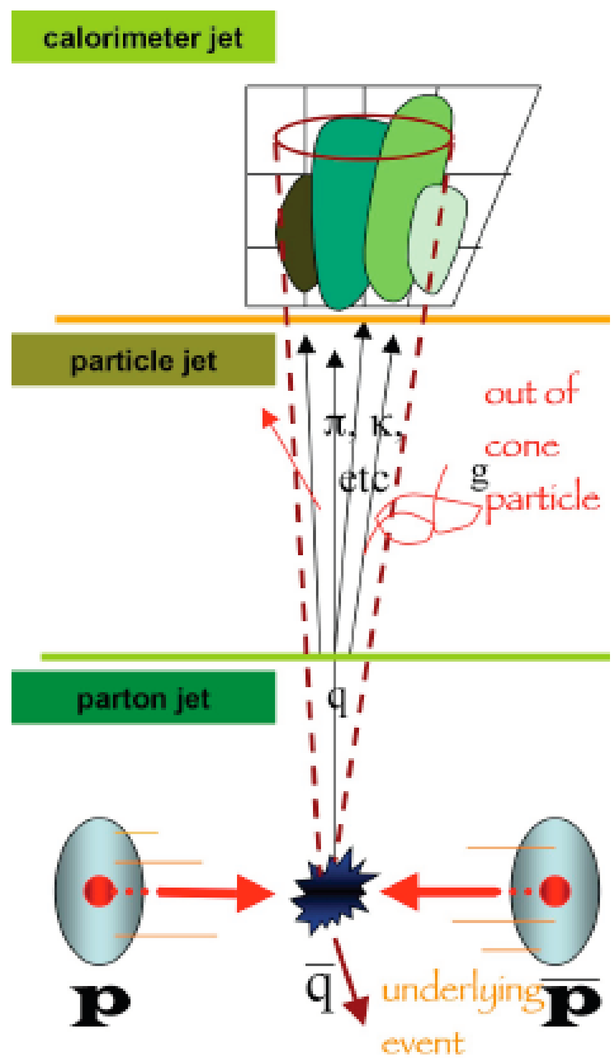
Jet Algorithms for the LHC
…framing the issues

Joey Huston
Michigan State University
…apologies for not being
there in person

MC4LHC

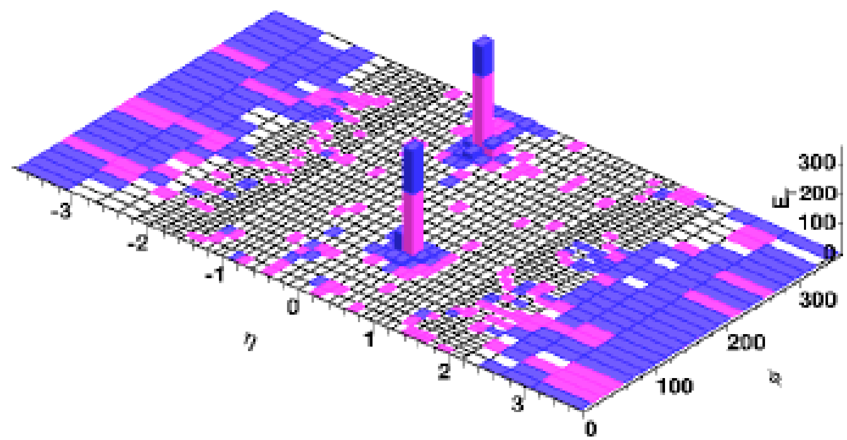
Jet algorithms

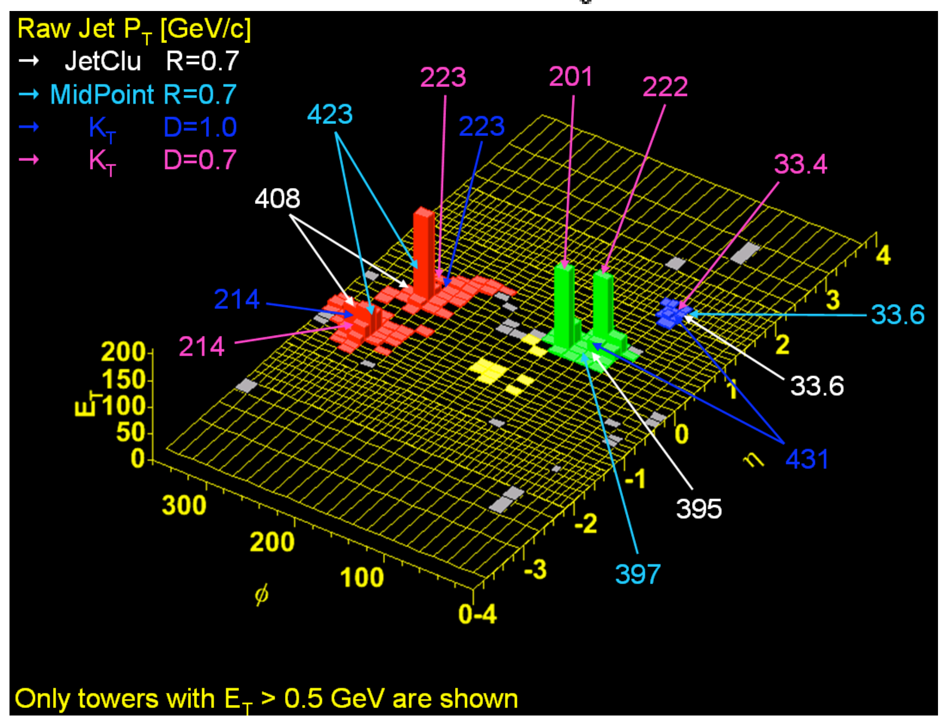
- References
  - www.pa.msu.edu/~huston/Les_Houches_2005/Les_Houches_SM.html and references therein
  - www.pa.msu.edu/~huston/seminars/Main.pdf
  - my talk during plenary session



Jet algorithms

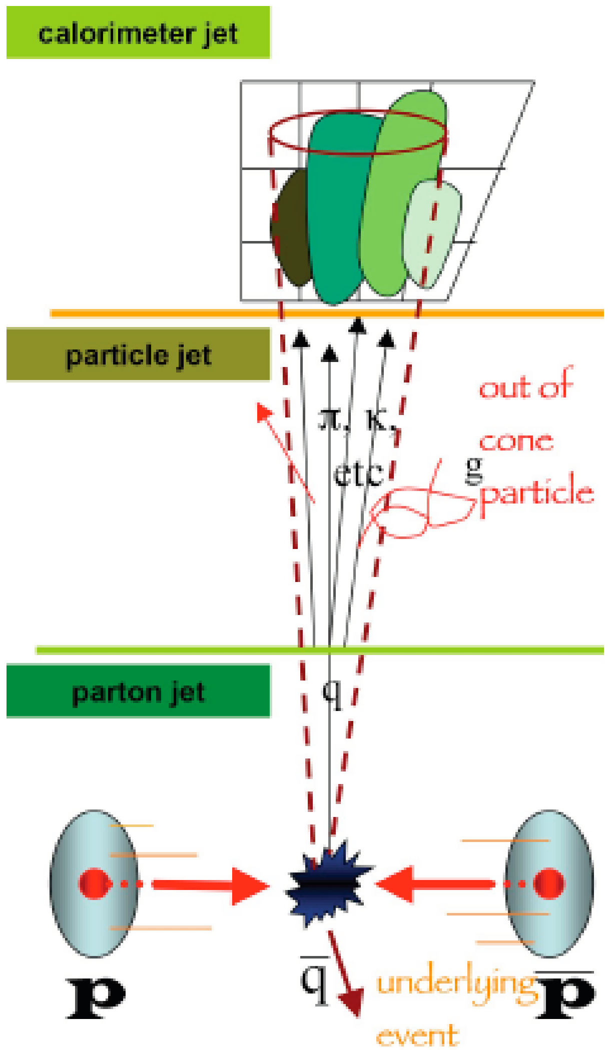
- For some events, the jet structure is very clear and there's little ambiguity about the assignment of towers to the jet
- But for other events, there is ambiguity and the jet algorithm must make decisions that impact precision measurements
- If comparison is to hadron-level Monte Carlo, then hope is that the Monte Carlo will reproduce all of the physics present in the data and influence of jet algorithms can be understood
  - but needs to be studied for precision physics topics, such as top mass determination
  - more difficulty when comparing to parton level (HO) calculations





Algorithms: some statements

- Jet algorithms should be able to operate on parton, particle and calorimeter levels
  - and corrections from one level to another should be clearly specified/determined → themes for workshop
  - where possible, data should be presented at the hadron level
  - where possible, results should be presented with both cone and k_T algorithms
    - other algorithms: Cambridge, JEF?
  - where possible, ATLAS and CMS should use jet algorithms with as many elements in common as possible
    - same algorithms for QCD/non-QCD analyses?
    - an LHC accord?



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
Coloured and hence correllated with the rest of the event →

Fragmentation

Perturbative:

Final-State Radiation (FSR)

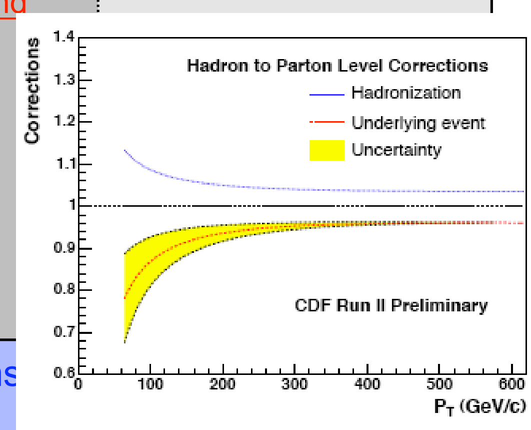
= Timelike Showers

= Jet Broadening and Hard Final-State Bremsstrahlung

Non-perturbative:

String / Cluster Hadronisation

(Colour Reconnections?)



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

So what's the problem(s) with cone algorithms

- Matching a cone algorithm at (NLO) parton level and at detector level
- To illustrate, construct a Snowmass potential which indicates where stable cone solutions can be found

•$z = p_T^{jet2}/p_T^{jet1}$; d=ΔR between partons
•At NLO; two partons within region I or II will be called one jet
•R_{sep} parameter was introduced into the theory because experiment reconstructs separate jets if $\Delta R > R_{sep}$*Rcone

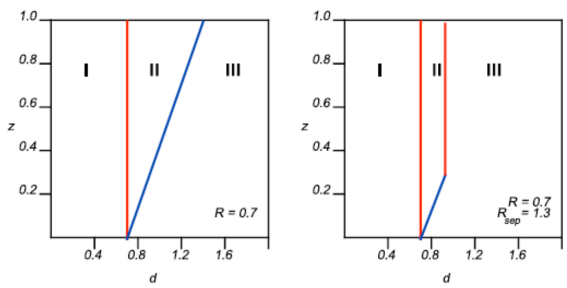
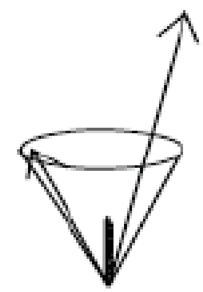


Figure 15. The parameter space (d,Z) for which two partons will be merged into a single jet.

midpoint seed was intended to remove need for R_{sep}



terms of the 2-dimensional vector $\overrightarrow{r} = (y, \phi)$ via

$$V(\overrightarrow{r}) = -\frac{1}{2}\sum_j p_{T,j}\left(R_{cone}^2 - (\overrightarrow{r_j} - \overrightarrow{r})^2\right)\Theta\left(R_{cone}^2 - (\overrightarrow{r_j} - \overrightarrow{r})^2\right). \quad (39)$$

The flow is then driven by the "force" $\overrightarrow{F}(\overrightarrow{r}) = -\overrightarrow{\nabla} V(\overrightarrow{r})$ which is thus given by,

$$\overrightarrow{F}(\overrightarrow{r}) = \sum_j p_{T,j}(\overrightarrow{r_j} - \overrightarrow{r})\Theta\left(R_{cone}^2 - (\overrightarrow{r_j} - \overrightarrow{r})^2\right)$$
$$= \left(\overrightarrow{r}_{C(\overrightarrow{r})} - \overrightarrow{r}\right)\sum_{j \subset C(r)} p_{T,j}, \quad (40)$$

where $\overrightarrow{r}_{C(\overrightarrow{r})} = \left(\overline{y}_{C(\overrightarrow{r})}, \overline{\phi}_{C(\overrightarrow{r})}\right)$ and the sum runs over $j \subset C(\overrightarrow{r})$ such that $\sqrt{(y_j - y)^2 + (\phi_j - \phi)^2} \le R_{cone}$. As desired, this force pushes the cone to the stable cone position.

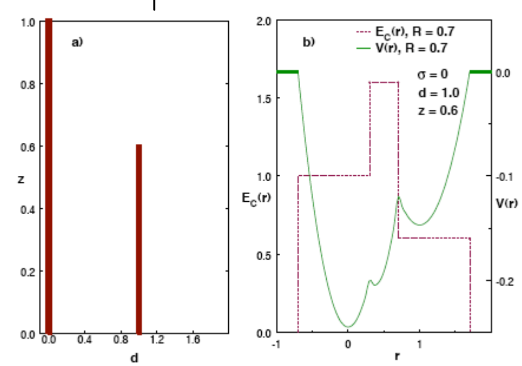


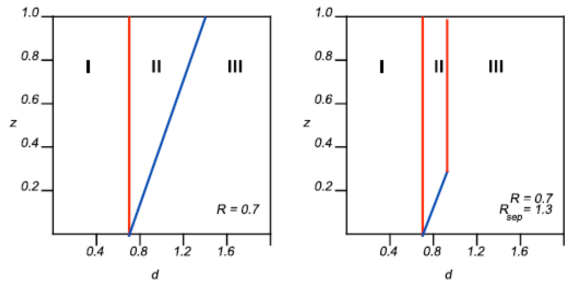
Figure 18. A schematic depiction of a specific parton configuration and the results of applying the midpoint cone jet clustering algorithm. The potential discussed in the text and the resulting energy in the jet are plotted.

stable solution at position of left parton, at right parton and at midpoint, but there's no parton seed at midpoint

So what's the problem(s)

- Matching a cone algorithm at (NLO) parton level and at detector level
- Parton configurations that will be included in a jet at NLO will not be at hadron level due to stochastic smearing because of parton showering/hadronization

•$z=p_T^{jet2}/p_T^{jet1}$; d=ΔR between partons
•At NLO; two partons within region I or II will be called one jet
•R_{sep} parameter was introduced into the theory because experiment reconstructs separate jets if $\Delta R > R_{sep}$*Rcone



midpoint seed was intended to remove need for R_{sep}
…but it's smearing not seeds

Figure 15. The parameter space (d,Z) for which two partons will be merged into a single jet.

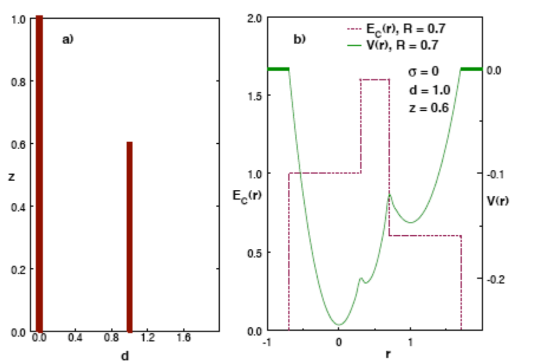


Figure 18. A schematic depiction of a specific parton configuration and the results of applying the midpoint cone jet clustering algorithm. The potential discussed in the text and the resulting energy in the jet are plotted.

have lost central solution (both partons) and right solution… some energy ends up unclustered in any jet

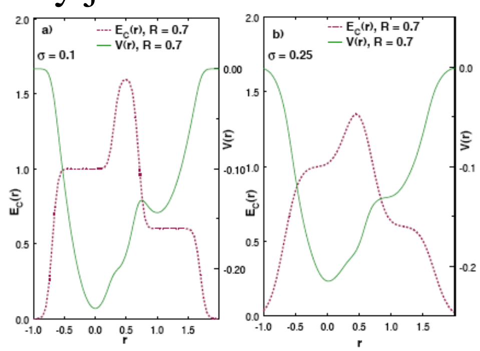
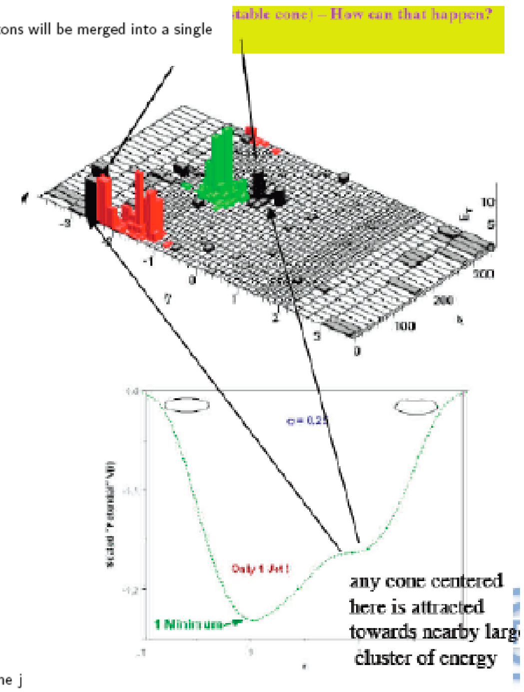


Figure 19. A schematic depiction of the effects of smearing on the midpoint cone j clustering algorithm.



Some major silliness

- Matching a cone algorithm at (NLO) parton level and at detector level
- Parton configurations that will be included in a jet at NLO will not be at hadron level due to stochastic smearing because of parton showering/hadronization
- Modified midpoint algorithm uses smaller initial search cone (R/2), reduces unclustered energy
  - recovers right solution, but in most cases not central
    - i.e. R_{sep} still needed
    - consider this an interim solution
  - default midpoint algorithm has ~2% of 400 GeV/c dijet events with >50 GeV/c of unclustered energy
- **All cone algorithms with seeds are IR-sensitive**
  - D0 version of midpoint algorithm has IR-sensitivity <1%
  - CDF version has IR-sensitivity of ~1%
    - but essentially no unclustered energy

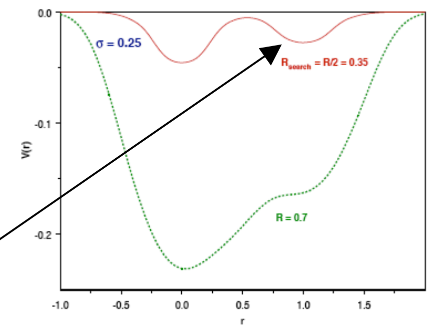
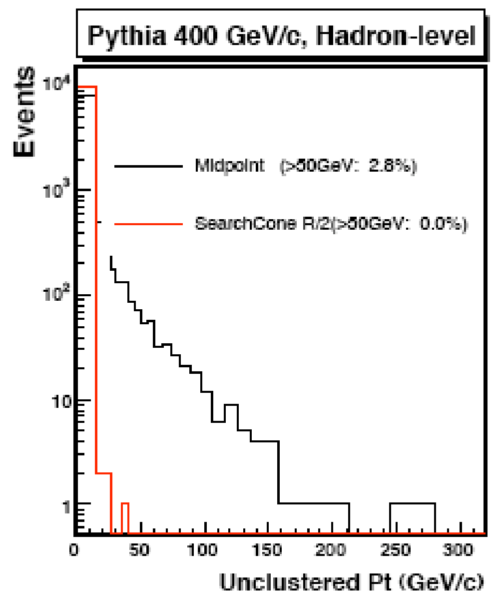
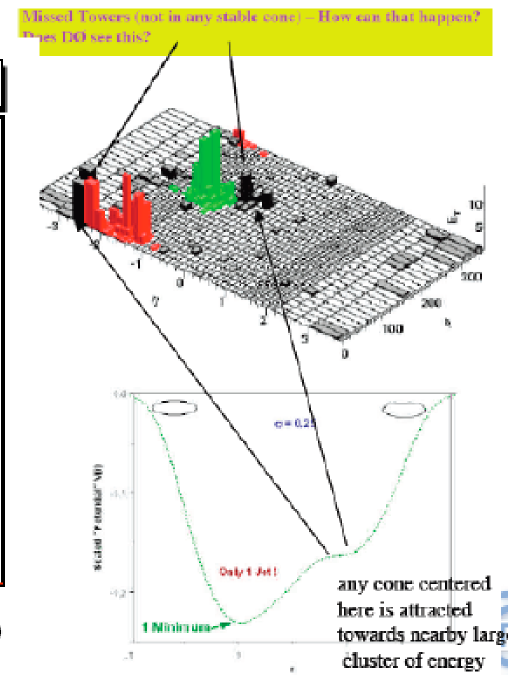


Figure 20. A schematic depiction of the effects of smearing on the midpoint cone jet clustering algorithm and the result of using a smaller initial search cone.

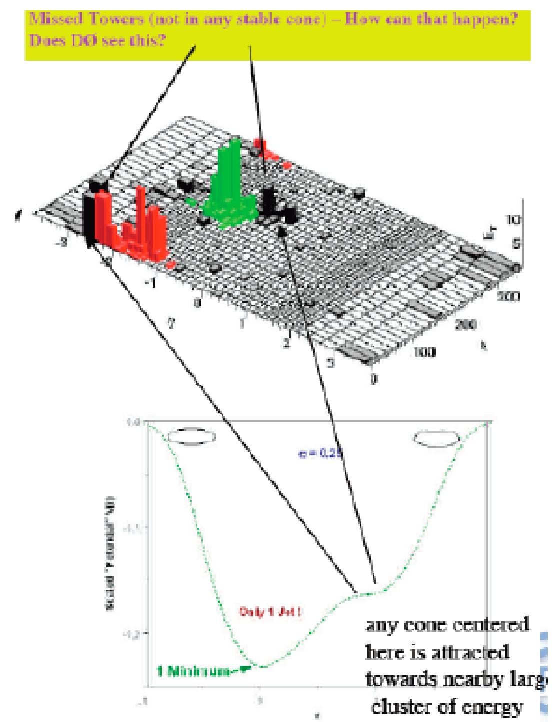




Solution(s)

- Experimental level
  - run standard (out-of-box) midpoint algorithm
  - after first pass, remove towers clustered into jets
  - run algorithm again on remaining towers
  - merge jet pairs in Region II on left…or
- Theoretical level
  - use appropriate R_{sep} (~1.3) in theory calculation

- What about f_{merge} parameter (% of overlap of two jets before they're merged)?



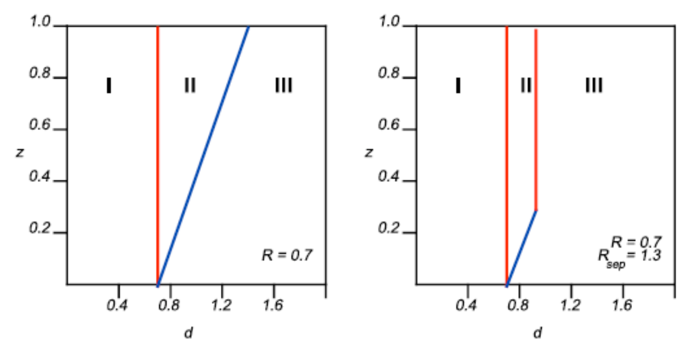
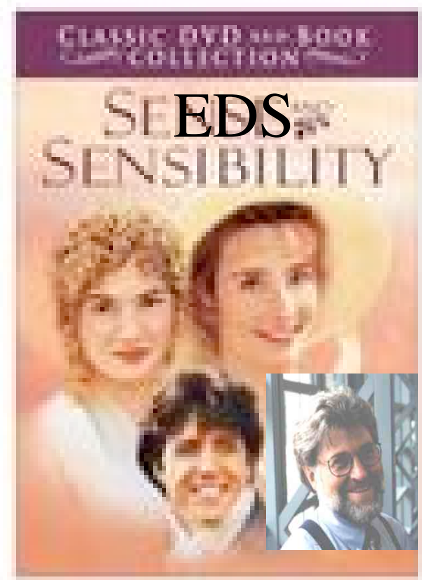


Figure 15. The parameter space (d,Z) for which two partons will be merged into a single jet.

Seeds and sensibility

- To save on computer time, experiments require seeds for initiation of jet cone searches
  - impact on experimental cross section compared to seedless algorithm is small
- Seeds have also been used in the theoretical calculations, but here the number of potential seeds is small
  - the requirement for seeds introduces a dependence on soft gluon emission
  - the midpoint algorithm removes this (logarithmic) dependence to NNLO, but not for higher orders
- Steve's suggestion: if you must use seeds in your experimental algorithm, correct to seedless level before comparison to data
  - see Steve's talk



- much larger corrections already performed by experiments

So what's the problem(s) with k_T algorithms

- Lack of experience at hadron-hadron colliders
  - little study of what the problems might be
- Underlying event and multiple interaction subtractions
  - little work has been done to date for dealing with a high luminosity environment
- Can we try to put the use of the k_T algorithm on a more robust basis?

Charge

- From discussions here, can we formulate a document/working group to carry these ideas forward?

Benchmark studies for LHC

- Goal: produce predictions/event samples corresponding to 1 and 10 fb^{-1}
- Cross sections will serve as
  - benchmarks/guidebook for SM expectations in the early running
    - are systems performing nominally? are our calorimeters calibrated?
    - are we seeing signs of “unexpected” SM physics in our data?
    - how many of the signs of new physics that we undoubtedly will see do we really believe?
  - feedback for impact of ATLAS data on reducing uncertainty on relevant pdf’s and theoretical predictions
  - venue for understanding some of the subtleties of physics issues
- Has gone (partially) into Les Houches proceedings
- *Companion* review article on hard scattering physics at the LHC by John Campbell, James Stirling and myself
  - “Hard Interactions at the LHC: a primer for LHC physics”
  - www.pa.msu.edu/~huston/seminars/Main.pdf

SM benchmarks for the LHC

LES HOUCHES

centre de physique

See www.pa.msu.edu/~huston/_
Les_Houches_2005/Les_Houches_SM.html

- pdf luminosities and uncertainties
- expected cross sections for useful processes
  - inclusive jet production
    - simulated jet events at the LHC
    - jet production at the Tevatron
      - a link to a CDF thesis on inclusive jet production in Run 2
      - CDF results from Run II using the kT algorithm
  - photon/diphoton
  - Drell-Yan cross sections
  - W/Z/Drell Yan rapidity distributions
  - W/Z as luminosity benchmarks
  - W/Z+jets, especially the *Zeppenfeld* plots
  - top pairs
    - ongoing work, list of topics (pdf file)

More…

- technical benchmarks
  - jet algorithm comparisons
    - midpoint vs simple iterative cone vs kT
      - top studies at the LHC
      - an interesting data event at the Tevatron that examines different algorithms
    - Building Better Cone Jet Algorithms
      - one of the key aspects for a jet algorithm is how well it can match to perturbative calculations; here is a 2-D plot for example that shows some results for the midpoint algorithm and the CDF Run 1 algorithm (JetClu)
      - here is a link to Fortran/C++ versions of the CDF jet code
  - fits to underlying event for 200 540, 630, 1800, 1960 GeV data
    - interplay with ISR in Pythia 6.3
    - establish lower/upper variations
    - extrapolate to LHC
    - effect on target analyses (central jet veto, lepton/photon isolation, top mass?)