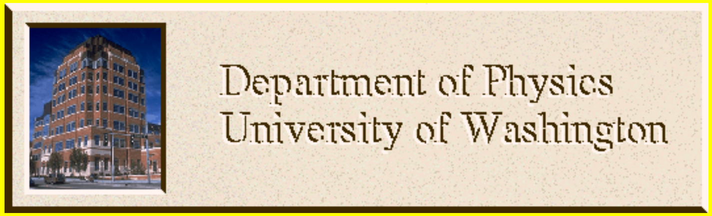
View From a(n Old!) Theorist

Stephen D. Ellis
University of Washington

A lot of Philosophy From Long Experience
– See the Next talk for Serious Details



MC4LHC CERN July 2006

The Goal at the LHC is a 1% (Precision) Description of Strong Interaction Physics (where Tevatron Run I is ~ 10%)

To this end we want to *precisely* map

- physics at 1 meter, *i.e.*, what we can measure in the detector, *e.g.*, $E(y,\phi)$

On To

- physics ⑥ 1 fermi, *i.e.*, what we can calculate with small numbers of partons, leptons and gauge bosons as functions of E, y, ϕ

We “understand” what happens at the level of short distance partons and leptons, *i.e.*, perturbation theory is simple, can reconstruct masses, *etc*.

Thus

We want to map the observed (hadronic) final states onto a representation that mimics the kinematics of the (short-distance) partons; ideally on a event-by-event basis.

But

We know that the (short-distance) partons shower (perturbatively) and hadronize (nonperturbatively), *i.e*., spread out as they evolve from short to long distances, and there *must be* color correlations.

"<u>SOLUTION</u>": associate "nearby" hadrons or partons into JETS via ALGORITHMS, *i.e.*, rules that can be applied to data and theory

- <u>Cone Algorithms</u>, *e.g.*, Snowmass, based on "fixed" geometry (well suited to hadron colliders with UEs)

- <u>k_T Algorithm,</u> based on pairwise merging, nearest, lowest p_T first (familiar at e^+e^- colliders), tends to "vacuum up" soft particles

👍 **Render PertThy IR & Collinear Safe**

👎 **But mapping of hadrons to partons can *never* be 1 to 1, event-by-event!   Colored states ≠ singlet states!**

Goals of IDEAL ALGORITHM (Motherhood)

- *Fully Specified: including defining in detail any preclustering, merging, and splitting issues*
- *Theoretically Well Behaved: the algorithm should be infrared and collinear safe (and insensitive) with no ad hoc clustering parameters (e.g., R_{SEP})*
- *Detector Independence: there should be no dependence on cell type, numbers, or size*
- *Order Independence: The algorithms should behave equally at the parton, particle, and detector levels.*
- *Uniformity: everyone uses the same algorithms*

Defining a Jet with Algorithm-

- Start with a list of particles (4-vectors) and/or calorimeter towers (energies and angles)

- End with lists of particles/towers, one list for each jet

- And a list of particles/towers not in any jet – the spectators – remnants of the initial hadrons not involved in the short distance physics (but there must be some correlations)

<u>Fundamental Issue</u> – Compare Experiments to each other & to Theory

Warning:

<u>We should all use the same algorithm!!</u>
(as closely as humanly possible), *i.e*. both ATLAS & CMS (and theorists).

This is NOT the case at the Tevatron, even in Run II!!

Observations:

- Iterative Cone Algorithm

  Has detailed issues (merge/split, seeds, dark towers), which only became clear with serious study (and this is a *good* thing)

  And now we know (most of) the issues and can correct for them

- The k_T Algorithm

  May have detailed issues ("vacuum" effect, UE and pile-up sensitivity,..), but much less mature experience at hadron colliders

  We need to find out with the same sort of serious study (history says issues *will* arise)

Cones: Seeds and Sensibility -

- Tension between desire

  To Limit analysis time (for experiments) with seeds

  To Use identical algorithms in data and perturbation theory

- Seeds are intrinsically IR sensitive (MidPoint Fix only for NNLO, not NNNLO)

$\Rightarrow$ DON'T use seeds in perturbation theory, correct for them in data analysis

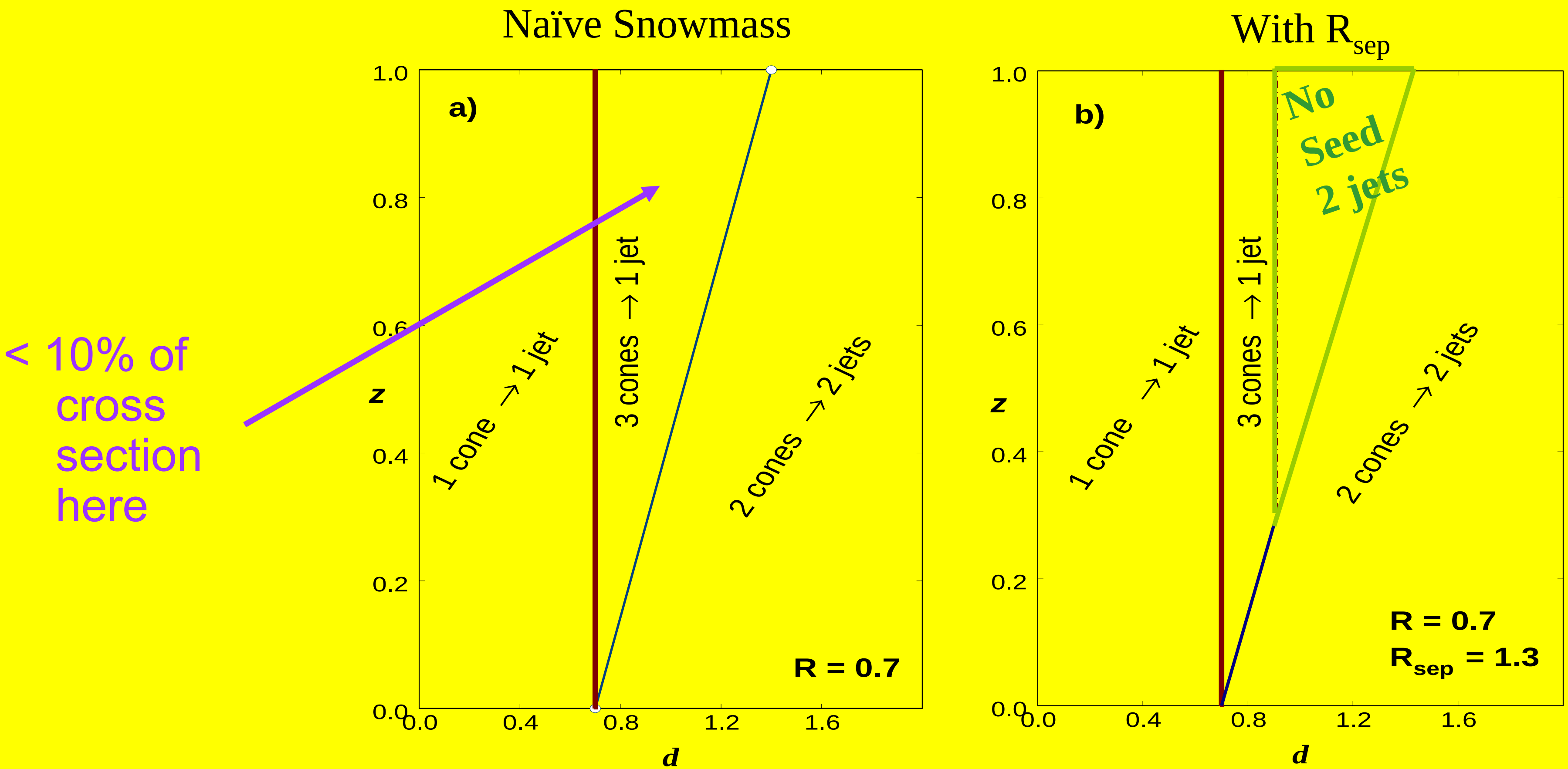
In the theory they are a big deal – IR UNsafety (Yikes)!!!!!!

In the data seeds vs seedless is a few % correction (e.g., lower the Seed p_T threshold) and this is small compared to other corrections – [Run I jets results are meaningful!!]

Numerical issue:

- Seeds can mean missed configurations with 2 partons in 1 Jet, NLO Perturbation Theory – d = parton separation, $z = p_2/p_1$,,

  Simulate the missed middle cones with R_{sep}



To understand this last issue consider Snowmass "Potential"

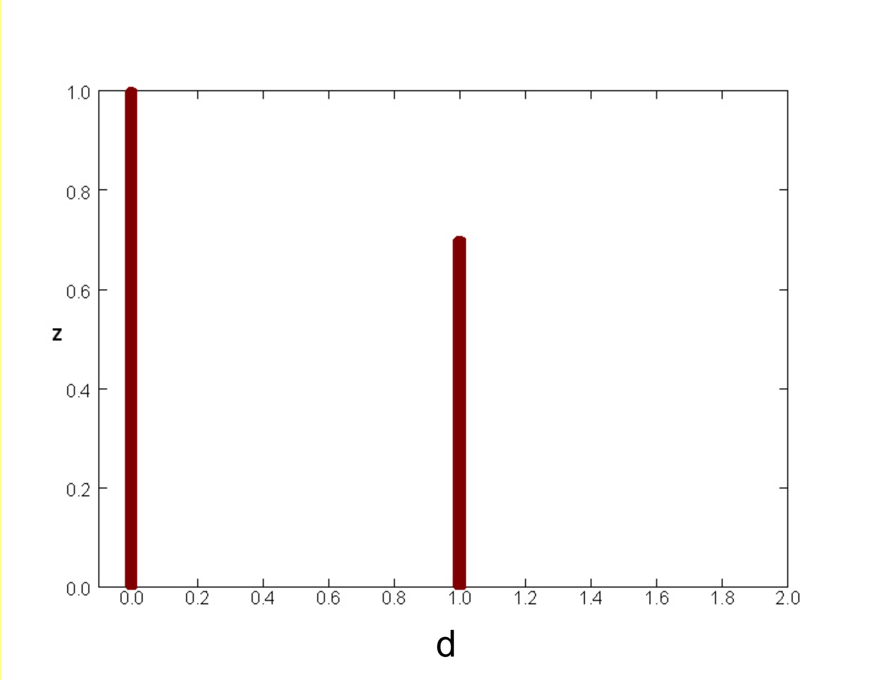
- In terms of 2-D vector $\vec{r} = (\eta, \varphi)$ or (y, φ) define a "potential"

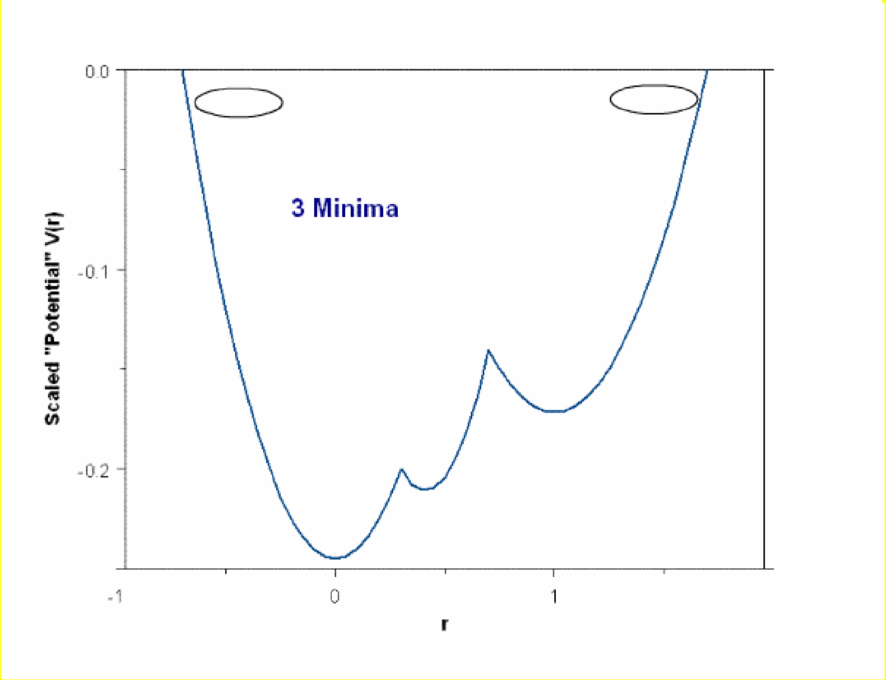
$$V(\vec{r}) \equiv -\frac{1}{2}\sum_i E_T^i \left(R^2 - \left(\vec{r}^{\,i} - \vec{r}\right)^2\right)\Theta\left(R^2 - \left(\vec{r}^{\,i} - \vec{r}\right)^2\right)$$

- Extrema are the positions of the stable cones; gradient is "force" that pushes trial cone to the stable cone, *i.e.,* the flow vector

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}) = \sum_i E_T^i \left(\vec{r}^{\,i} - \vec{r}\right)\Theta\left(R^2 - \left(\vec{r}^{\,i} - \vec{r}\right)^2\right)$$

(THE) Simple Theory Model - 2 partons (separated by < 2R): yield potential with 3 minima – trial cones will migrate to minima from seeds near original partons ⇒ miss central minimum





$z = p_{\min} / p_{\max}$, d = separation

Smearing of order R

A NEW issue for Iterative Cone Algorithms – DARK TOWERS

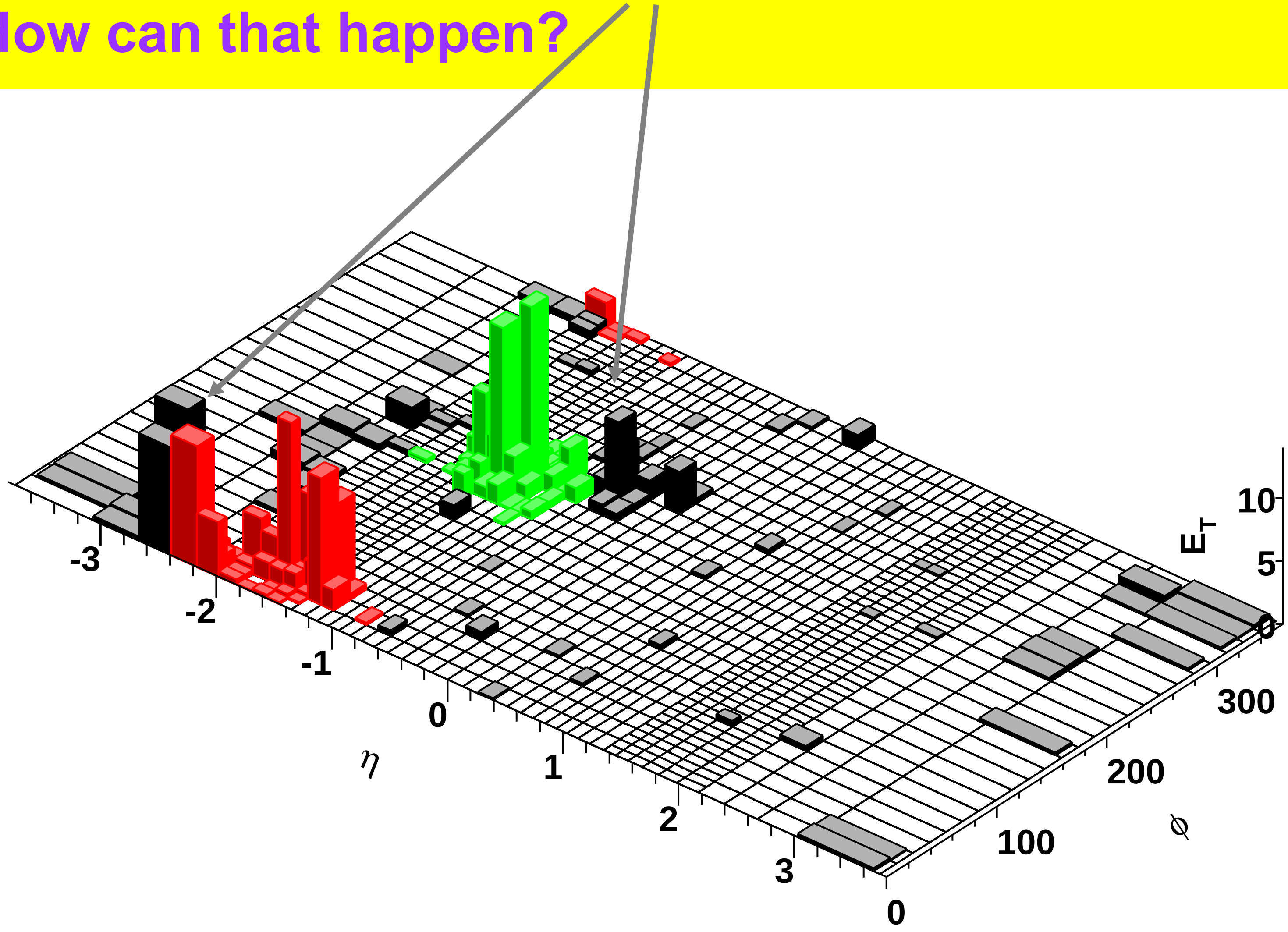
- Compare jets found by JETCLU (with ratcheting) to those found by MidPoint and Seedless Algorithms

- "Missed Energy" – when energy is smeared by showering/hadronization do not always find stable cones expected from perturbation theory

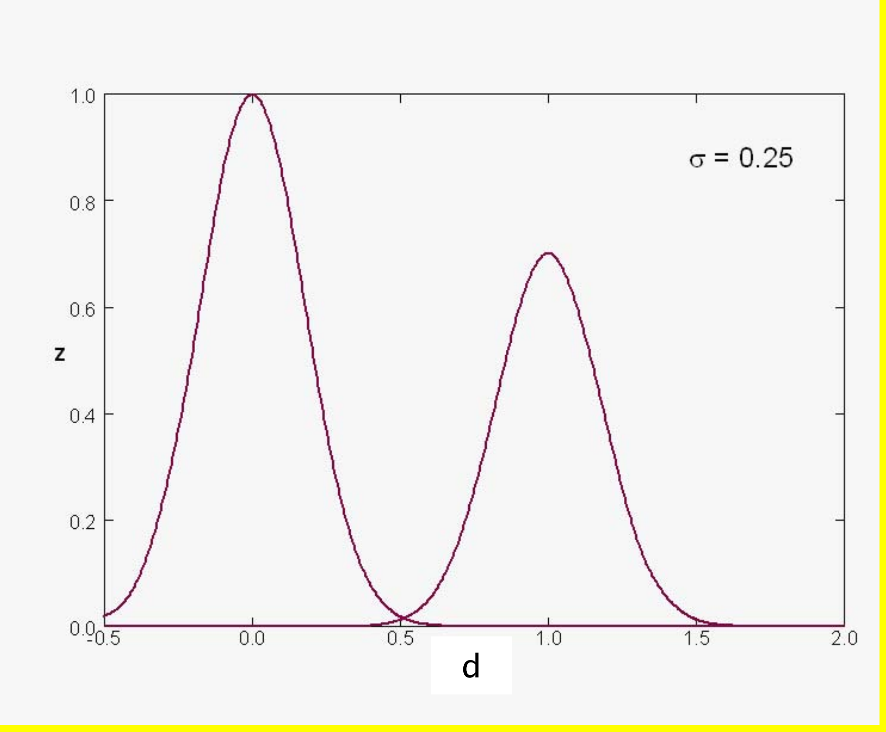
  $\Rightarrow$ 2 partons in 1 cone solutions
  $\Rightarrow$ or even second cone

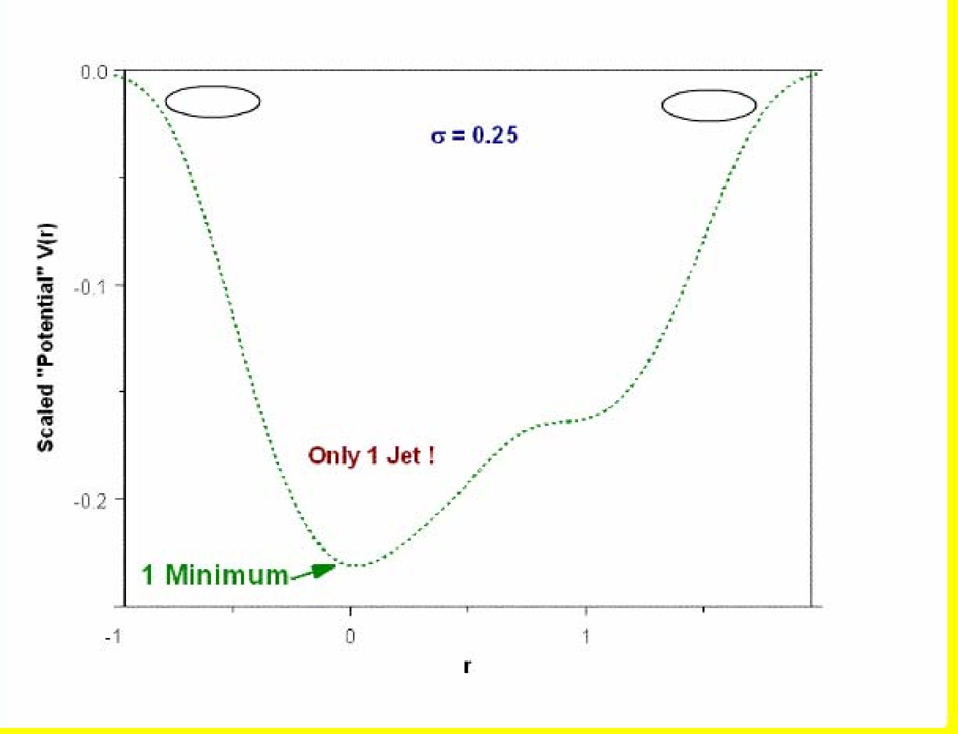
Under-estimate E_T – new kind of Splashout

Missed or Dark Towers (not in any stable cone) – How can that happen?



Why Dark towers? Include smearing (~ showering & hadronization) in simple picture, find only 1 stable cone (no midpoint stable cone & dark towers)



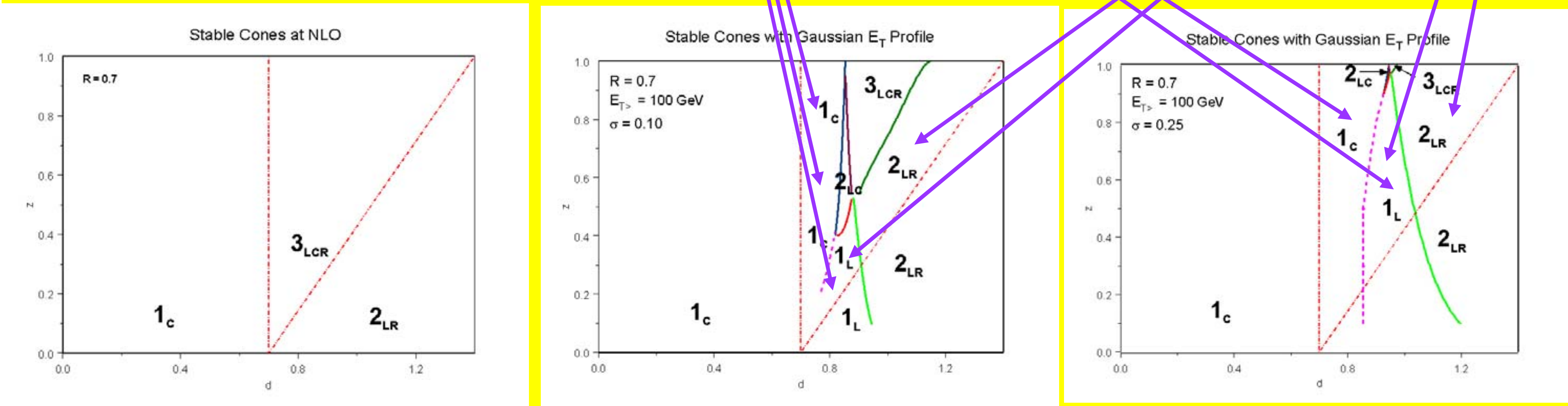


Compare with smearing: MidPoint will still miss 2-in-1 Jets (R_{sep} < 2)

Dark towers (no R stable cone)

Missing MidPoint (no C stable cone)

σ = 0 σ = 0.1 σ = 0.25



Proposed Fix with smaller Radius Search Cone – Used by CDF

- Over compensates with (too) many found stable cones, so use larger f_merge (f_CDF > f_D0)

- (Re)Introduces IR-sensitivity through soft stable search cones (R' < R) that, when expanded to R, can envelop and merge nearby pairs of energetic partons, which themselves do not correspond to a stable cone (R)

- NOT A COMPLETE SOLUTION!!

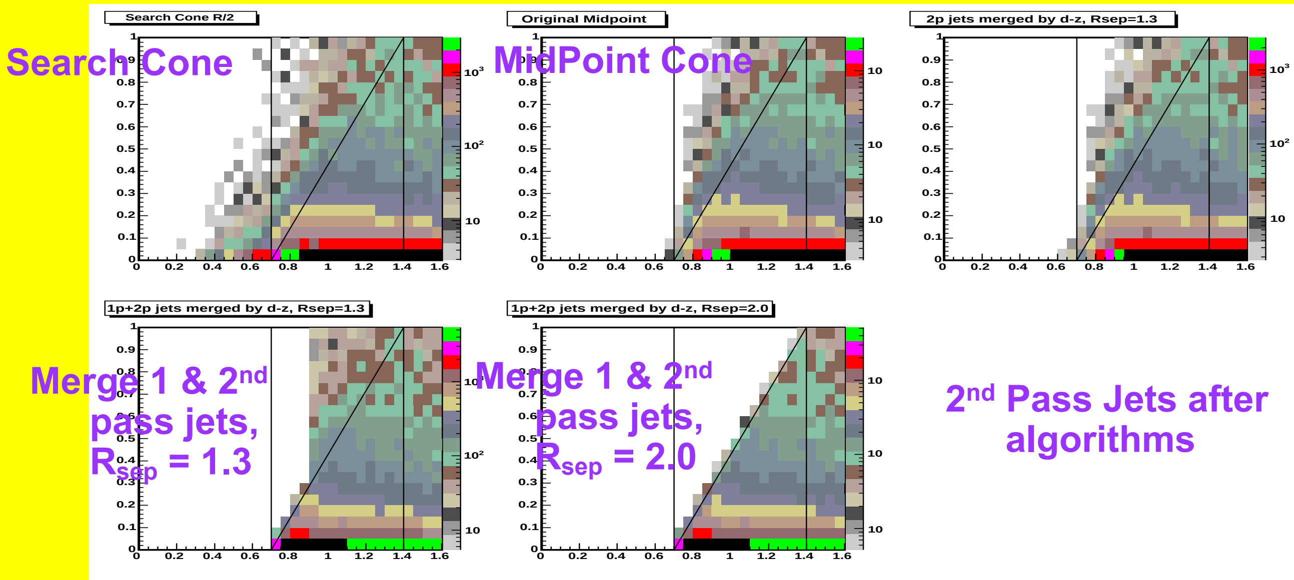
Better(?) - Consider a Dark Tower Correction based on Comparison to pQCD

- Take multiple passes at data

  1st pass jets = found by Cone Algorithm
  2nd pass jets = missed by Cone Algorithm (but found if remove 1st pass jet)

- Merge if in correct region of (d, z) plane

  ⇒ Correction to data



The k_T Algorithm

- Merge partons, particles or towers pair-wise based on "closeness" defined by minimum value of

$$d_{ij}^2 \equiv \mathrm{Min}\left[p_{T,i}^2, p_{T,j}^2\right]\frac{\left(y_i - y_j\right)^2 + \left(\phi_i - \phi_j\right)^2}{D^2}, d_i^2 = p_{T,i}^2$$

  If d_{ij}^2 is the minimum, merge pair and redo list;
  If d_i^2 is the minimum -> *i* is a jet! (no more merging for *i*),

  1 parameter *D* (?), at NLO $R = 0.7$, $R_{sep} = 1.3 \Leftrightarrow D = 0.83$

- Jet identification is unique – no merge/split stage 👍

- Resulting jets are more amorphous, energy calibration difficult (subtraction for UE?), and analysis can be very computer intensive (time grows like N^3, recalculate list after each merge) 👎
  But new version goes like N ln N (only recalculate nearest neighbors) 👍

In the future: (comments, not criticisms)

- When we look carefully will we find problems and add details ? History says yes!
- The (official?) k_T webpage has 5 parameters to specify the implementation, resolution variable, combination scheme, etc.
- Recall the Cambridge k_T (e^+e^-) algorithm that added angular ordering to get rid of “junk jets” (resolution variable 🕒 ordering variable) and “soft-freezing” to reduce mis-clustering

Jet Algorithm Summary:

- Seeds & pQCD are a bad mix (not IRS). **It is better to correct for seeds during the analysis of the data (a small correction) and compare to theory w/o seeds (so no IRS issue) !!**

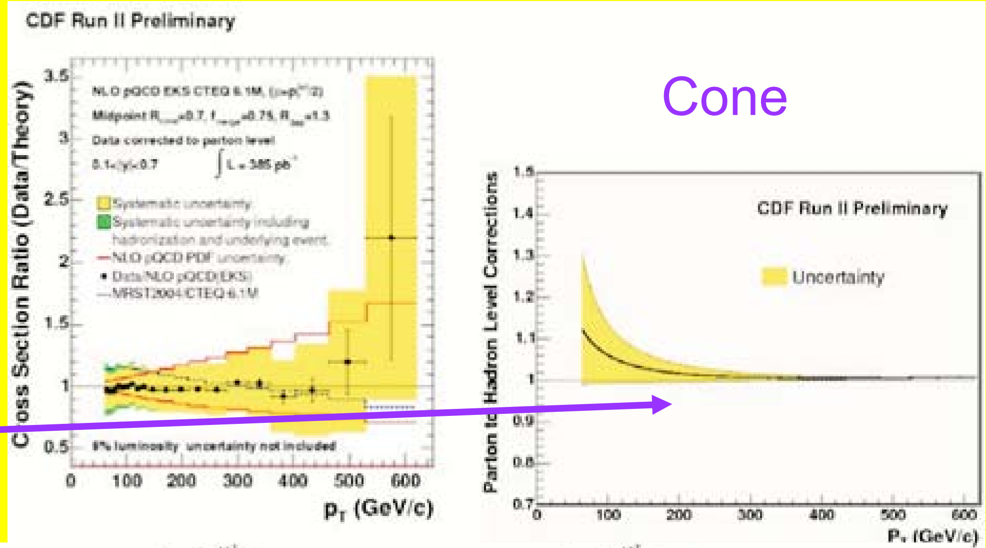
- Dark towers are a real 5 - 10% effect, but the search cone fix aggravates the IRS issue – better to recognize as a correction during the analysis of the data (or the theory), along with corrections for detector, UE, hadronization, seeds, and missing 2-in-1 configurations

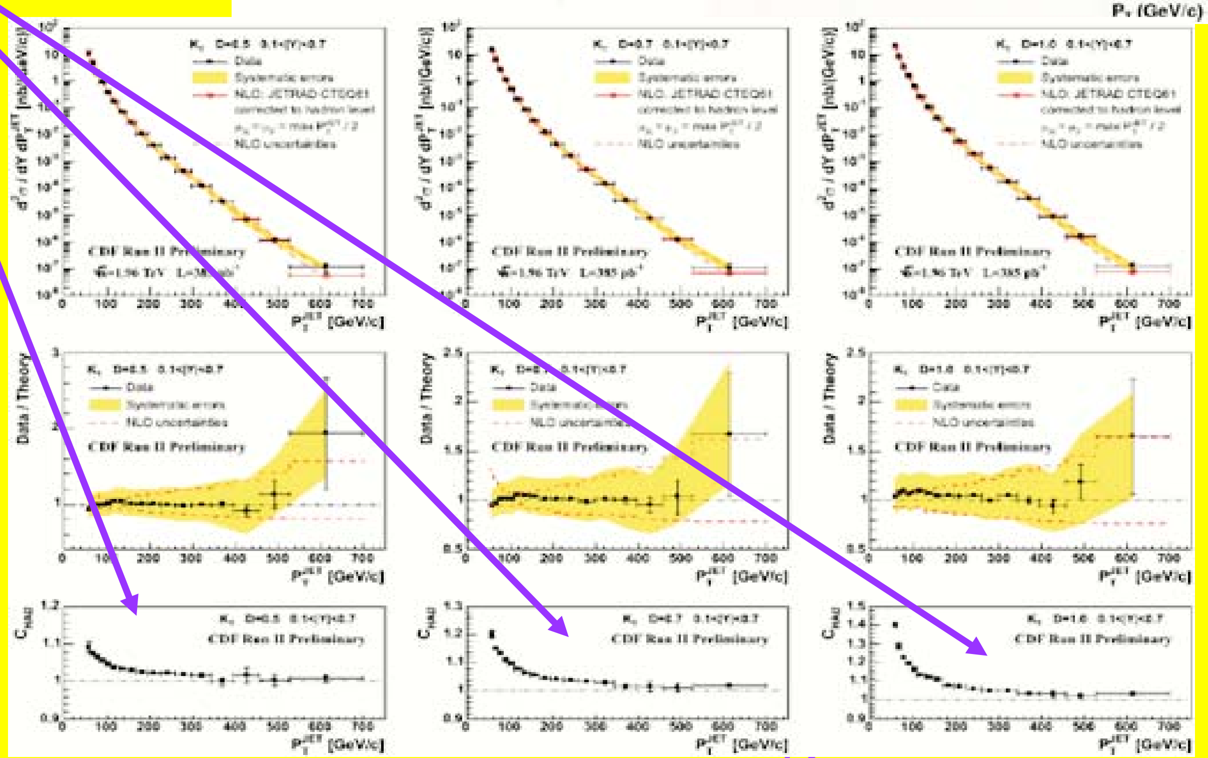
- Compare corrected experimental numbers to pQCD without seeds and $R_{sep} = 2$

- Need serious phenomenology study of the k_T algorithm

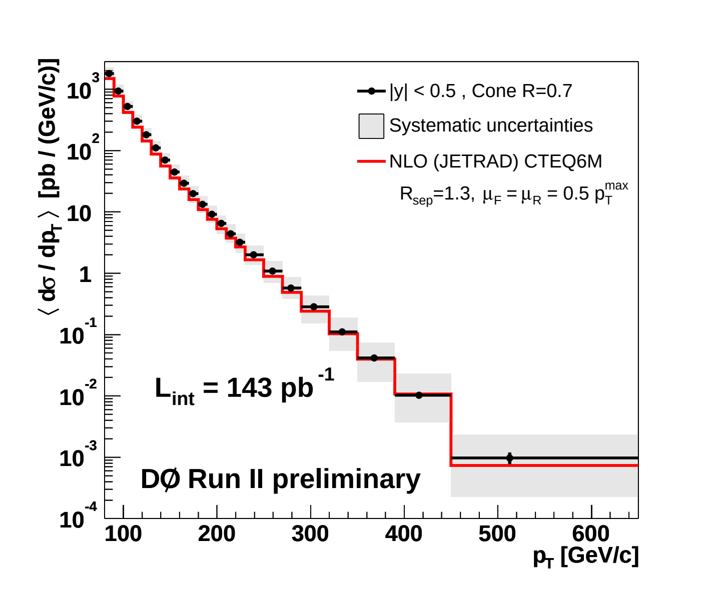
Seed and Dark Tower corrections ≤ current CDF corrections for hadrons → partons

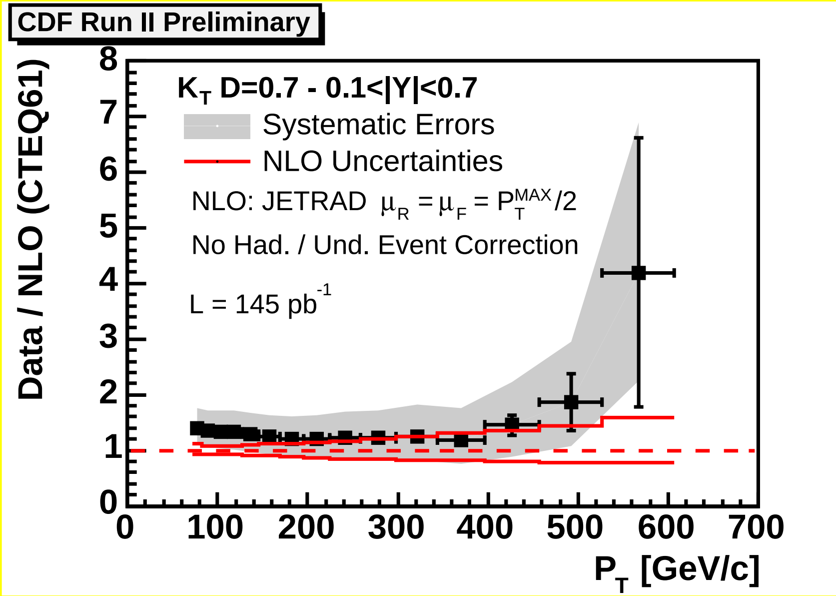
Cone

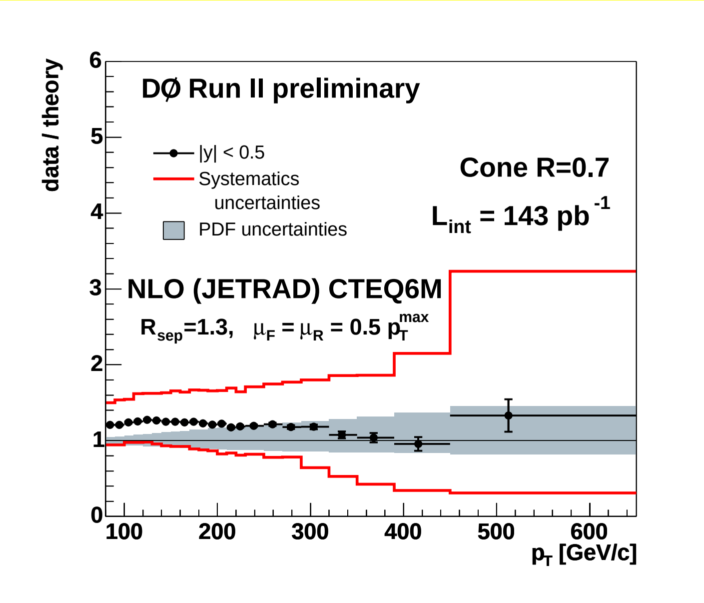


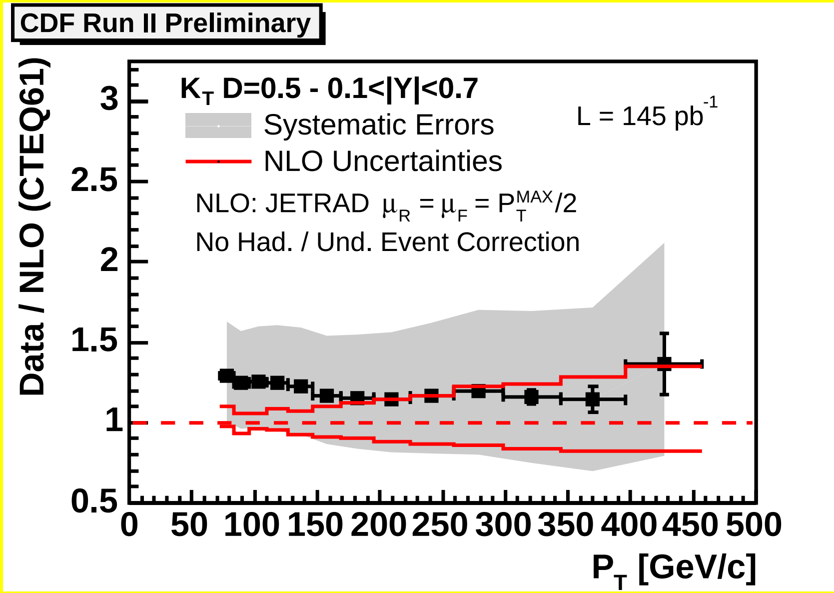


K_T


|y| < 0.5 , Cone R=0.7
Systematic uncertainties
NLO (JETRAD) CTEQ6M
R_{sep}=1.3, $\mu_F = \mu_R = 0.5\ p_T^{max}$
L_{int} = 143 pb^{-1}
DØ Run II preliminary
$\langle d\sigma / dp_T \rangle$ [pb / (GeV/c)]
p_T [GeV/c]


CDF Run II Preliminary
K_T D=0.7 - 0.1<|Y|<0.7
Systematic Errors
NLO Uncertainties
NLO: JETRAD $\mu_R = \mu_F = P_T^{MAX}/2$
No Had. / Und. Event Correction
L = 145 pb^{-1}
Data / NLO (CTEQ61)
P_T [GeV/c]


DØ Run II preliminary
|y| < 0.5
Systematics uncertainties
PDF uncertainties
Cone R=0.7
L_{int} = 143 pb^{-1}
NLO (JETRAD) CTEQ6M
R_{sep}=1.3, $\mu_F = \mu_R = 0.5\ p_T^{max}$
data / theory
p_T [GeV/c]


CDF Run II Preliminary
K_T D=0.5 - 0.1<|Y|<0.7
Systematic Errors
L = 145 pb^{-1}
NLO Uncertainties
NLO: JETRAD $\mu_R = \mu_F = P_T^{MAX}/2$
No Had. / Und. Event Correction
Data / NLO (CTEQ61)
P_T [GeV/c]

Extra Detail Slides

Dictionary of Hadron Collider Terminology

EVENT

HADRON-HADRON COLLISION

Primary (Hard) Parton-Parton Scattering

Initial-State Radiation (ISR) = Spacelike Showers associated with Hard Scattering

Underlying Event

Multiple Parton-Parton Interactions: Additional parton-parton collisions (in principle with showers etc) in the same hadron-hadron collision.

= Multiple Perturbative Interactions (MPI)

= Spectator Interactions

Beam Remnants: Left over hadron remnants from the incoming beams. Colored and hence correlated with the rest of the event →

Fragmentation

Perturbative:

Final-State Radiation (FSR)

= Timelike Showers

= Jet Broadening and Hard Final-State Bremsstrahlung

Non-perturbative:

String / Cluster Hadronization

(Color Reconnections?)

PILE-UP: Additional hadron-hadron collisions recorded as part of the same event.

From Peter Skands

Example Lego & Flow