event

→ Multiple Interactions per crossing (MI)

Worry from Run 1

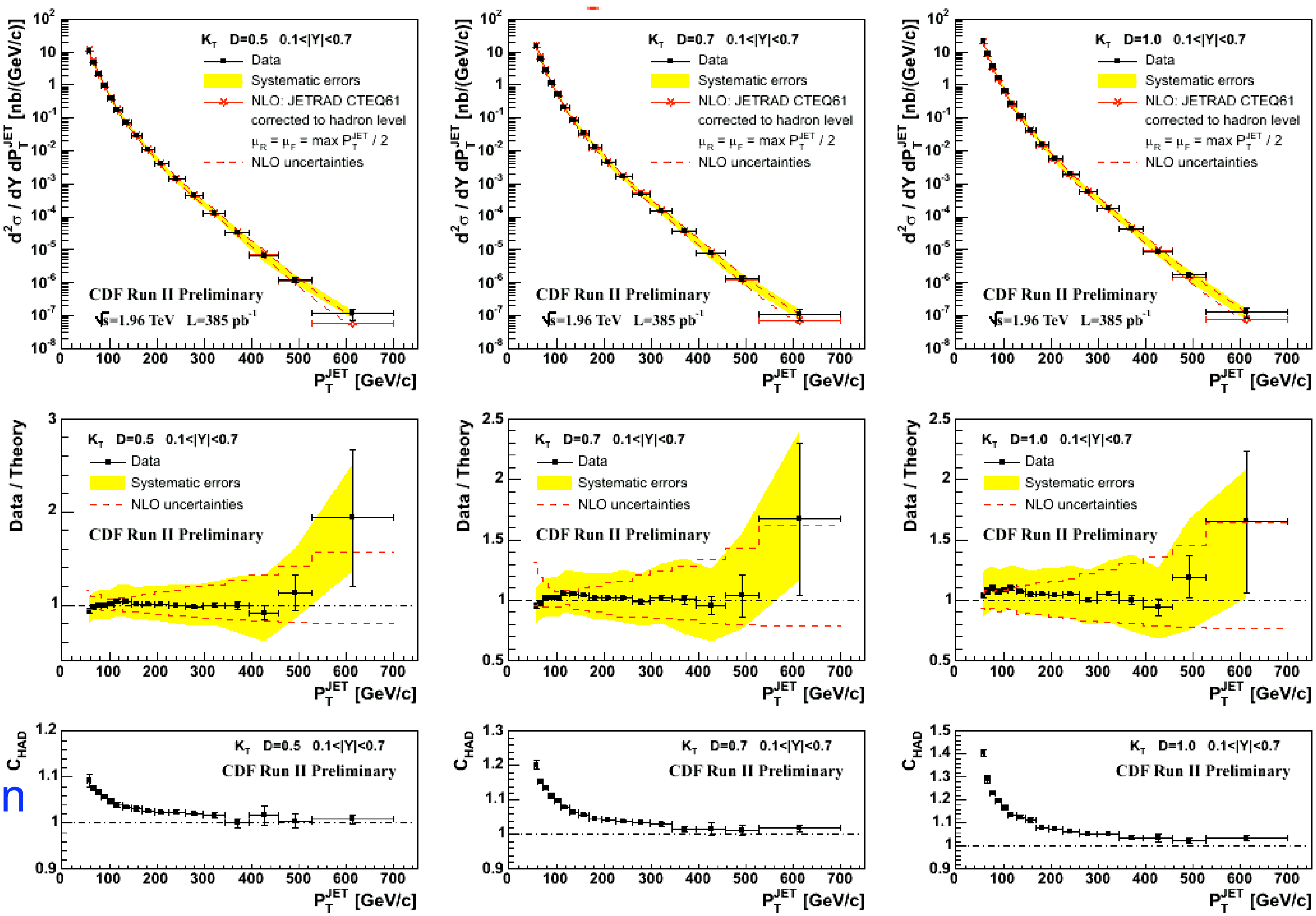
- D0 k_T result did not agree nearly as well with NLO predictions as did the D0 (and CDF) cone result
  - CDF had no k_T result in Run 1
- Does k_T algorithm reach out and grab stuff?
  - hadrons: ok, maybe reduces hadronization correction
  - partons: not so good, may be higher order effect and beyond scope of NLO calculation



CDF: k_T jet cross section results in Run 2

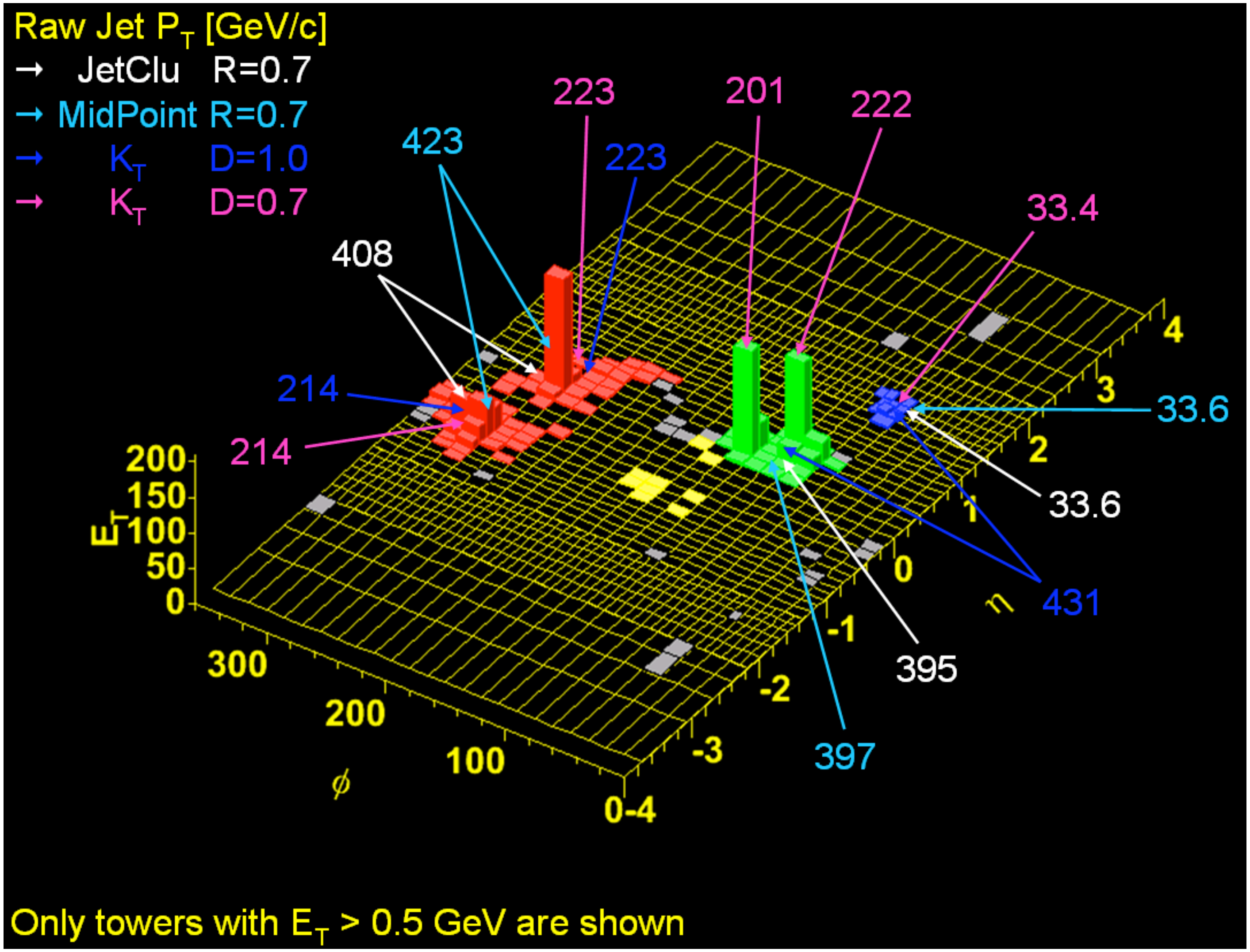
$$d_{ij} = \min(P^2_{T,i}, P^2_{T,j}) \frac{\Delta R^2}{D^2} \qquad d_i = (P_{T,i})^2$$

k_T algorithm seems to work well at a hadron collider



underlying + hadronization correction

Interesting event to study algorithm differences

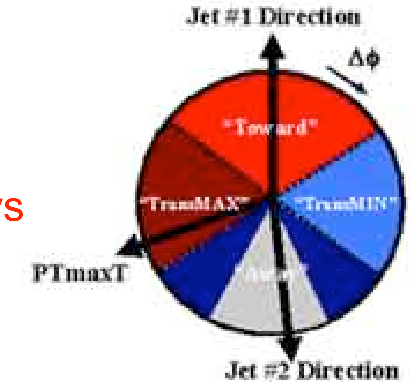


…project to examine what different experimental algorithms (CDF, D0, ATLAS CMS) do with *interesting* events

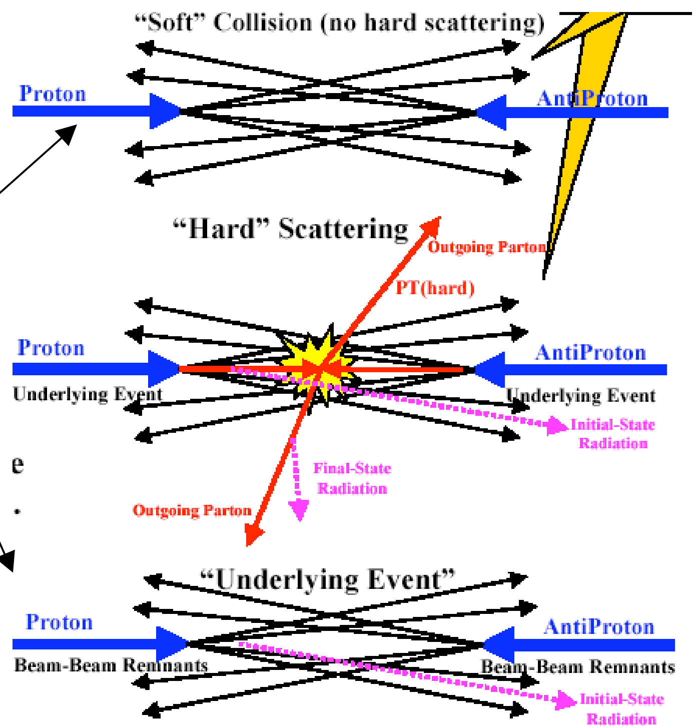
Importance of underlying event

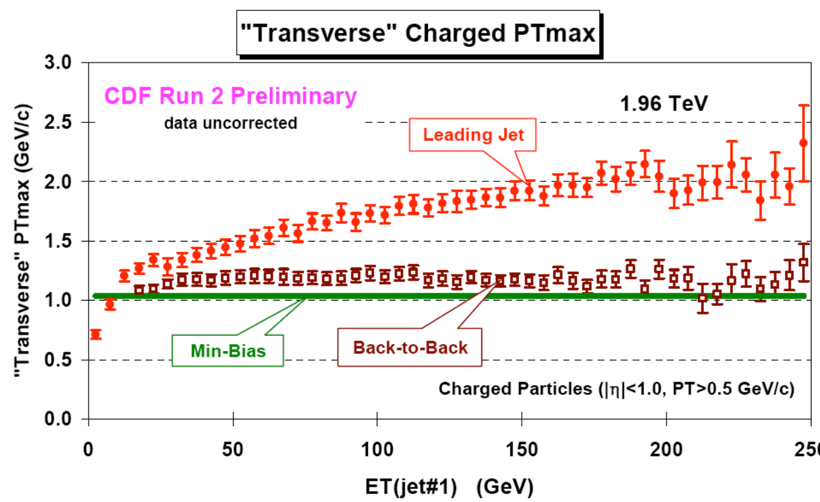
- Have to subtract underlying event from hard scatter in order to compare jet cross sections to parton-level calculations

Σp_T in max region increases as jet E_T increases
Σp_T in min region stays flat, at level similar to min bias



how similar are these two?



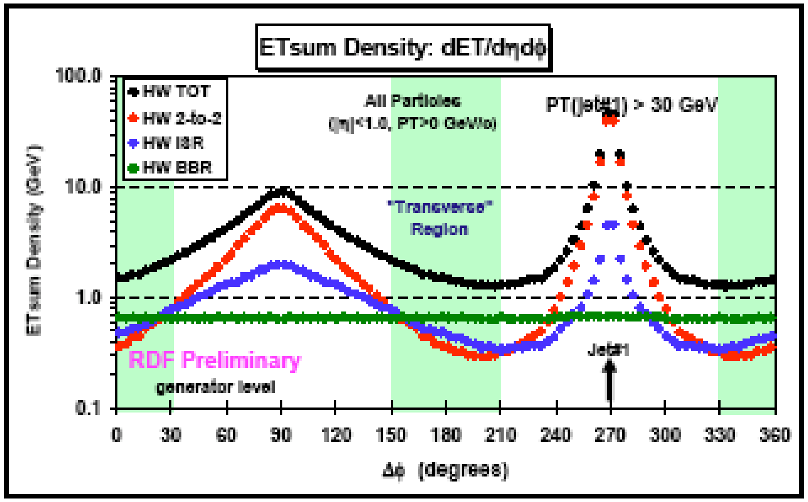


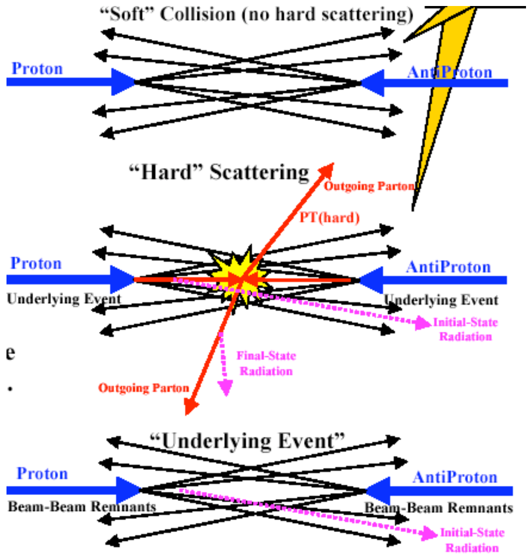
need inclusive jet production in MCatNLO
currently underway, but slowly

Other issues

- Is there an ISR contribution to the jet cross section not accounted for in the NLO calculation?
- And what's with that Tune A, anyway? It seems to represent an extreme
  - the colors of the MI partons in the final state are almost 100% correlated with the ones from the hard scattering
  - MI scatterings are supposed to be perturbatively independent

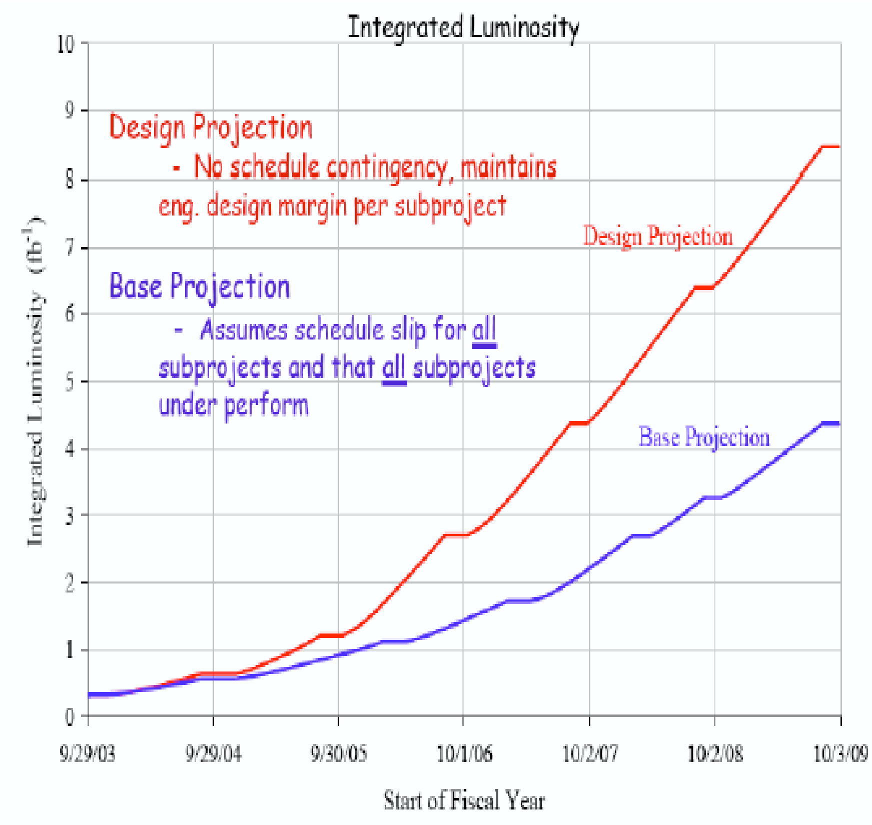
R. Field at TeV4LHC meeting in Dec





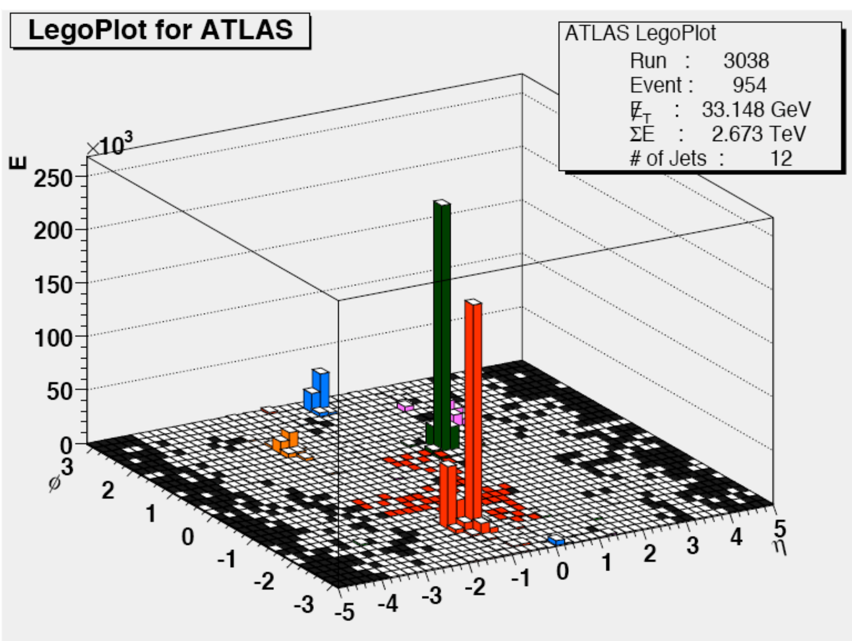
Summary

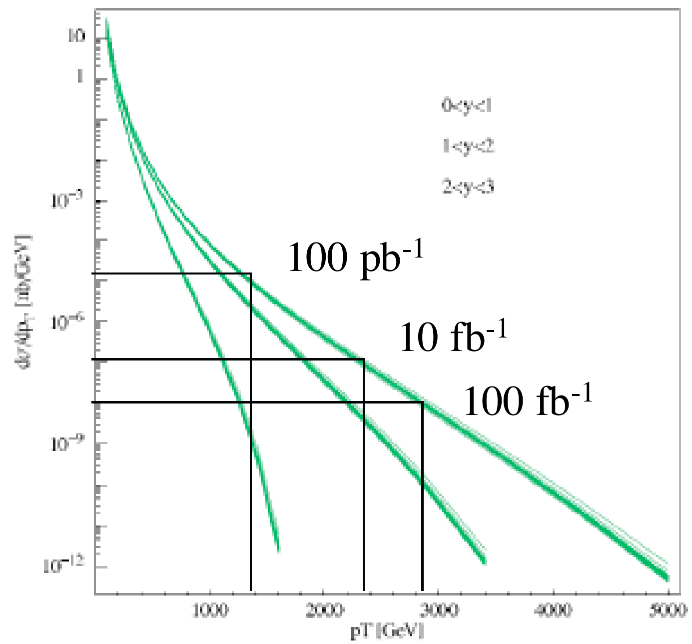
- Tevatron, CDF and D0 all working well
  - will have 1 fb^{-1} jet analyses in near future
    - pushing to forward region
    - and to lower p_T
- ~1.2 fb^{-1} down and > 8 fb^{-1} to go
  - 2 fb^{-1} by 2006
  - 4 fb^{-1} by 2007
  - 8 fb^{-1} by 2008



LHC

- Looking forward to the LHC
- Events will be more active than at Tevatron due to
  - more gg initial states
  - more phase space for gluon emission
  - underlying event more capable of producing jets



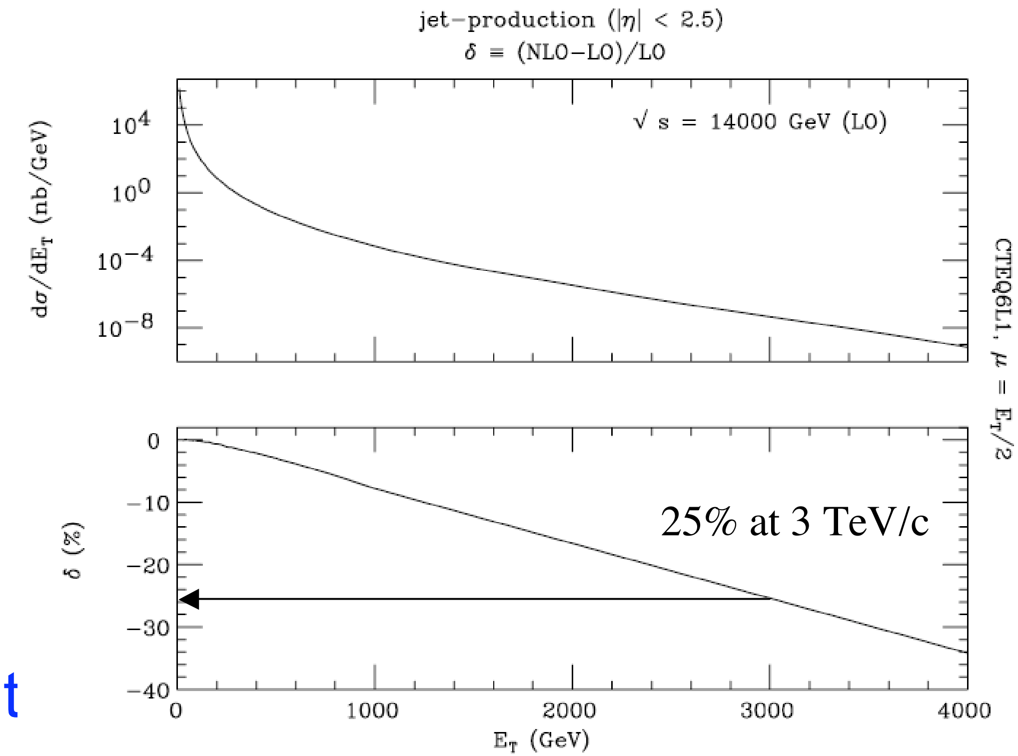


-meeting in Glasgow in early July to discuss jet physics at the LHC
-would like common algorithms for ATLAS, CMS
-would like both cone and k_T in common use

Extras

Unexpected new SM physics

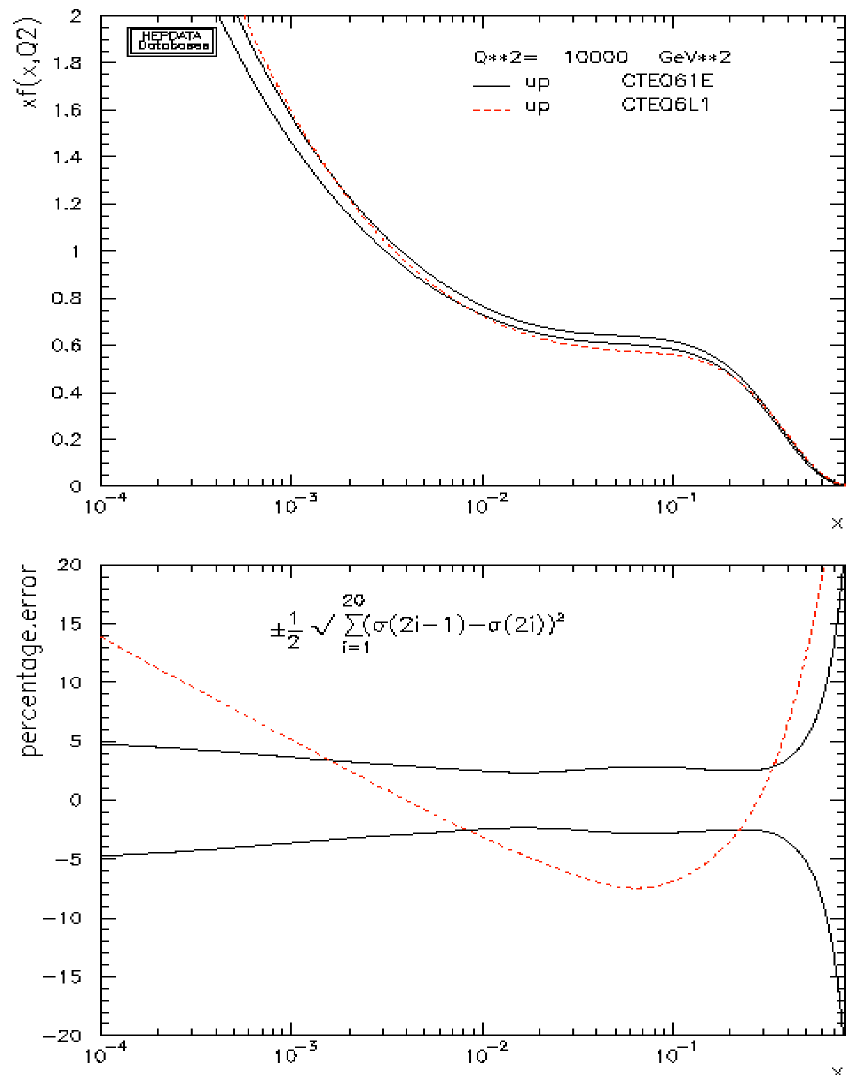
- In a recent paper (hep-ph/0503152), Stefano Moretti and Douglas Ross have shown large 1-loop weak corrections to the inclusive jet cross section at the LHC
- Effect goes as $\alpha_W \log^2(E_T^2/M_Z^2)$
- Confirmation is important
- Other (unsuspected) areas where weak corrections are important?



In *Rumsfeldese*, this is now one of the "known unknowns".
What are our unknown unknowns?

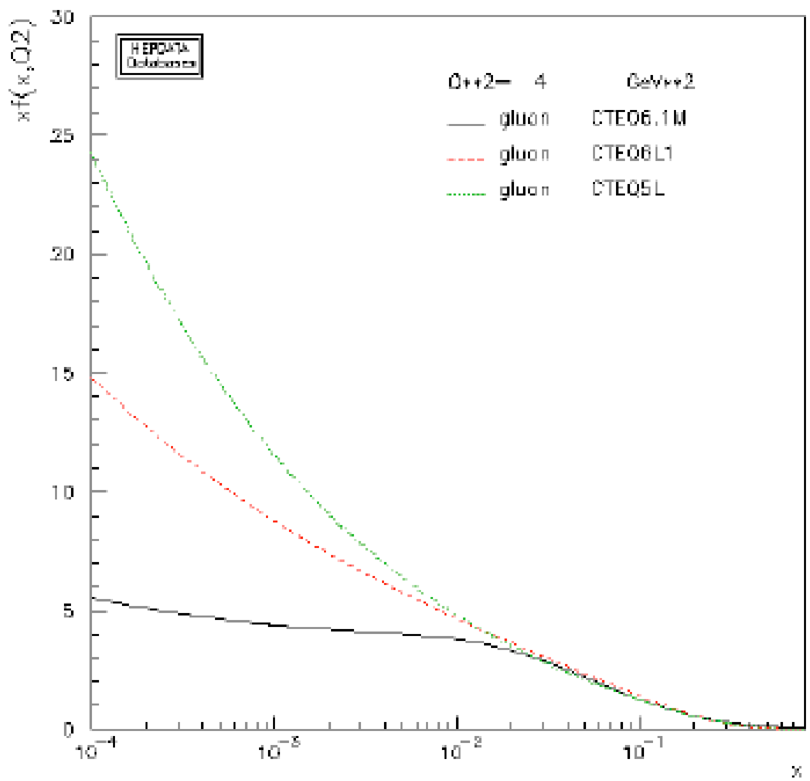
LO vs NLO pdf's for parton shower MC's

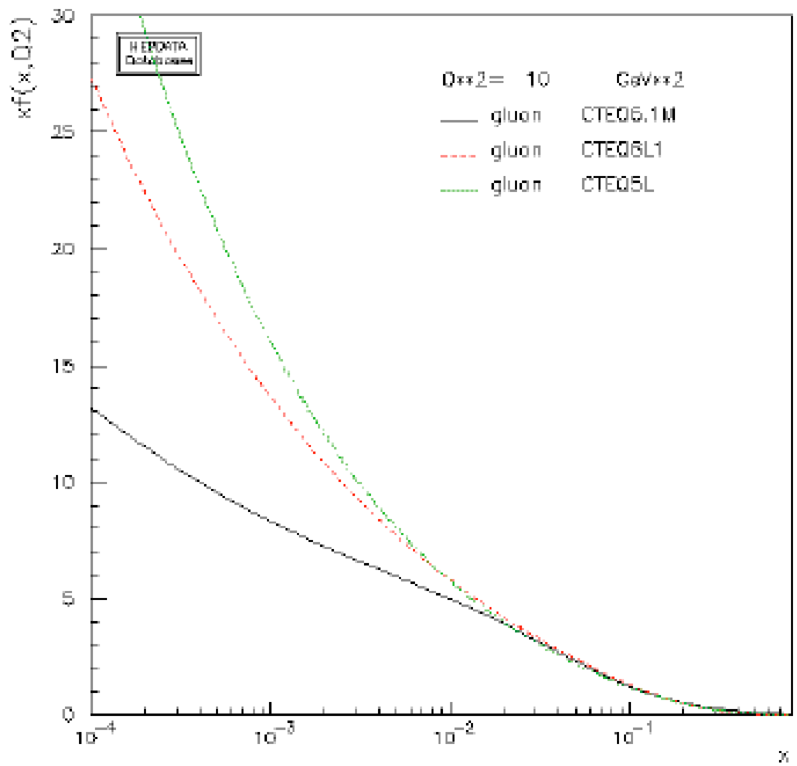
- For NLO calculations, use NLO pdf's (duh)
- What about for parton shower Monte Carlos?
  - somewhat arbitrary assumptions (for example fixing Drell-Yan normalization) have to be made in LO pdf fits
  - DIS data in global fits affect LO pdf's in ways that may not directly transfer to LO hadron collider predictions
  - LO pdf's for the most part are outside the NLO pdf error band
  - LO matrix elements for many of the processes that we want to calculate are not so different from NLO matrix elements
  - by adding parton showers, we are partway towards NLO anyway
  - any error is formally of NLO
- (my recommendation) use NLO pdf's
  - pdf's must be + definite in regions of application (CTEQ is so by def'n)
- Note that this has implications for MC tuning, i.e. Tune A uses CTEQ5L
  - need tunes for NLO pdf's



...but at the end of the day this is still LO physics;
There's no substitute for honest-to-god NLO.

- 5L significantly steeper at low x and Q^2
- Rick Field has produced a tune based on CTEQ6.1





CTEQ6.1 Tune

I used LHAPDF! See the next talk by Craig Group!

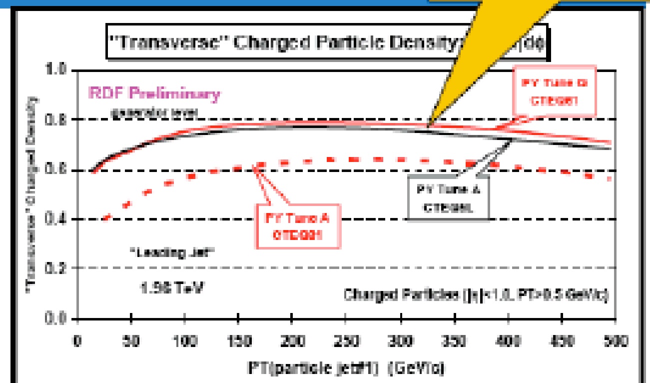
PYTHIA 6.2 CTEQ6.1

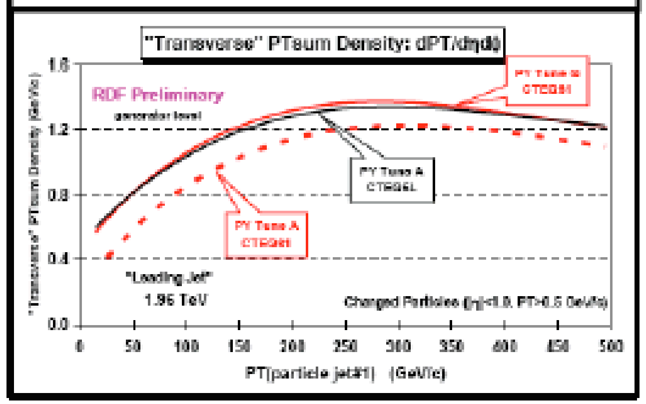
Parameter	Tune Q	Tune QW
MSTP(81)	1	1
MSTP(82)	4	4
PARP(82)	1.2 GeV	1.2 GeV
PARP(83)	0.5	0.5
PARP(84)	0.4	0.4
PARP(85)	0.9	0.9
PARP(86)	0.95	0.95
PARP(89)	1.8 TeV	1.8 TeV
PARP(90)	0.25	0.25
PARP(62)	1.0	1.25
PARP(64)	1.0	0.2
PARP(67)	4.0	4.0
MSTP(91)	1	1
PARP(91)	1.0	2.1
PARP(92)	5.0	15.0

UE Parameters

ISR Parameters

Intrensic KT





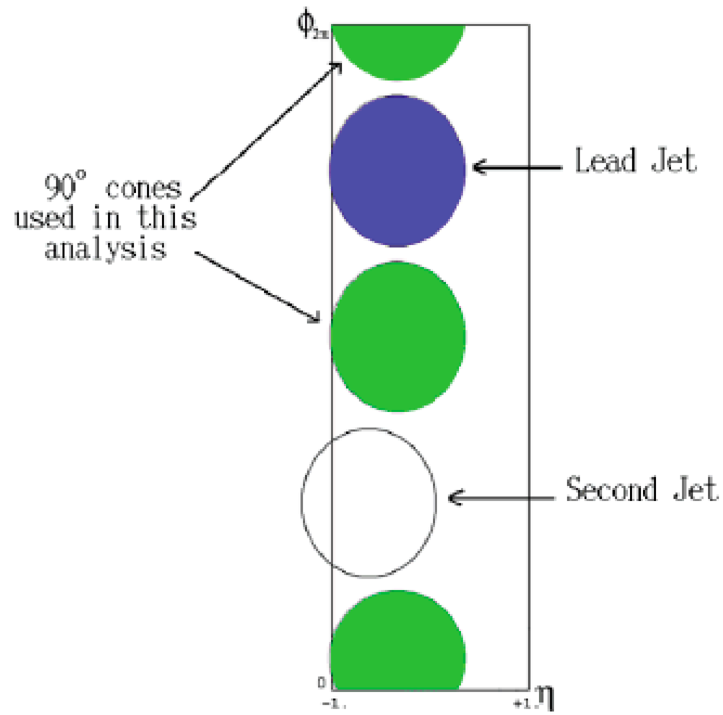
TeV4LHC - Fermilab
October 20, 2005

Rick Field - Florida/CDF

Page 18

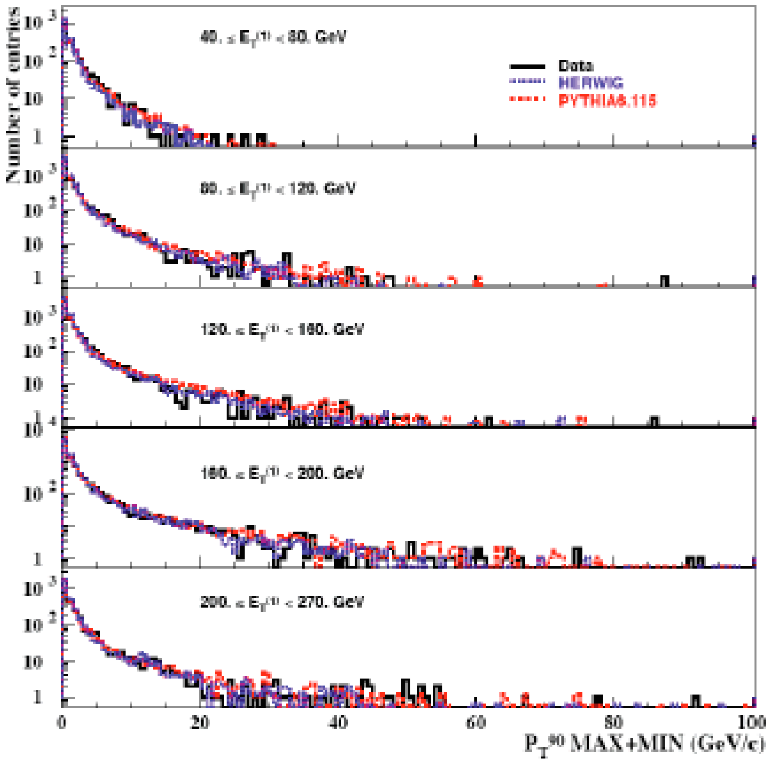
Run 1 study

- Inclusive jet cross sections at CDF are compared to NLO QCD calculations at the parton level
- The (mostly) non-perturbative underlying event has to be subtracted in order for the comparison to be made
- The assumption made by CDF is that the underlying event measured in an *active* (class 12 vertex) should be subtracted
- This analysis sought to check this assumption and to understand how well Monte Carlo programs predicted the event structure
- In this analysis, we used the Run I inclusive cross section data but restricted there to be only 1 vertex of class 10,11 or 12 in the event



By definition there is at least one jet in the central rapidity region; we construct 2 cones (R=0.7) at the same rapidity as the lead jet and 90 degrees away in phi. One of these cones has more energy (max cone) and one has less (min cone)

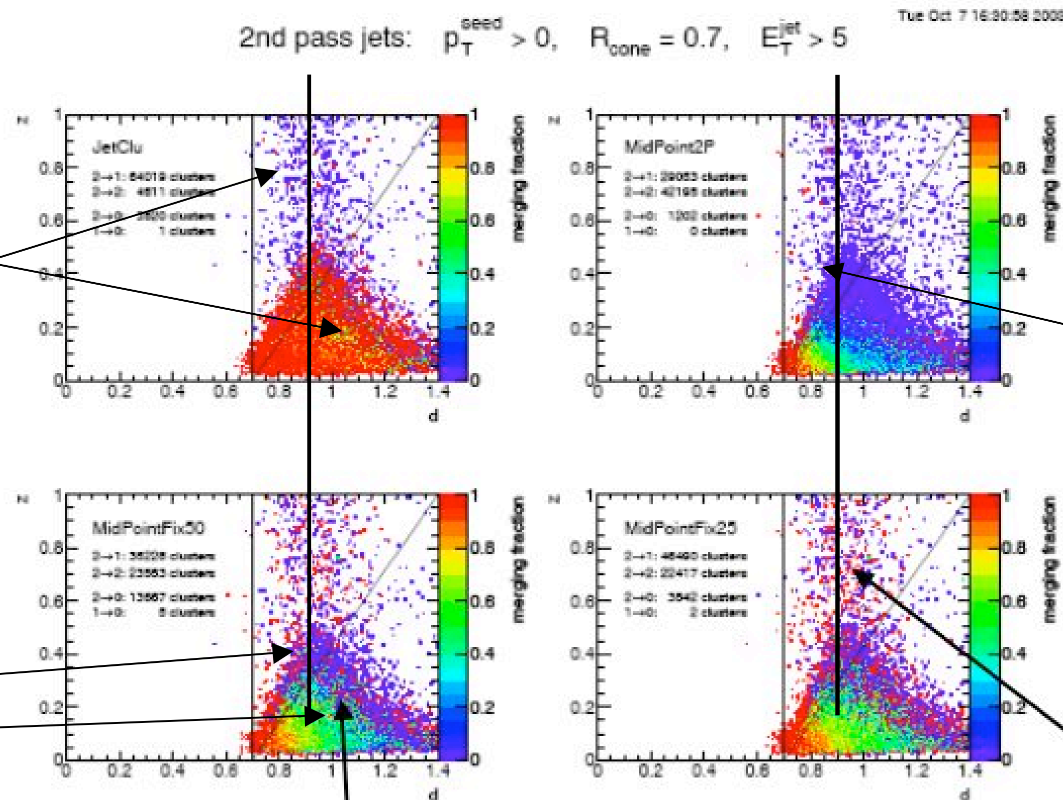
- Look at distributions of momentum in max+min cones for different lead jet bins
- Contains contributions from underlying event plus gluon radiation
  - double-log enhanced
    - basis of parton shower Monte Carlos
  - single-log enhanced
    - partially in MC's; new area of much theoretical effort
    - expect major contribution when log (E_T^{jet}/E_T^{cones}) is large
- Conclusion
  - Pythia and Herwig ain't so bad



R_{sep}

With the introduction of an R_{sep} parameter of 1.3 into the NLO calculation, an ideal cone algorithm would merge any jets above the diagonal and to the left of the line.

z (E_T^{jet2}/E_T^{jet1}) vs d (ΔR between jets)



JetClu merges lots of jets down here due to racheting and misses some here.

midpoint with no initial smaller search cone misses some jets here

Midpoint with a smaller initial search cone merges more jets here but also here.

An ideal cone jet algorithm would merge all nearby jets in this region and none in this region.