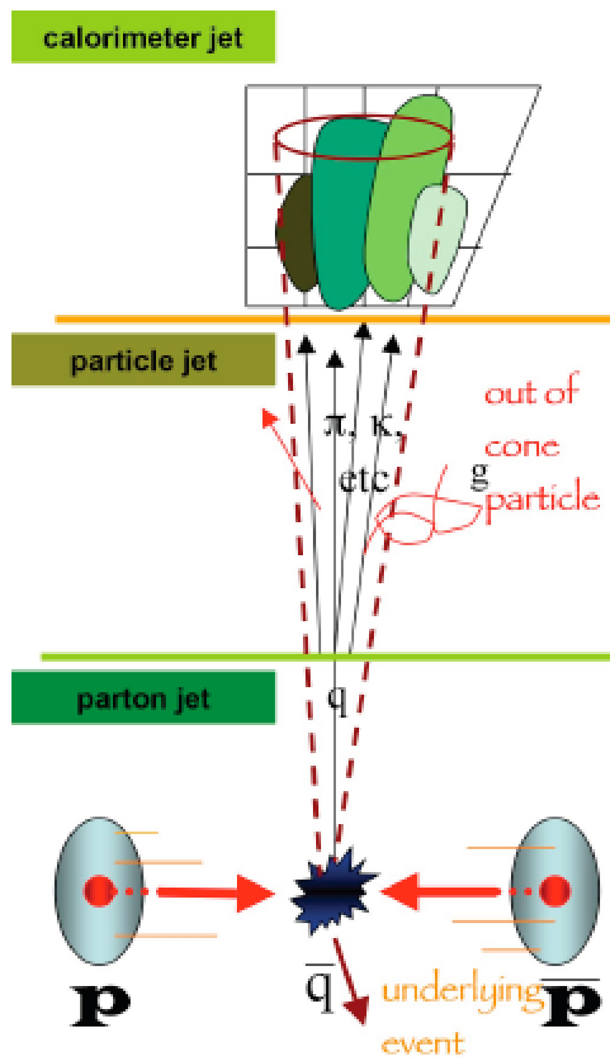
Jet Algorithms for the LHC
…and some other stuff

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

MC4LHC

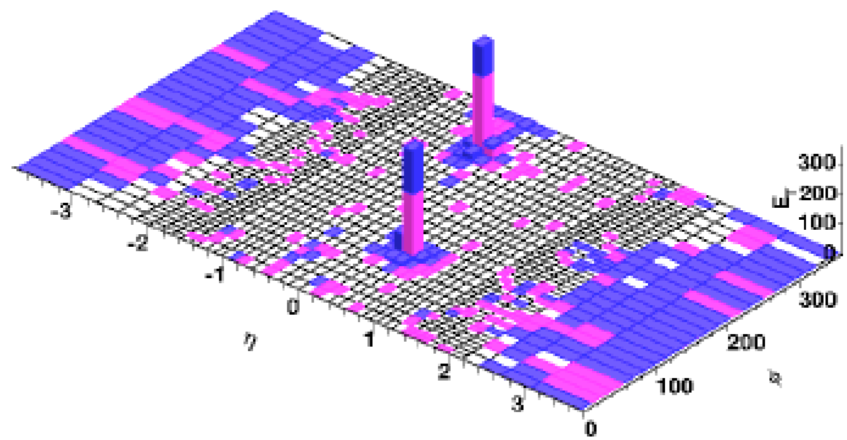
Jet algorithms

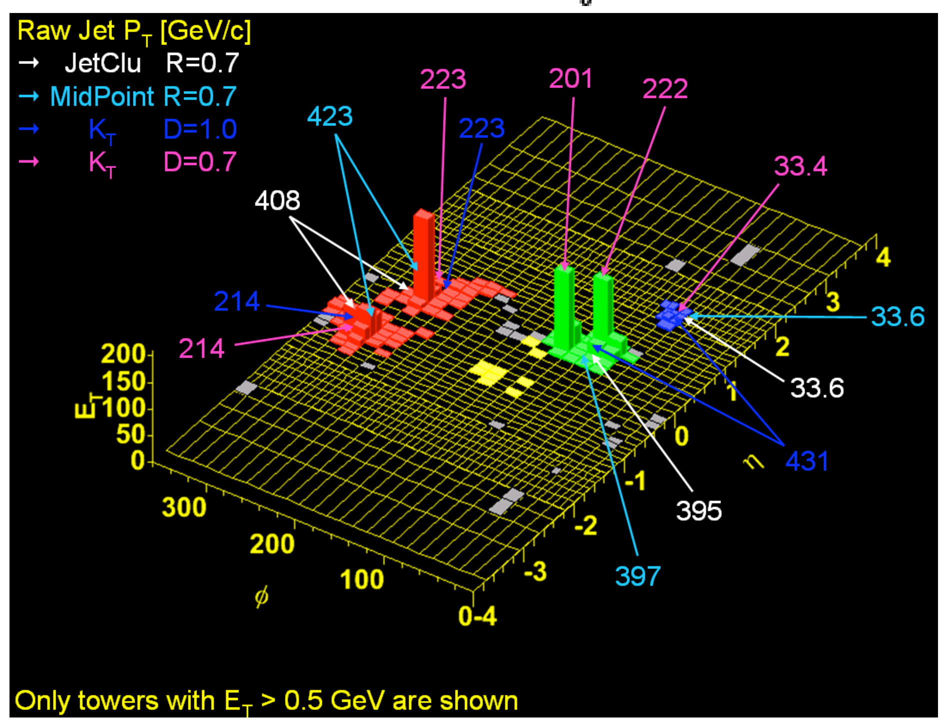
- To date, emphasis in ATLAS and CMS has been (deservedly so) on jet energy calibration and not on details of jet algorithms
- But attention to the latter will be necessary for precision physics at the LHC
  - this is a good time for ATLAS, CMS and theorists to be talking
- Big effort at Les Houches 2005 on this aspect
  - I've been collecting some information on jet algorithms and other aspects of LHC physics at some benchmark webpages
  - www.pa.msu.edu/~huston/Les_Houches_2005/Les_Houches_SM.html
- Also part of TeV4LHC workshop



Jet algorithms

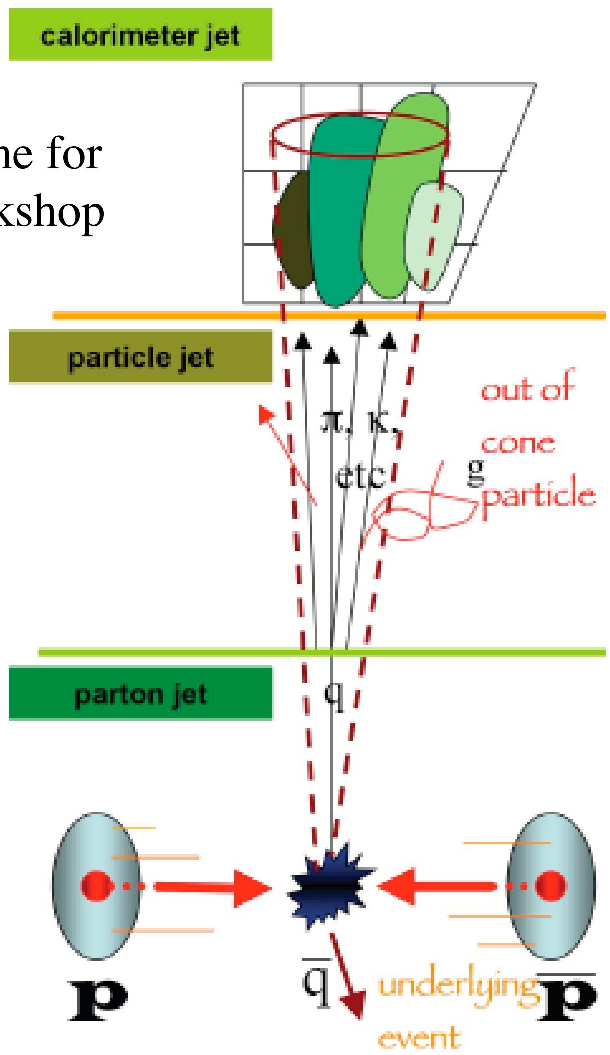
- For some events, the jet structure is very clear and there's little ambiguity about the assignment of towers to the jet
- But for other events, there is ambiguity and the jet algorithm must make decisions that impact precision measurements
- If comparison is to hadron-level Monte Carlo, then hope is 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





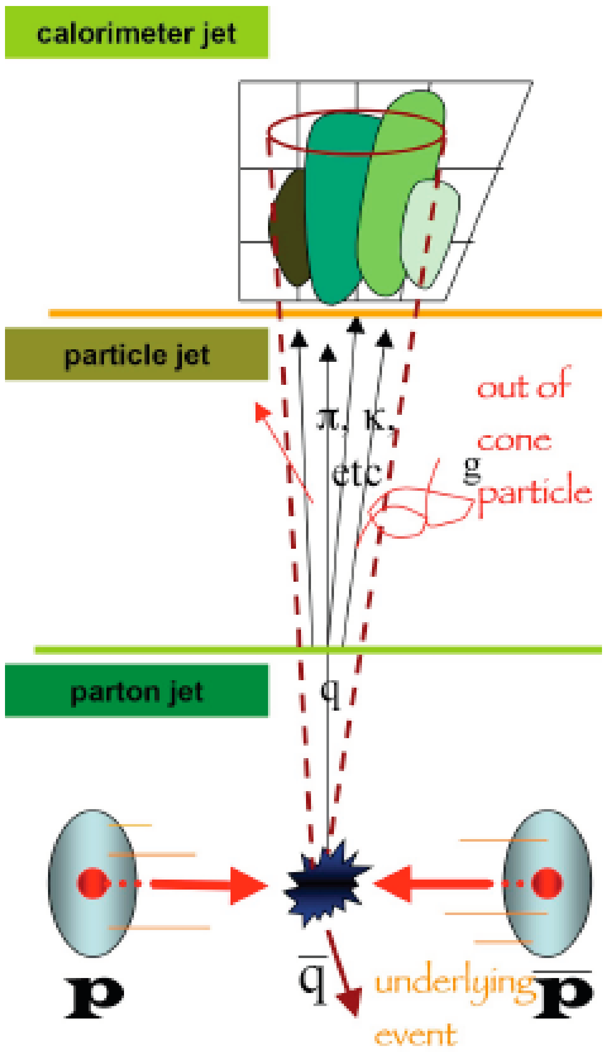
Algorithms

- 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 → theme for workshop
- Jet algorithms can either measure closeness in coordinate space (cone) or in momentum space (k_T)
  - connection between the two, as we'll hear in Matteo Cacciari's talk later in this workshop
  - almost all experience at the Tevatron is with cone algorithm; mostly k_T at LEP/HERA
    - this is a problem as we rely on the Tevatron for (recent) experience with hadron-hadron colliders



Cone Algorithms

- As mentioned before, almost all experience at the Tevatron is with cone algorithms
- Why?
  - trigger uses cone-like algorithm
  - underlying event and multiple-interaction correction is trickier with k_T than with cone algorithm
  - comfort-level with cone algorithm due to long experience
- CDF (JetClu) and D0 used own versions of iterative cone algorithm in Run 1
- Midpoint cone algorithm developed for (joint use in) Run 2
- Differences have developed between implementation of midpoint in last few years
- NB: both CDF and D0 cone algorithms require presence of seeds to start searches for jet cones
  - implicit infra-red sensitivity, but numerically small



Midpoint algorithm

* Midpoint algorithm: cone algorithm (R=0.7) (clusters particles whose trajectories/towers are close together). In contrast to JetClu's use of E_T and η, Midpoint uses p_T and ~~ϕ~~ y

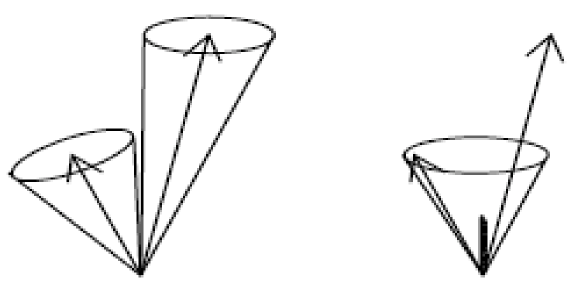
* Need to do this consistently at parton, hadron and calorimeter level.

* Calor Jets: begin with 1 GeV seed towers† → cluster towers ($P_T >$ 100MeV) into a centroid if $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta Y)^2} < 0.7$.

* Start new search cones at the midpoints of stable cones‡

* Overlapping jets- merge jets if overlapping energy is $>$ 0.75 the energy of the smaller jet (choice: D0 uses 0.50)

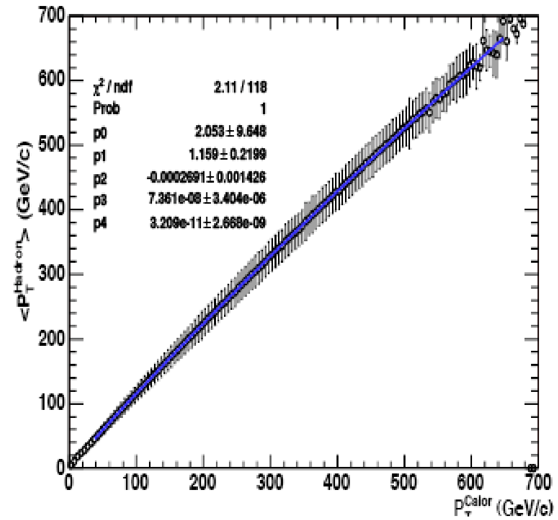
* Calculate jet quantites from final stable cones: P_T, ~~E_T~~, Y, ϕ etc



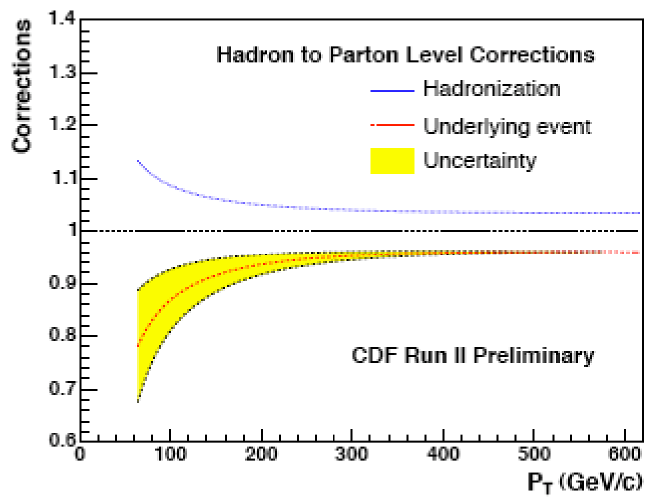
† Clustering begins around seeds, presence of soft radiation can cause merging of jets
* Ideally algroithm is insensitive to soft radiation.
‡ Addition of midpoints lessens the sensitivity

Jet Corrections

- **Need to correct from calorimeter to hadron level** → 2nd theme for this workshop. See Ben Cooper's talk.



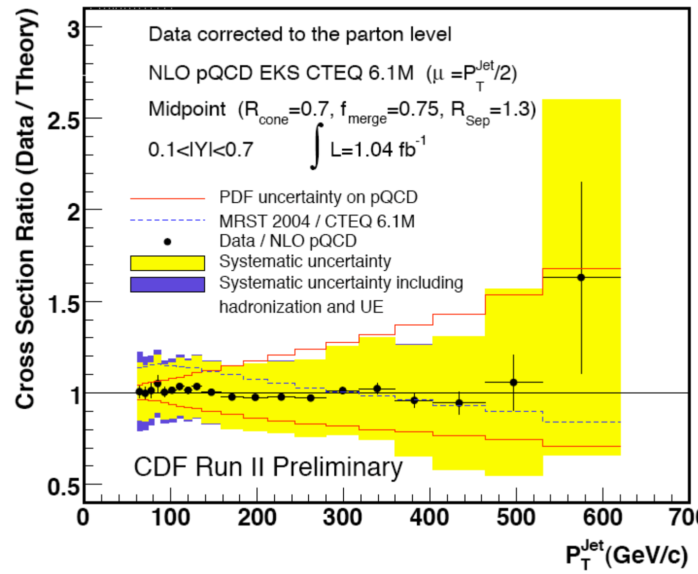
- And for
  - underlying event and out-of-cone for some observables
  - resolution effects
  - hadron to parton level for other observables (such as comparisons to parton level cross sections)
    - can correct data to parton level or theory to hadron level…or both and be specific about what the corrections are
  - note that loss due to hadronization is basically constant at 1 GeV/c for all jet p_T values at the Tevatron (for a cone of radius 0.7)
    - for a cone radius of 0.4, the two effects cancel to within a few percent
  - interesting to check over the jet range at the LHC

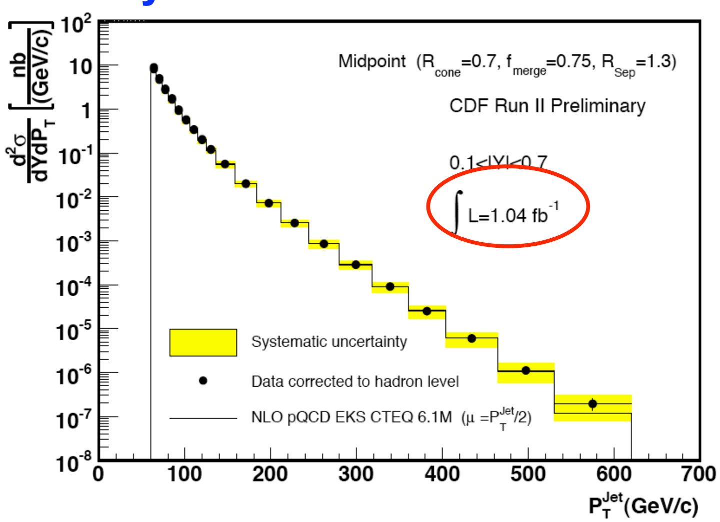


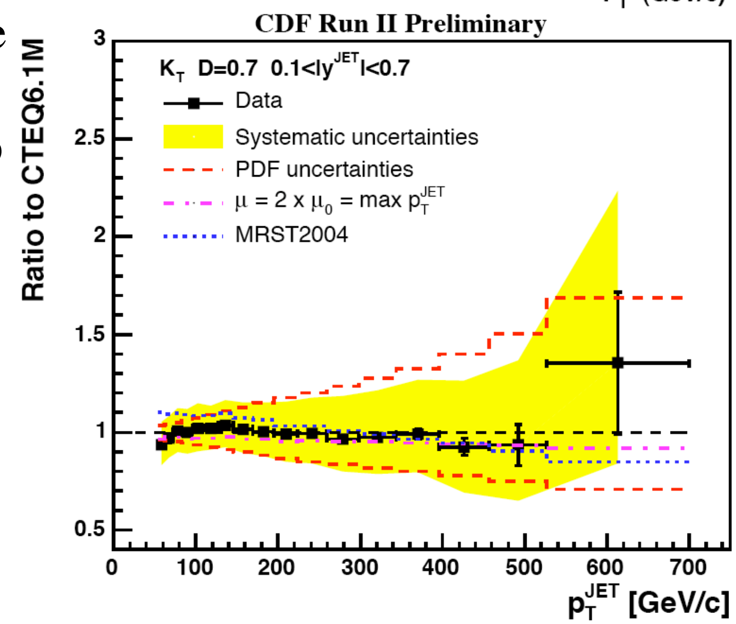
note Pythia and Herwig hadronization corrections are the same

CDF Run 2 results

- CDF Run II result in good agreement with NLO predictions using CTEQ6.1 pdf's
  - enhanced gluon at high x
  - I've included them in some new CTEQ fits leading to new pdf's
- …and with results using k_T algorithm
  - the agreement would appear even better if the same scale were used in the theory (k_T uses $p_T^{max}/2$)
- **need to have the capability of using different algorithms in analyses as cross-checks** → 3rd theme of workshop

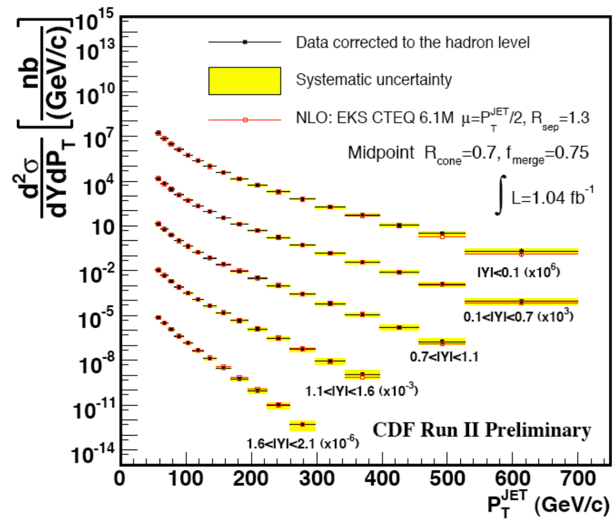


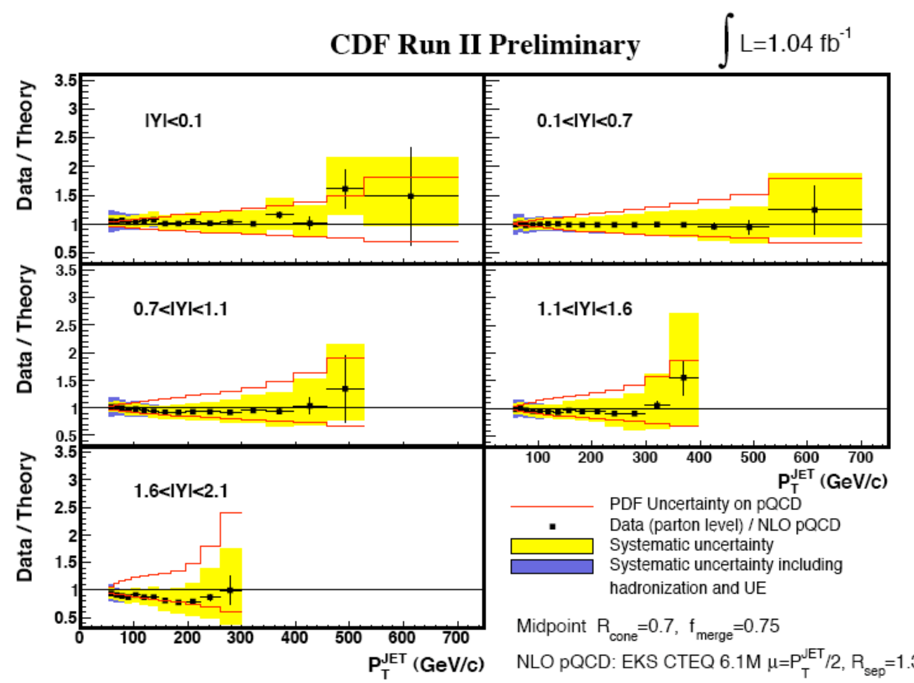




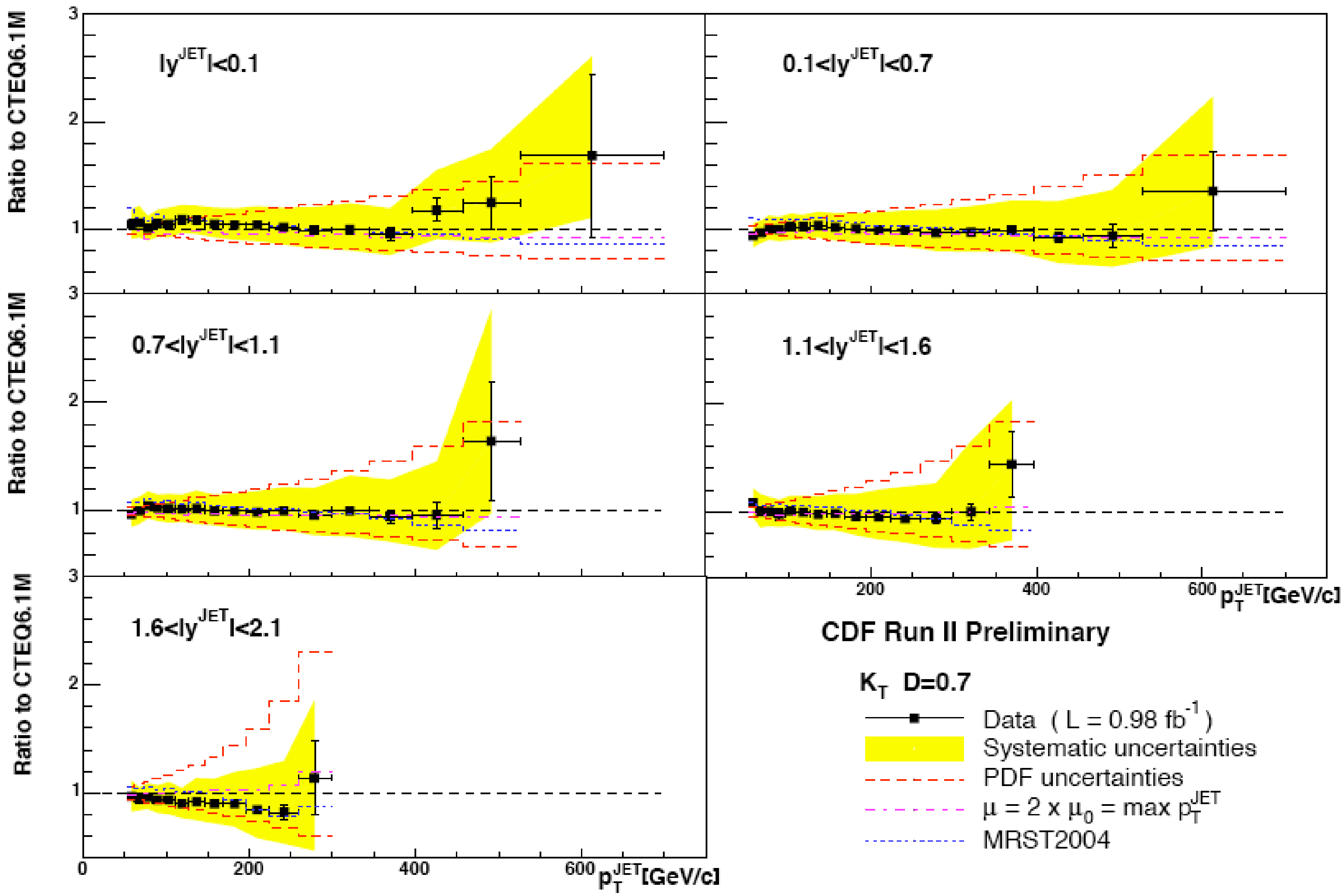
CDF Run 2 cone results

- Precise results over a wide rapidity range
- Good agreement with CTEQ6.1 predictions using CDF midpoint algorithm
- PDF uncertainties are on the same order or less than systematic errors
- Should reduce uncertainties for next round of CTEQ fits
  - so long to eigenvector 15?





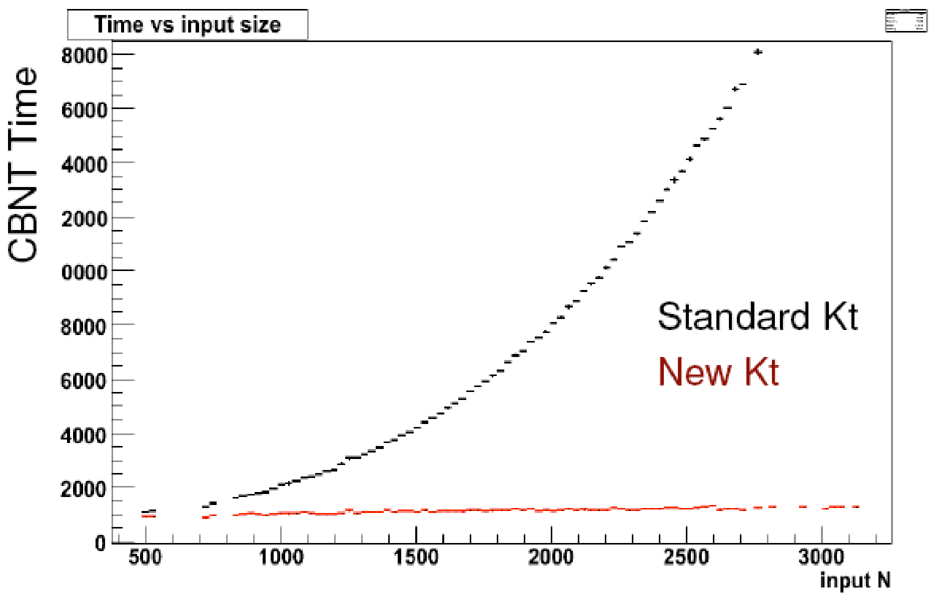
Forward jets with the k_T algorithm



Need to go lower in p_T for comparisons of the two algorithms, apply k_T to other analyses. Would love to see tT with k_T algorithm, for example.

New k_T algorithm

- k_T algorithms are typically slow because speed goes as $O(N^3)$, where N is the number of inputs (towers, particles,…)
- Cacciari and Salam (hep-ph/0512210) have shown that complexity can be reduced and speed increased to O(NlnN) by using information relating to geometric nearest neighbors
  - should be useful for LHC
  - already implemented in ATLAS
  - see Matteo's talk on Thurs



So what's the problem(s)

- 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

•$z = p_T^{jet2}/p_T^{jet1}$; $d = \Delta R$ 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

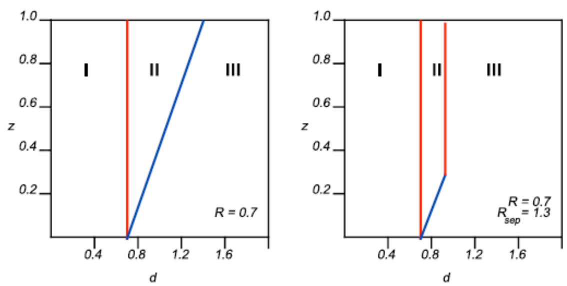
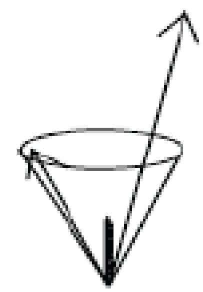


Figure 15. The parameter space (d,Z) for which two partons will be merged into a single jet.

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



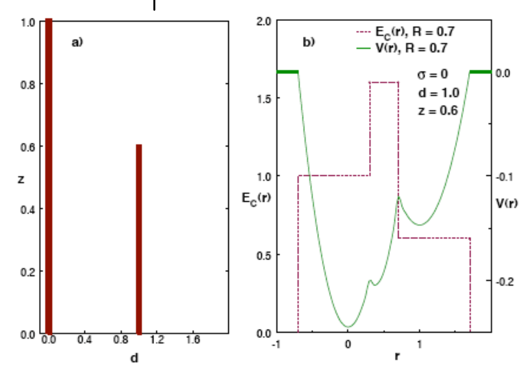
terms of the 2-dimensional vector $\vec{r} = (y, \phi)$ via

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

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

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

where $\vec{r}_{C(\vec{r})} = \left(\overline{y}_{C(\vec{r})}, \overline{\phi}_{C(\vec{r})}\right)$ and the sum runs over $j \subset C(\vec{r})$ 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.



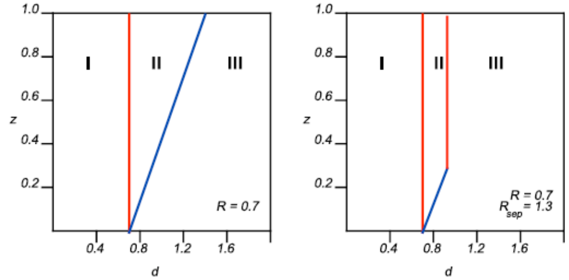
stable solution at position of left parton, at right parton and at midpoint, but there's no parton seed at midpoint

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.

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.

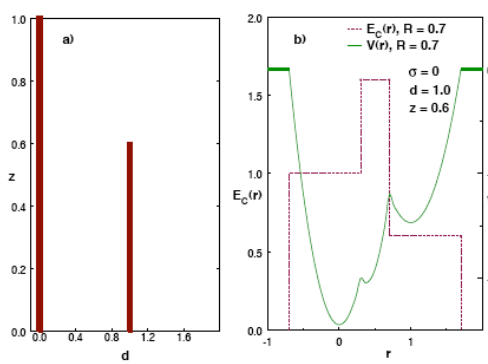


Figure 18. A schematic depiction of a specific parton configuration and the results of applying the midpoint cone jet clustering algorithm. The potential discussed in the text and the resulting energy in the jet are plotted.

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

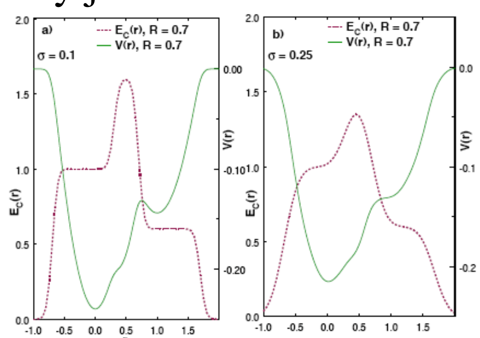
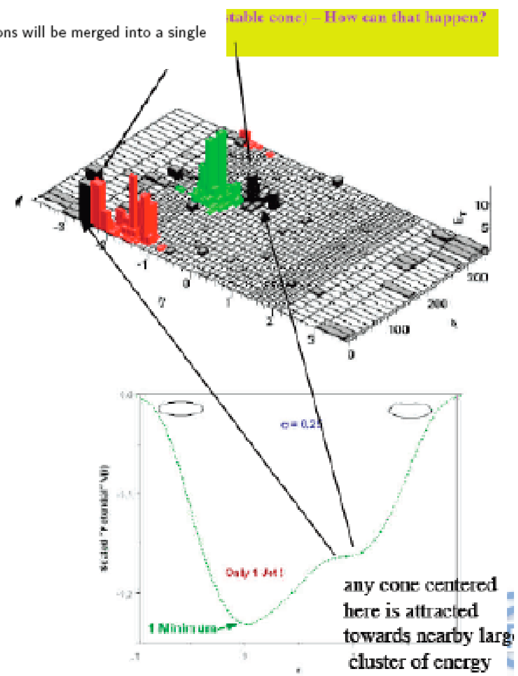


Figure 19. A schematic depiction of the effects of smearing on the midpoint cone j clustering algorithm.



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

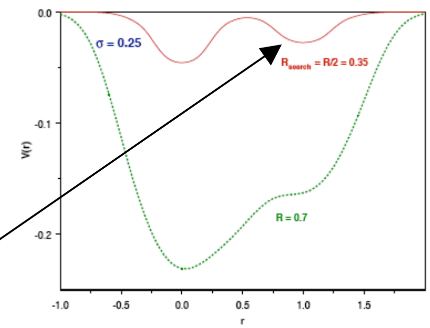
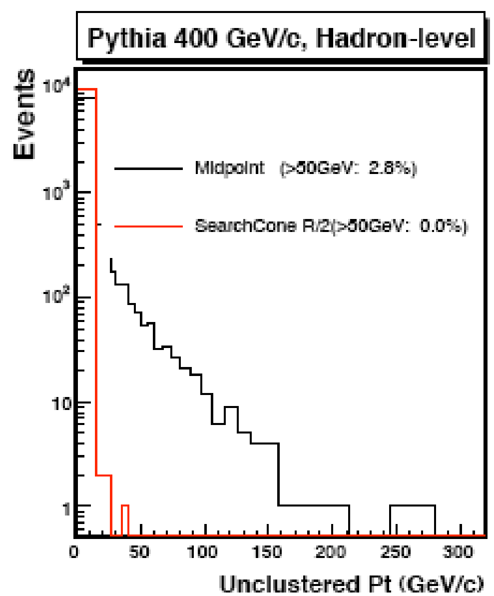
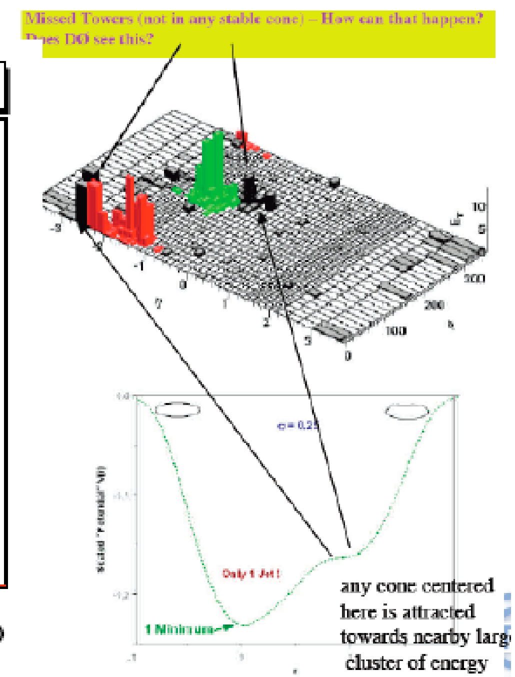


Figure 20. A schematic depiction of the effects of smearing on the midpoint cone jet clustering algorithm and the result of using a smaller initial search cone.



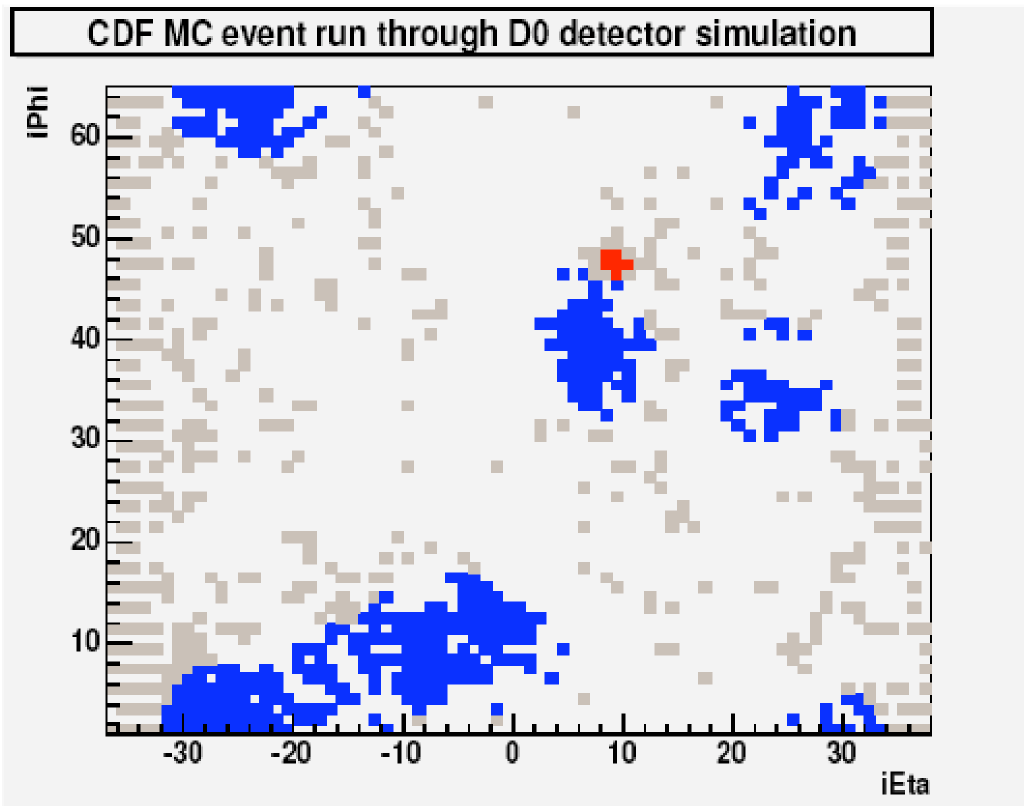


Jet algorithms

- The idea of the mid-point cone algorithm was to
  - provide more perturbative stability for the theoretical calculations
  - provide a jet algorithm common to CDF, D0 and theorists
- But to the strong disappointment of at least one theorist, CDF and D0 are using different implementations of the midpoint algorithm in Run 2
- Let's not disappoint him at the LHC

D0 report at the TeV4LHC meeting at CERN

- To address CDF observation of unclustered E_T



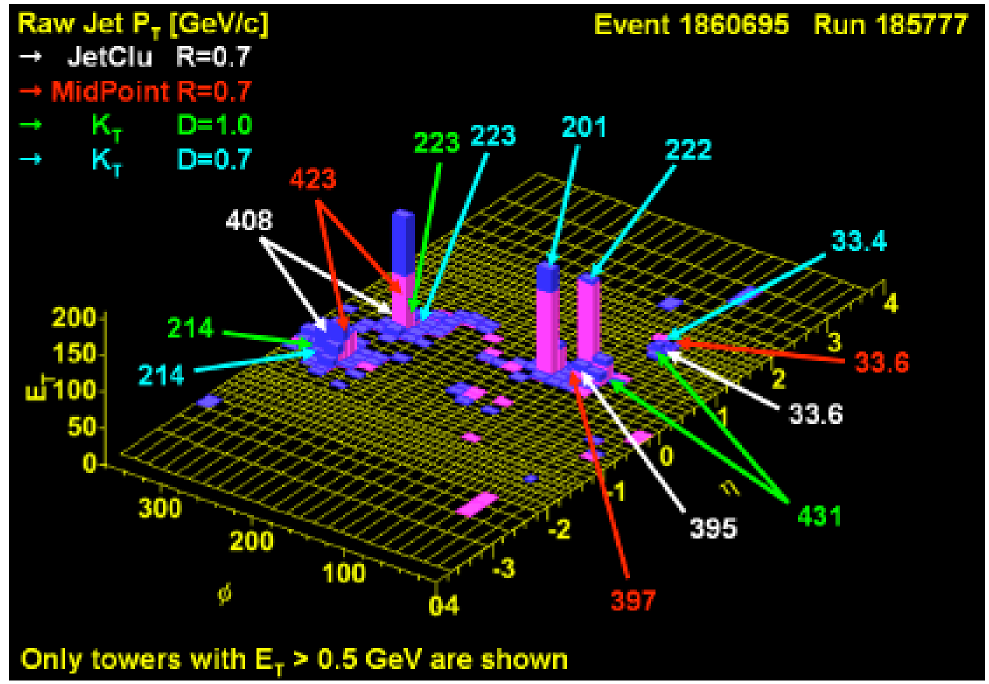
- RunII cone $R = 0.7$
- **Jet** towers
- **Unclustered** towers $pT < 2GeV$
- **Unclustered** towers $pT > 2GeV$

We see it too!

What about ATLAS and CMS? Currently investigating.

Can't we all just get along?

- I still believe that at the LHC, need both k_T and cone jet algorithms
- I'm working now on a version of the jet cone algorithm that matches as closely as possible seedless pQCD
    - trying to bypass both Scylla and Charybdis

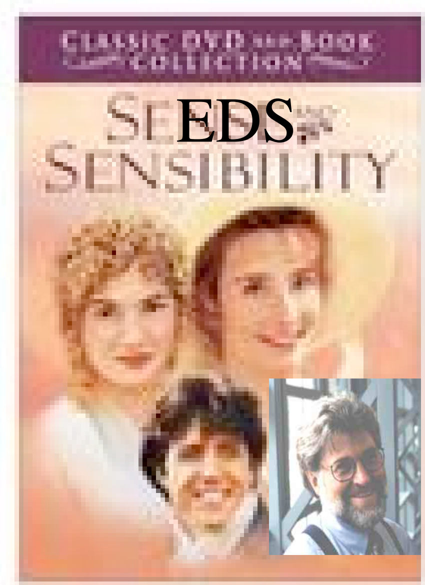


- Trying to summarize/think for TeV4LHC writeup
- Further discussion this workshop

Steve, Ken Hatakeyama and myself are working on a review paper on this subject

Seeds and sensibility

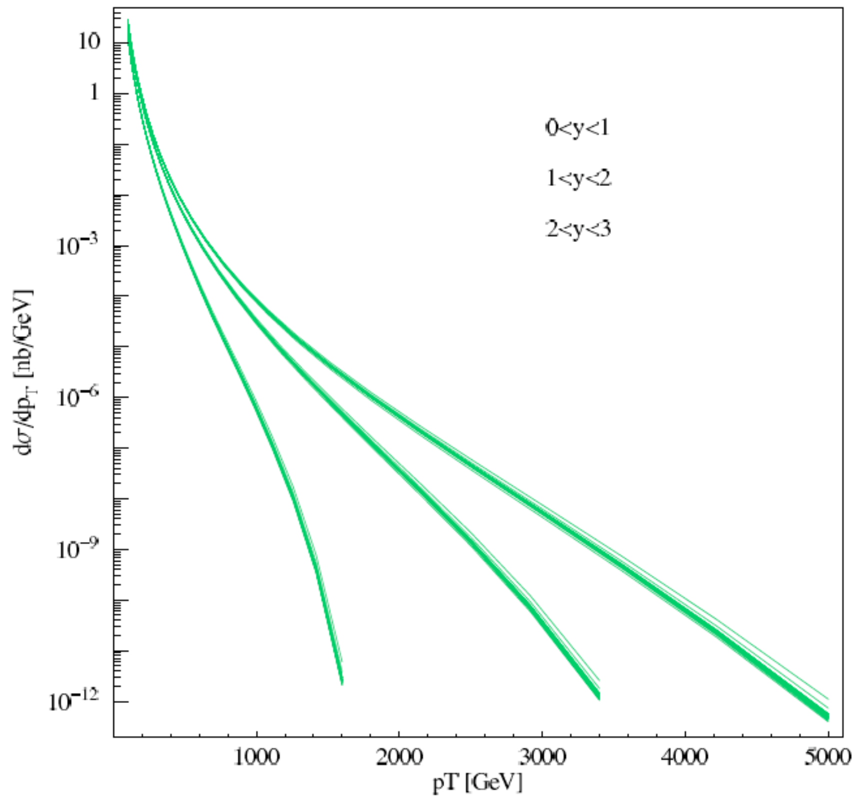
- To save on computer time, experiments require seeds for initiation of jet cone searches
  - impact on experimental cross section compared to seedless algorithm is small
- Seeds have also been used in the theoretical calculations, but here the number of potential seeds is small
  - the requirement for 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

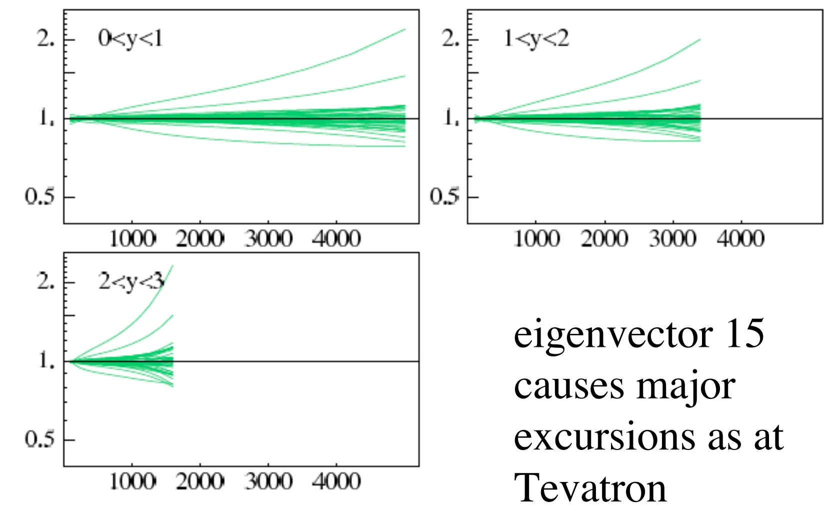


- much larger corrections already performed by experiments

Cross section predictions for LHC

These are predictions for ATLAS based on the CTEQ6.1 central pdf and the 40 error pdf's using the midpoint jet algorithm.





eigenvector 15 causes major excursions as at Tevatron

FIG. 31: The uncertainty range of the inclusive jet cross section at the LHC. The curves are graphs of the ratios of the cross sections for the 40 eigenvector basis sets compared to the central (CTEQ6.1M) prediction (ordinate) versus p_T in GeV (ordinate).

Need to have jet measurements over full rapidity range and good control over rapidity variations of jet systematics.

- γ+jet balancing
- dijet balancing

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

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

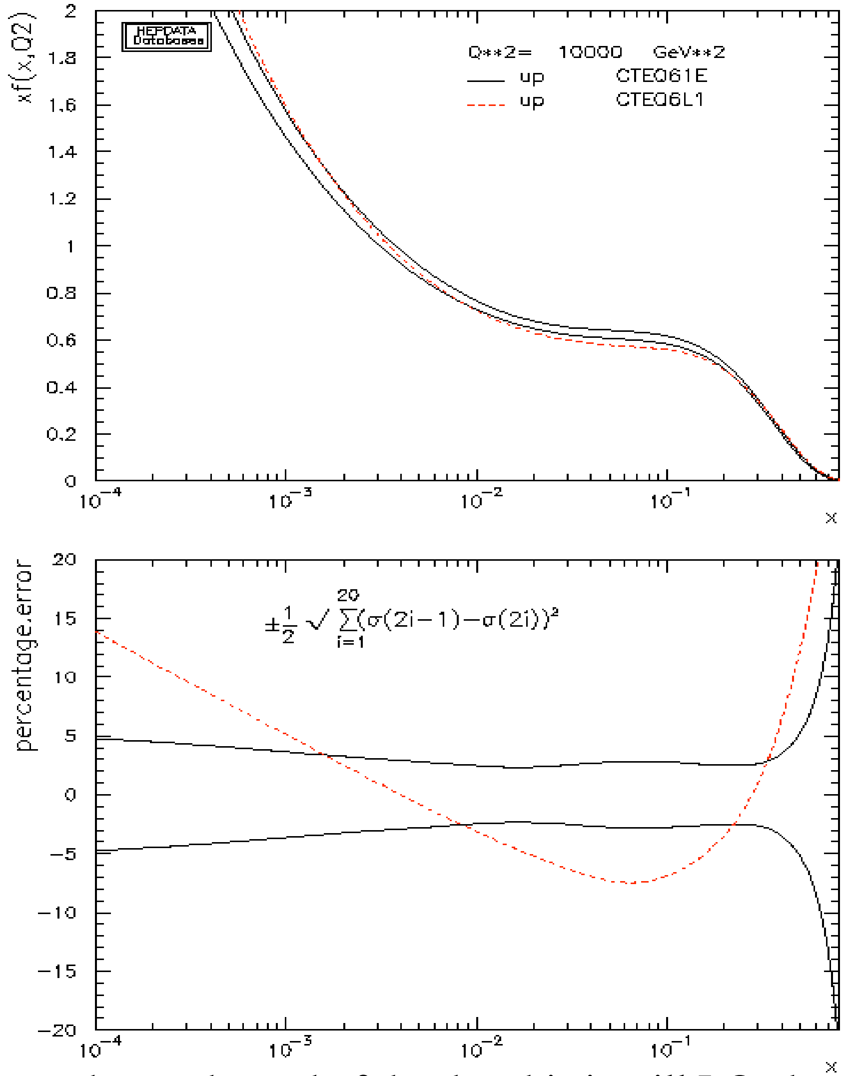
- pdf luminosities and uncertainties
- expected cross sections for useful processes
  - inclusive jet production
    - simulated jet events at the LHC
    - jet production at the Tevatron
      - a link to a CDF thesis on inclusive jet production in Run 2
      - CDF results from Run II using the kT 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?)

Last issue: LO vs NLO pdf's for parton shower MC's

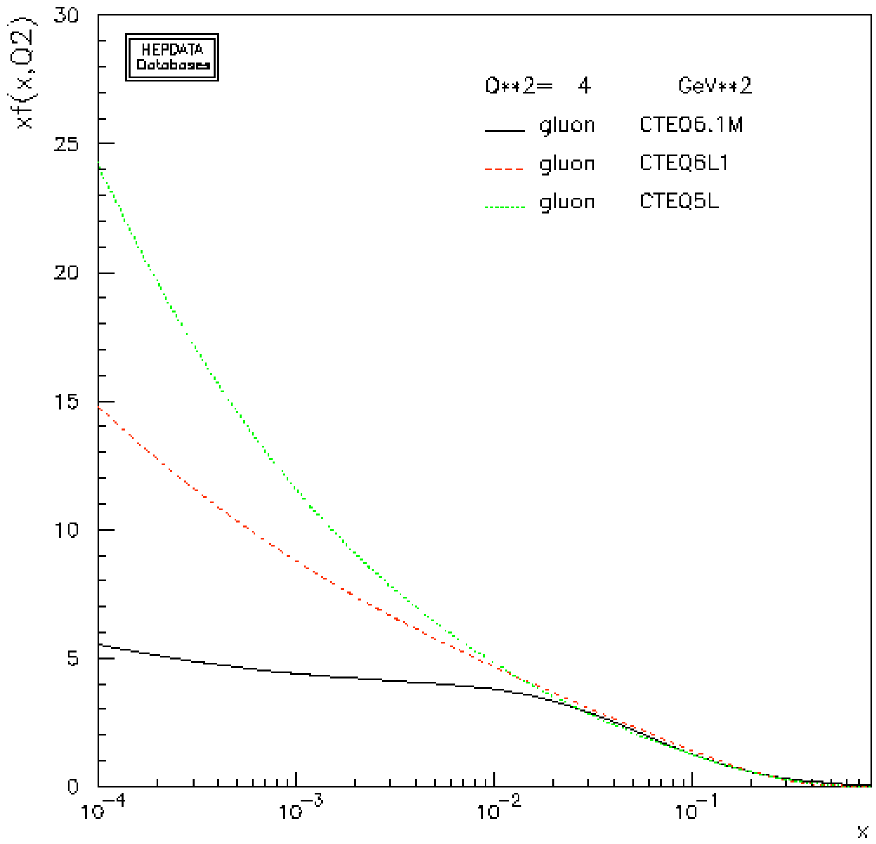
- For NLO calculations, use NLO pdf's (duh)
- What about for parton shower Monte Carlos?
  - somewhat arbitrary assumptions (for example fixing Drell-Yan normalization) have to be made in LO pdf fits
  - DIS data in global fits affect LO pdf's in ways that may not directly transfer to LO hadron collider 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

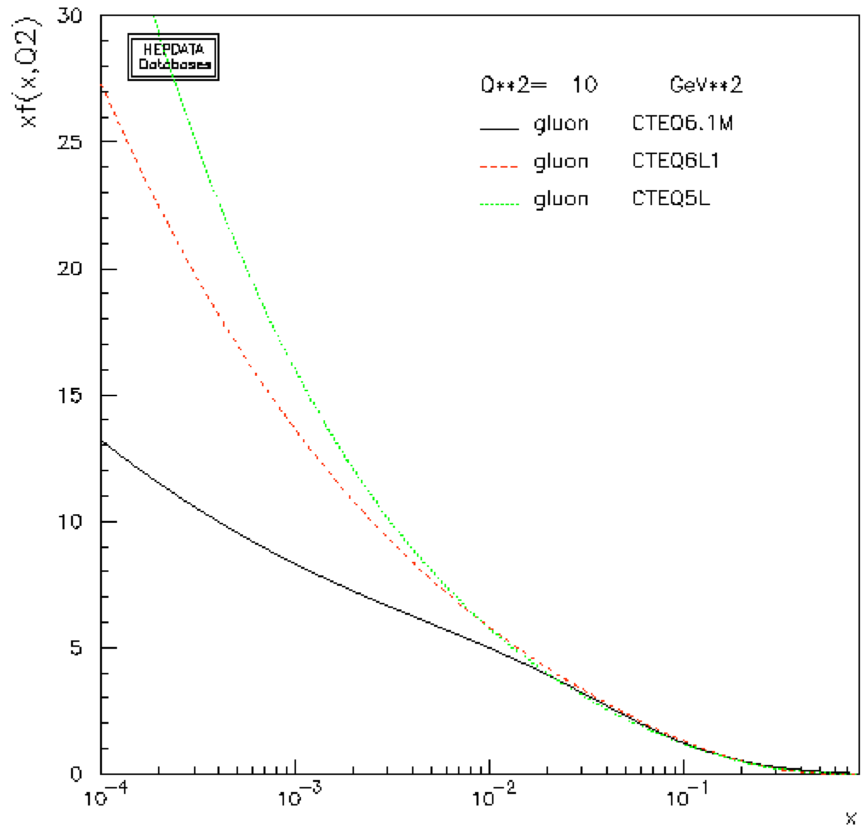


. . .but at the end of the day this is still LO physics; There's no substitute for honest-to-god NLO.

Impact on UE tunes

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





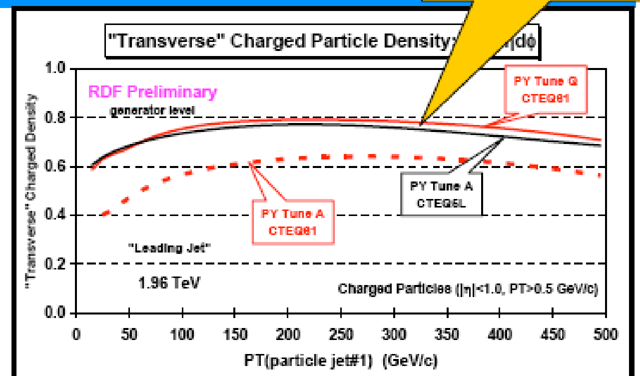
Rick’s tune

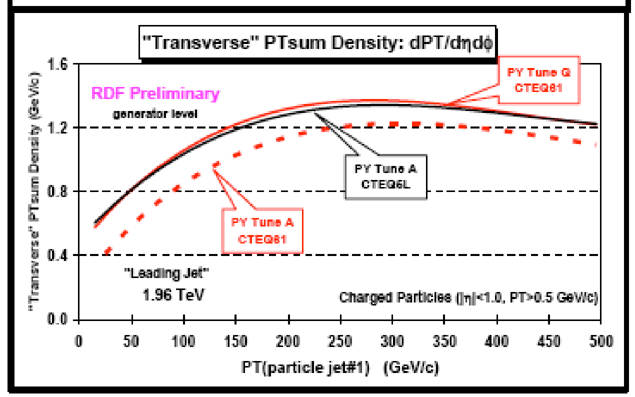
CTEQ6.1 Tune

PYTHIA 6.2 CTEQ6.1

	Parameter	Tune Q	Tune QW
UE Parameters	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
ISR Parameters	PARP(62)	1.0	1.25
	PARP(64)	1.0	0.2
	PARP(67)	4.0	4.0
Intrensic KT	MSTP(91)	1	1
	PARP(91)	1.0	2.1
	PARP(92)	5.0	15.0

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





TeV4LHC - Fermilab
October 20, 2005
Rick Field - Florida/CDF
Page 18

…will be discussed in detail in TeV4LHC writeup

Summary

- Now is the time to set up the SM tools and measurement program we need for the first few years of the LHC running
  - still great deal of preparation for early SM analyses needed
- Theoretical program to develop a broad range of tools for LHC
  - up to us (experimentalists) to make use of them/drive the development of what we need
- Program for SM benchmarks for LHC underway
  - www.pa.msu.edu/~huston/Les_Houches_2005/Les_Houches_SM.html
  - longer version of this talk available there
- Review paper available now
  - www.pa.msu.edu/~huston/seminars/Main.pdf
  - one of the authors has been honored in advance for his role on the paper

- Meeting on Thurs July 20 in 513-1-204 at 14:30 (Sky room on vrvs)
- Talks on
  - Tevatron
  - CMS/ATLAS experience/plans
  - a faster k_T algorithm
  - viewpoint from a theorist
  - hopefully useful discussions/conclusions
- Followup meeting on Monday July 24 at 15:00 in 40-4-C01 (also in Sky room)

Parton kinematics

- To serve as a handy "look-up" table, it's useful to define a parton-parton luminosity
  - this is from a contribution to Les Houches
- Equation 3 can be used to estimate the production rate for a hard scattering at the LHC

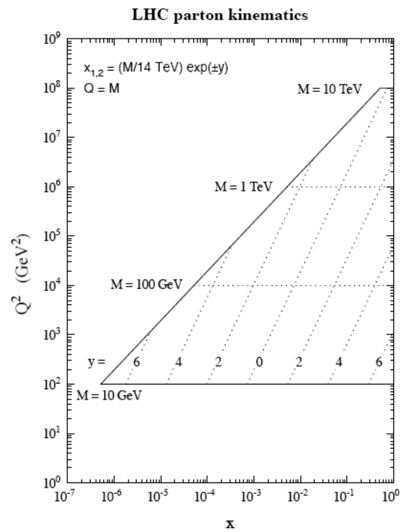


Fig. 1: Parton kinematics for the LHC.

$$\frac{dL_{ij}}{d\hat{s}\,dy} = \frac{1}{s}\frac{1}{1+\delta_{ij}}\left[f_i(x_1,\mu)f_j(x_2,\mu) + (1 \leftrightarrow 2)\right] . \quad (1)$$

The prefactor with the Kronecker delta avoids double-counting in case the partons are identical. The generic parton-model formula

$$\sigma = \sum_{i,j}\int_0^1 dx_1\, dx_2\, f_i(x_1,\mu)\, f_j(x_2,\mu)\, \hat{\sigma}_{ij} \quad (2)$$

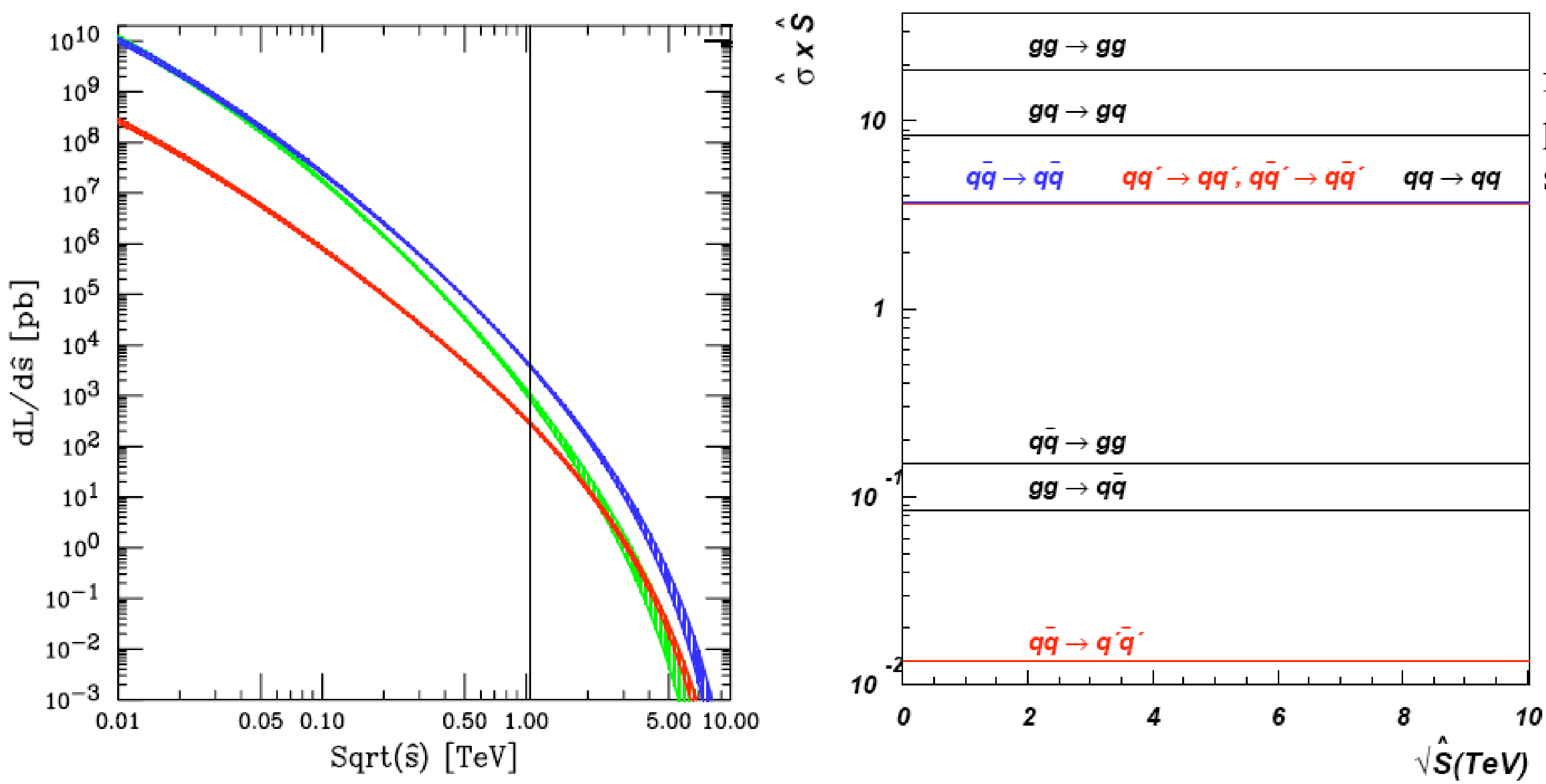
can then be written as

$$\sigma = \sum_{i,j}\int \left(\frac{d\hat{s}}{\hat{s}}\,dy\right)\left(\frac{dL_{ij}}{d\hat{s}\,dy}\right)\left(\hat{s}\,\hat{\sigma}_{ij}\right) . \quad (3)$$

Cross section estimates

$$\sigma = \frac{\Delta\hat{s}}{\hat{s}}\left(\frac{dL_{ij}}{d\hat{s}}\right)(\hat{s}\,\hat{\sigma}_{ij})$$

for the gluon pair production rate for $\hat{s}$=1 TeV and $\Delta\hat{s} = 0.01\hat{s}$, we have $\frac{dL_{gg}}{d\hat{s}} \simeq 10^3$ pb and $\hat{s}\,\hat{\sigma}_{gg} \simeq 20$ leading to $\sigma \simeq 200$ pb



for p_T=0.1* sqrt(s-hat)

Fig. 2: Left: luminosity $\left[\frac{1}{\hat{s}}\frac{dL_{ij}}{d\tau}\right]$ in pb integrated over y. Green=gg, Blue=$g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$, Red=$d\bar{d}+u\bar{u}+s\bar{s}+c\bar{c}+b\bar{b}+\bar{d}d+\bar{u}u+\bar{s}s+\bar{c}c+\bar{b}b$. Right: parton level cross sections $[\hat{s}\hat{\sigma}_{ij}]$ for various processes

Luminosities as a function of y

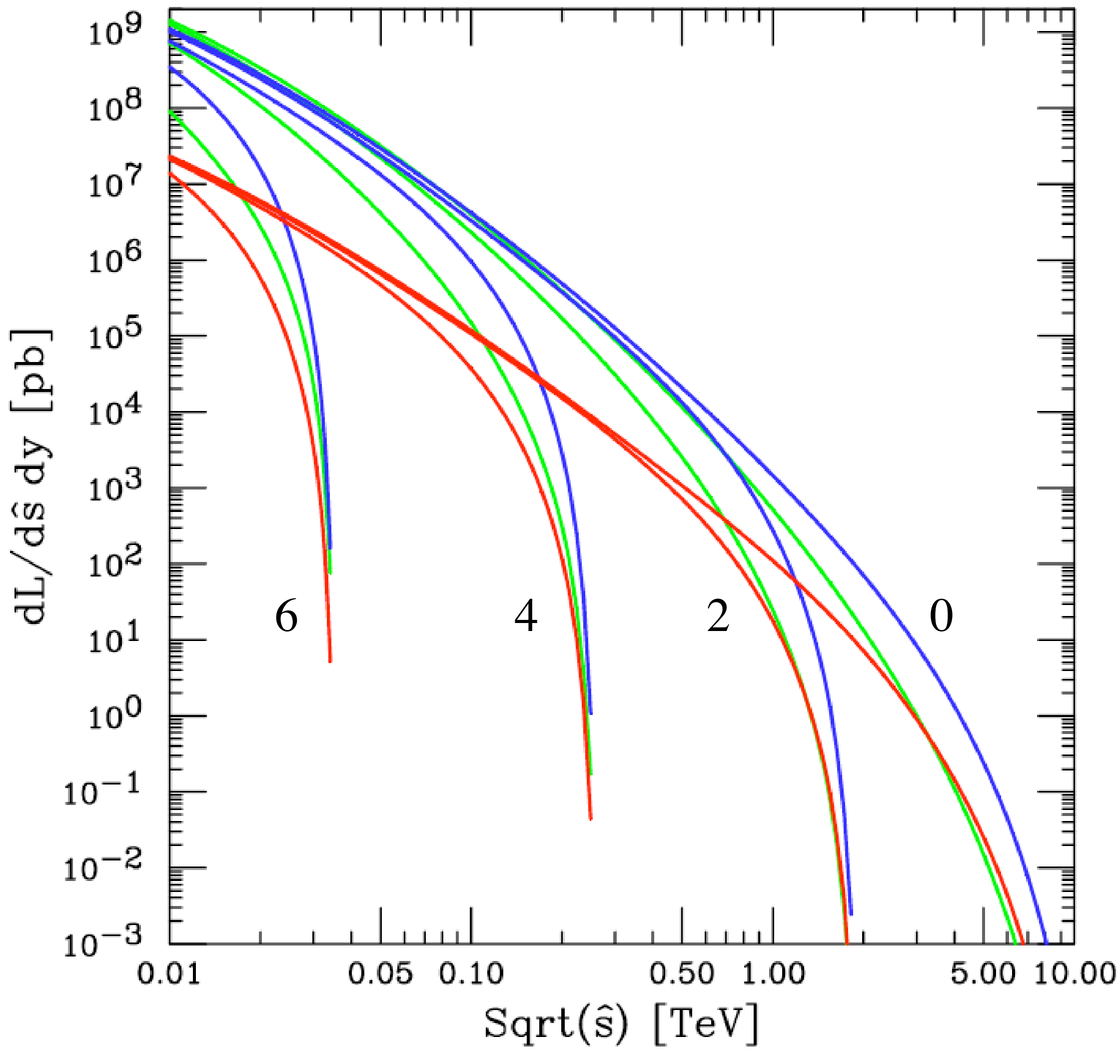


Fig. 3: dLuminosity/dy at $y = 0, 2, 4, 6$. Green=gg, Blue=$g(d+u+s+c+b) + g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b}) + (d+u+s+c+b)g + (\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$, Red=$d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$.

LHC to Tevatron pdf luminosities

- Processes that depend on qQ initial states (chargino pair production) have small enchancements
- Most backgrounds have gg or gq initial states and thus large enhancement factors (500 for W + 4 jets for example, which is primarily gq) at the LHC
- Luckily tT has a gg initial state as well as qQ so enhancement at the LHC is a factor of 100
  - but increased W + jets background means that a higher jet cut is necessary at the LHC

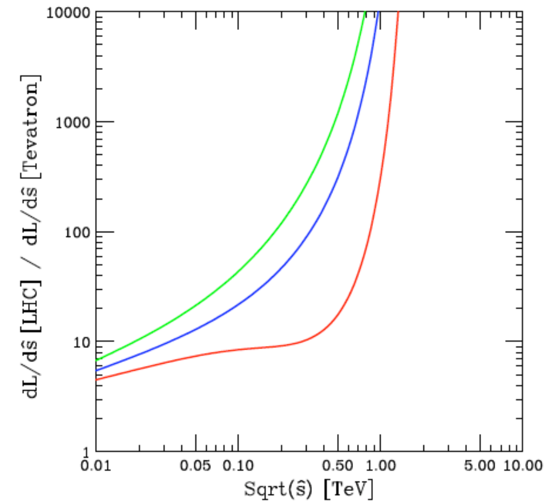


Figure 11. The ratio of parton-parton luminosity $\left[\frac{1}{\hat{s}}\frac{dL_{ij}}{d\tau}\right]$ in pb integrated over y at the LHC and Tevatron. Green=gg (top), Blue=$g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$ (middle), Red=$d\bar{d}+u\bar{u}+s\bar{s}+c\bar{c}+b\bar{b}+\bar{d}d+\bar{u}u+\bar{s}s+\bar{c}c+\bar{b}b$ (bottom).

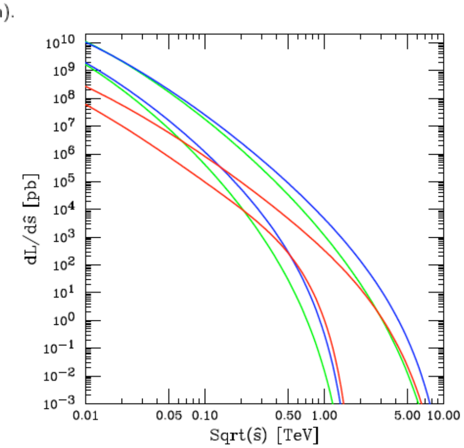


Figure 10. The parton-parton luminosity $\left[\frac{1}{\hat{s}}\frac{dL_{ij}}{d\tau}\right]$ in pb integrated over y. Green=gg, Blue=$g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$, Red=$d\bar{d}+u\bar{u}+s\bar{s}+c\bar{c}+b\bar{b}+\bar{d}d+\bar{u}u+\bar{s}s+\bar{c}c+\bar{b}b$. The top family of curves are for the LHC and the bottom for the Tevatron.

gg luminosity uncertainties

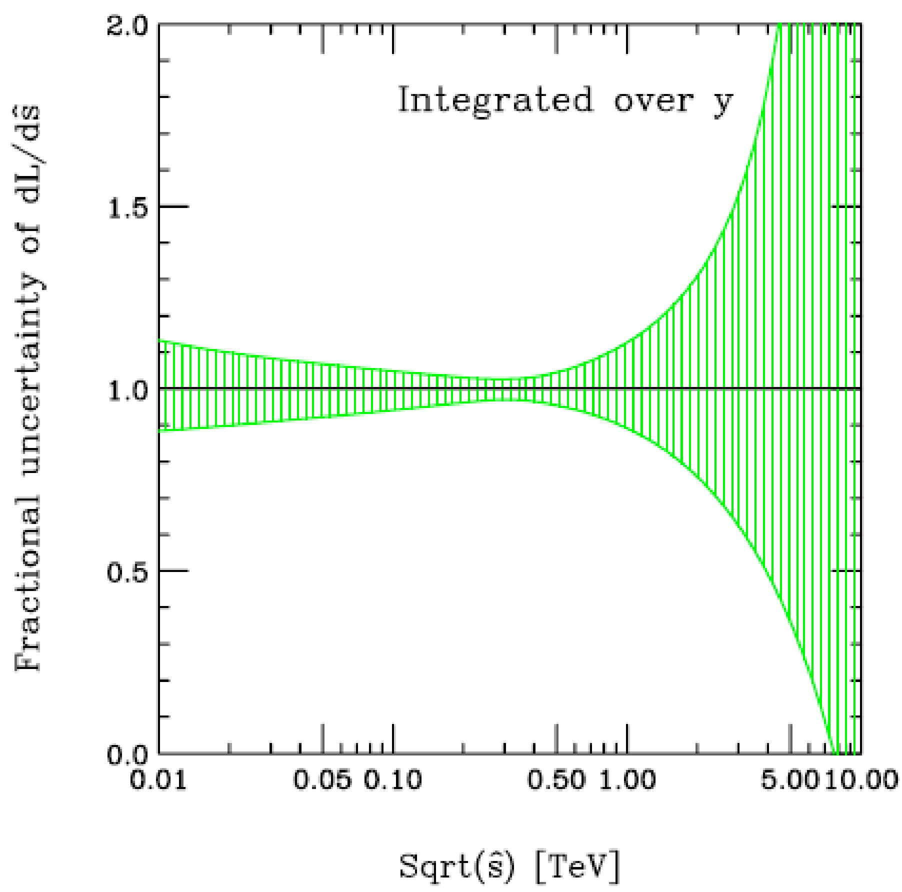


Fig. 4: Fractional uncertainty of gg luminosity integrated over y.

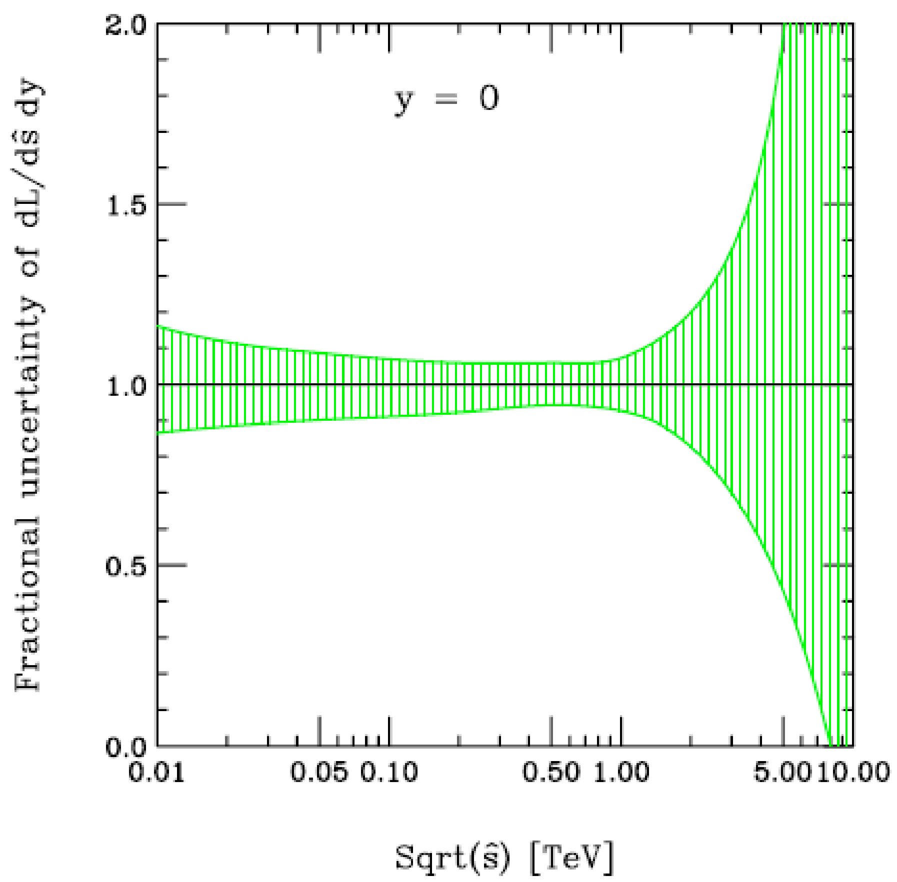
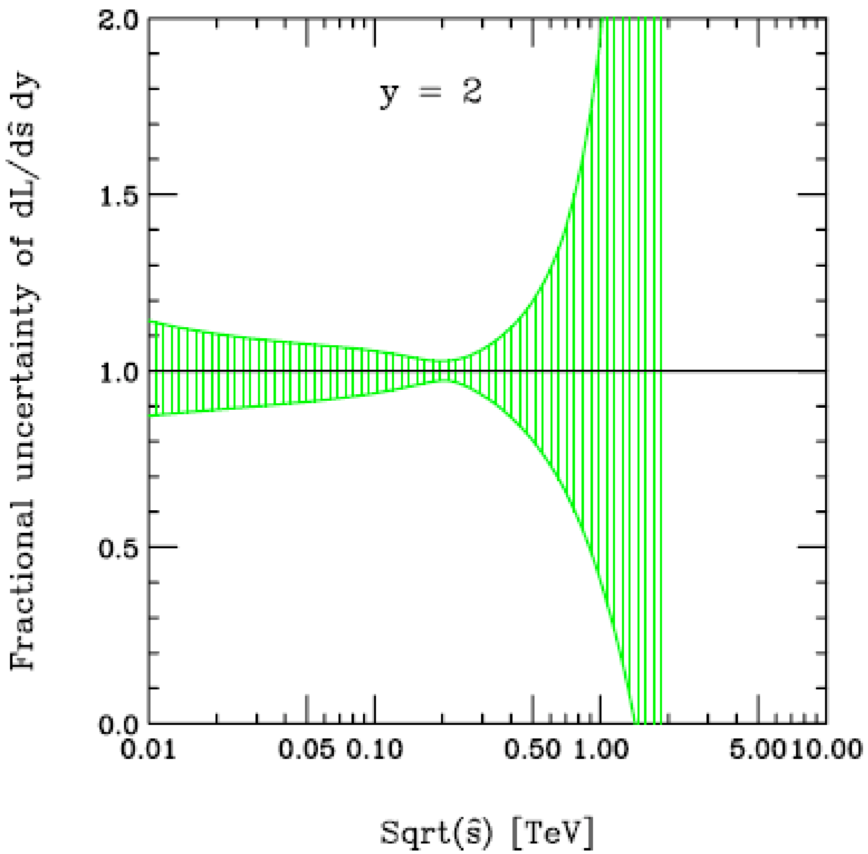
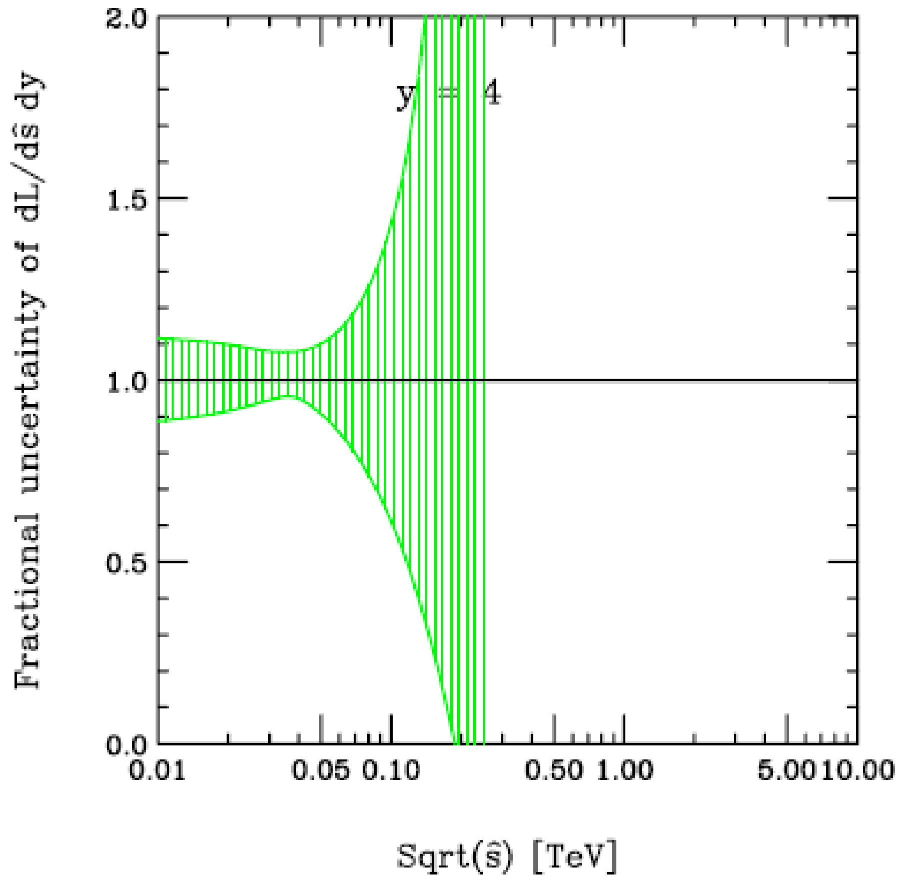


Fig. 5: Fractional uncertainty of gg luminosity at $y = 0$.

gg luminosity uncertainties





gq luminosity uncertainties

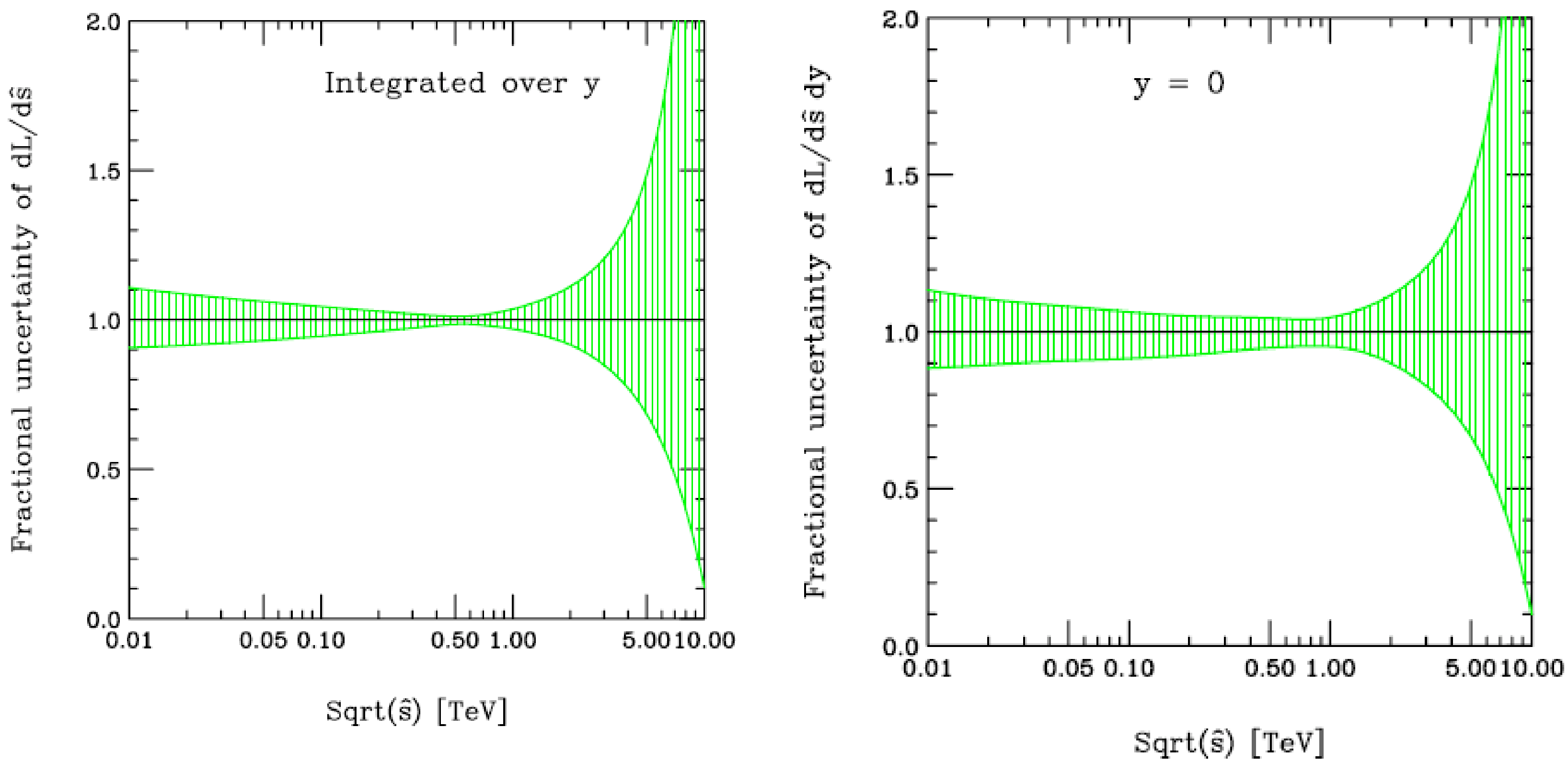
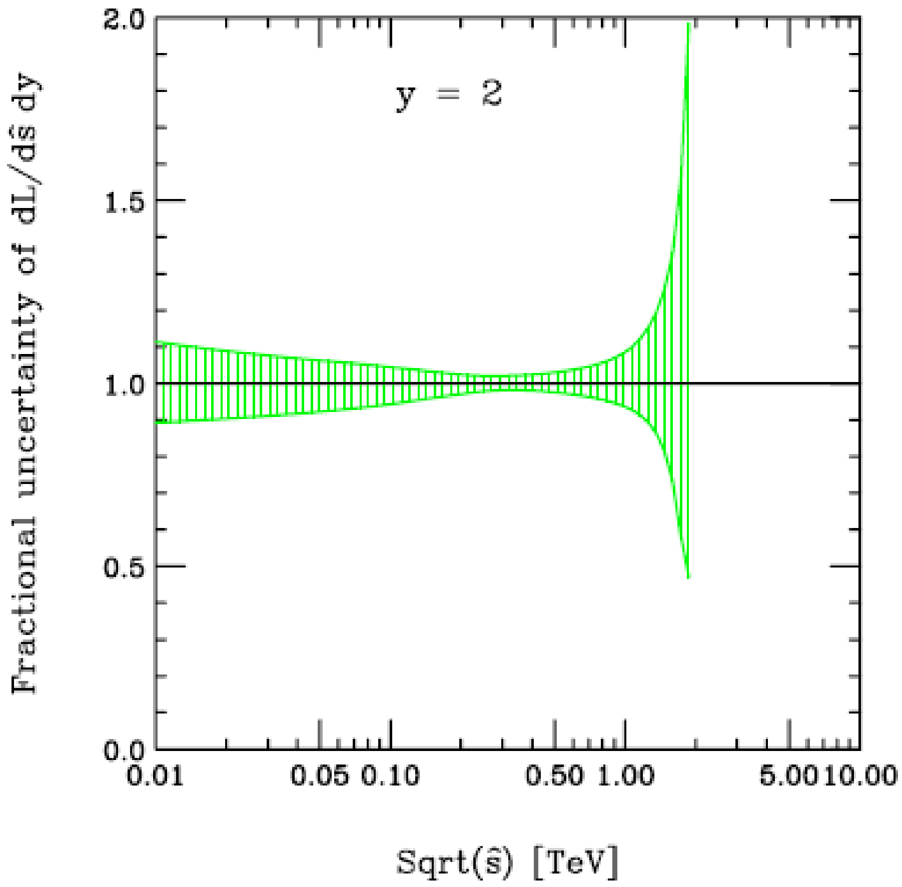
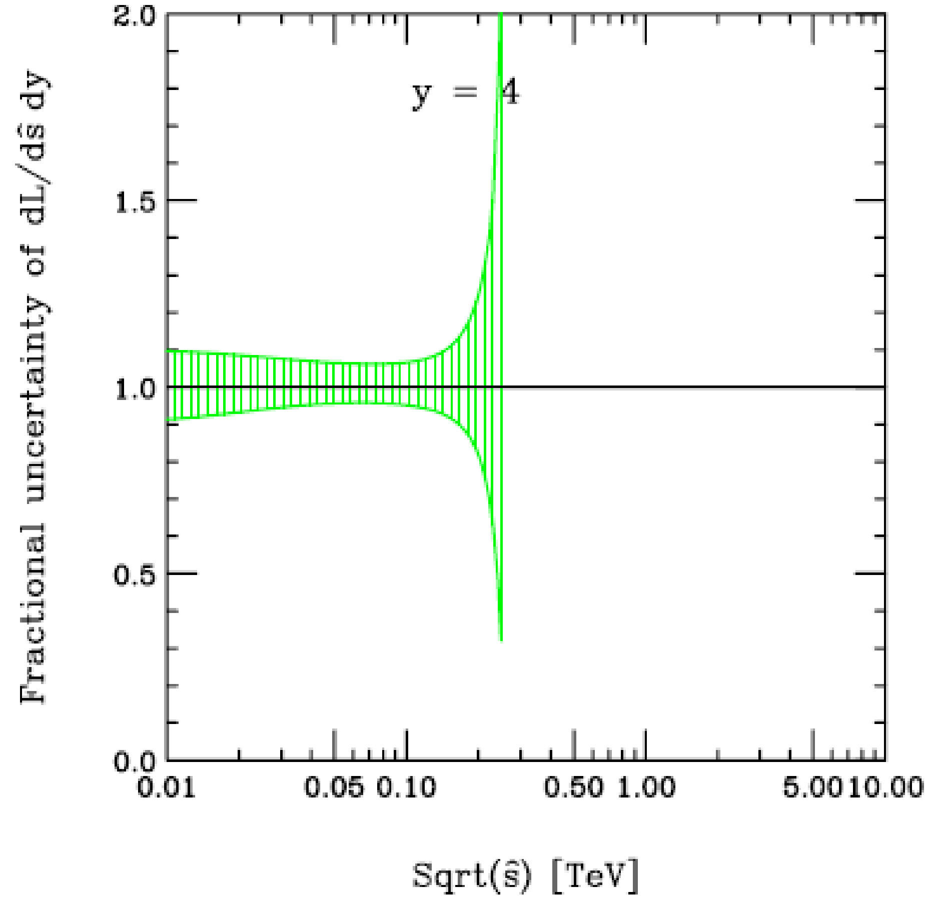


Fig. 6: Fractional uncertainty for Luminosity integrated over y for $g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$,

gq luminosity uncertainties





qQ luminosity uncertainties

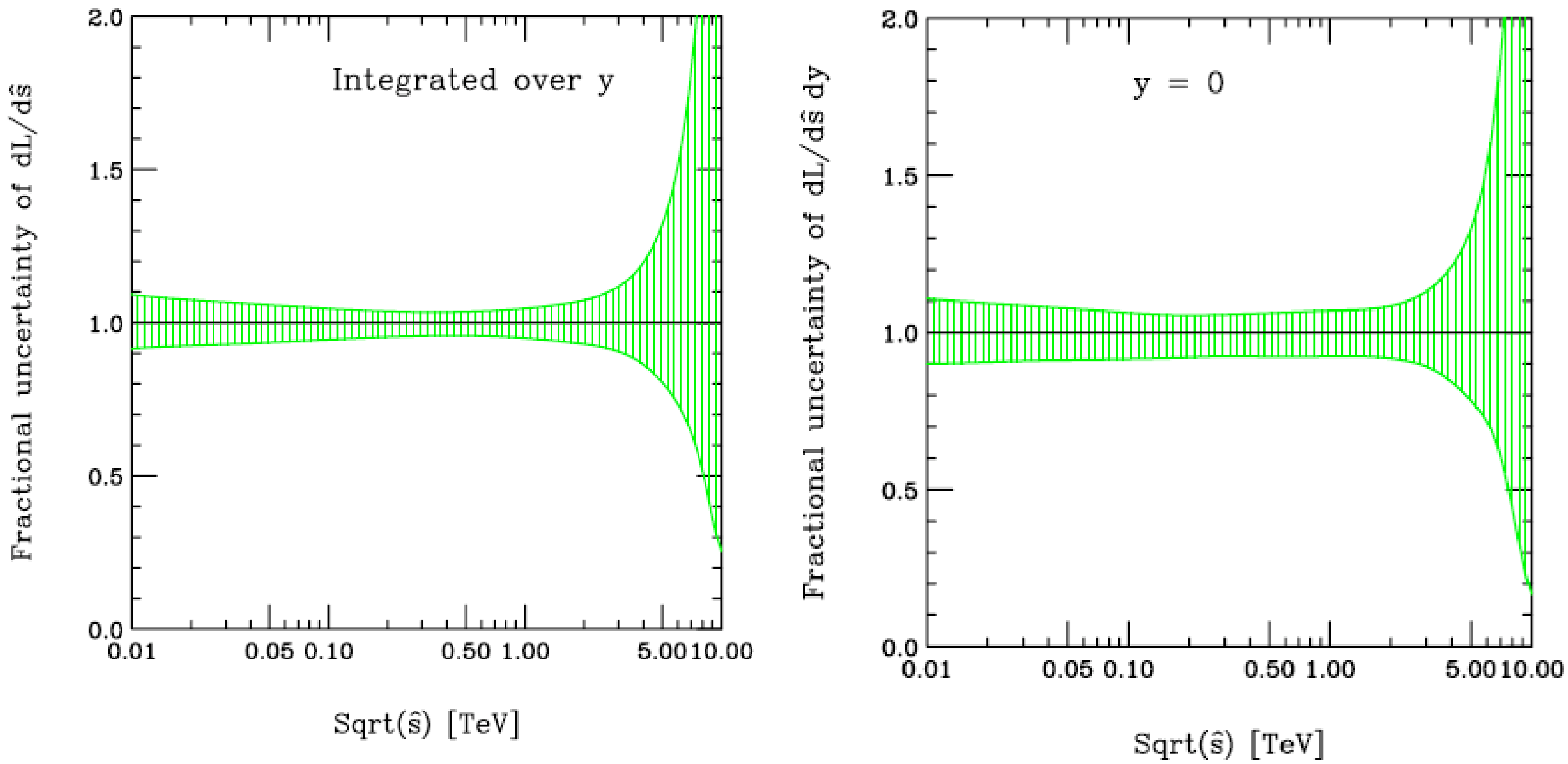
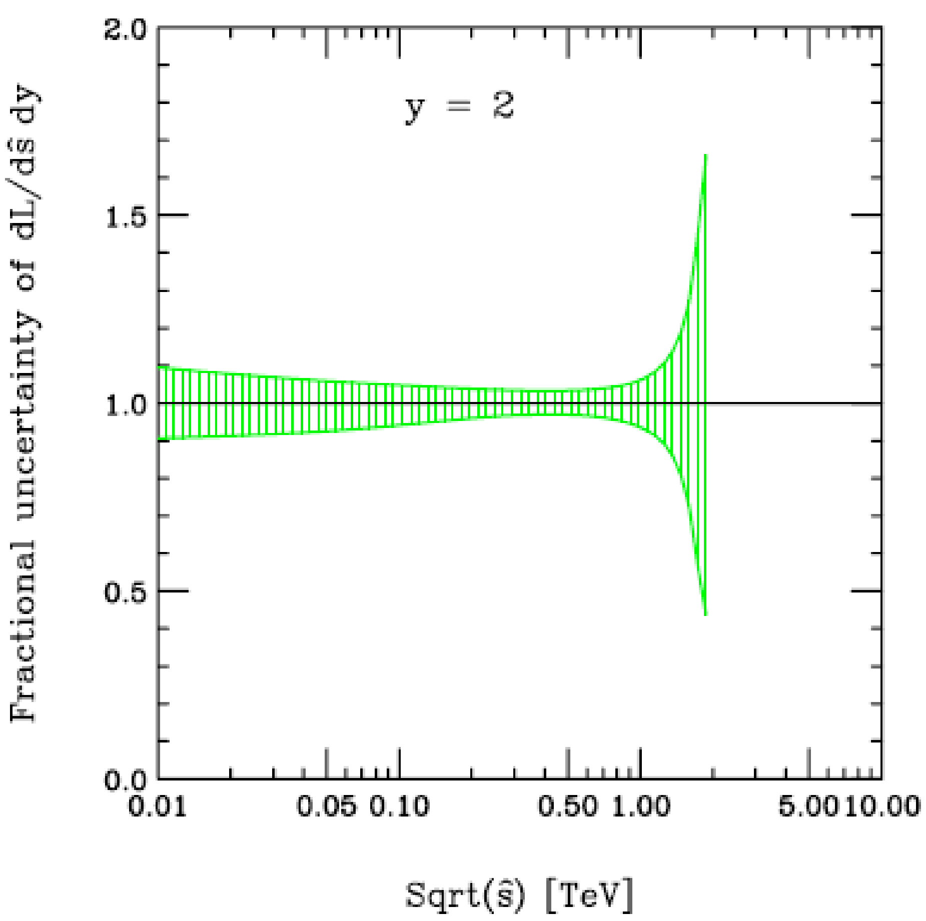
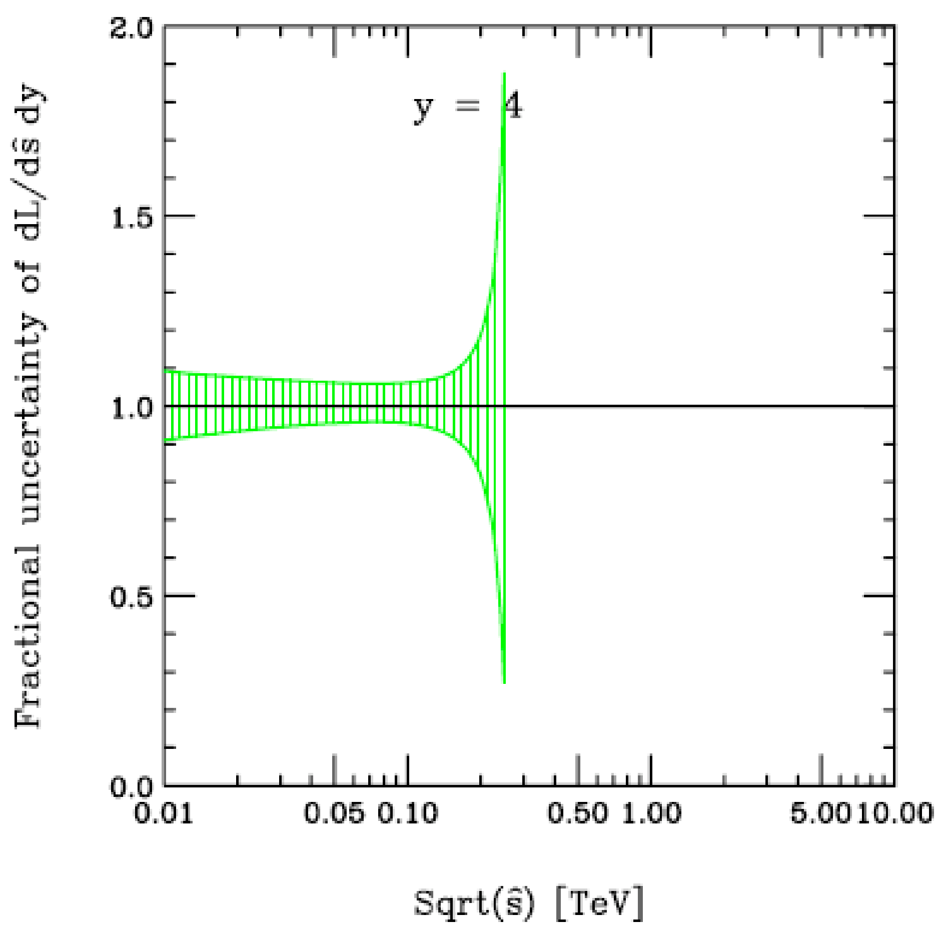


Fig. 7: Fractional uncertainty for Luminosity integrated over y for $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$.

qQ luminosity uncertainties





The “maligned” experimenter’s wishlist

- Missing many needed NLO computations Campbell

An experimenter’s wishlist

- Hadron collider cross-sections one would like to know at NLO

Run II Monte Carlo Workshop, April 2001

Single boson	Diboson	Triboson	Heavy flavour
$W + \leq 5j$	$WW + \leq 5j$	$WWW + \leq 3j$	$t\bar{t} + \leq 3j$
$W + b\bar{b} + \leq 3j$	$WW + b\bar{b} + \leq 3j$	$WWW + b\bar{b} + \leq 3j$	$t\bar{t} + \gamma + \leq 2j$
$W + c\bar{c} + \leq 3j$	$WW + c\bar{c} + \leq 3j$	$WWW + \gamma\gamma + \leq 3j$	$t\bar{t} + W + \leq 2j$
$Z + \leq 5j$	$ZZ + \leq 5j$	$Z\gamma\gamma + \leq 3j$	$t\bar{t} + Z + \leq 2j$
$Z + b\bar{b} + \leq 3j$	$ZZ + b\bar{b} + \leq 3j$	$WZZ + \leq 3j$	$t\bar{t} + H + \leq 2j$
$Z + c\bar{c} + \leq 3j$	$ZZ + c\bar{c} + \leq 3j$	$ZZZ + \leq 3j$	$t\bar{b} + \leq 2j$
$\gamma + \leq 5j$	$\gamma\gamma + \leq 5j$		$b\bar{b} + \leq 3j$
$\gamma + b\bar{b} + \leq 3j$	$\gamma\gamma + b\bar{b} + \leq 3j$		
$\gamma + c\bar{c} + \leq 3j$	$\gamma\gamma + c\bar{c} + \leq 3j$		
	$WZ + \leq 5j$		
	$WZ + b\bar{b} + \leq 3j$		
	$WZ + c\bar{c} + \leq 3j$		
	$W\gamma + \leq 3j$		
	$Z\gamma + \leq 3j$		

NLO calculation priority list from Les Houches 2005: theory benchmarks

- Note have to specify how inclusive final state is
  - what cuts will be made?
  - how important is b mass for the observables?
- How uncertain is the final state?
  - what does scale uncertainty look like at tree level?
  - new processes coming in at NLO?
- Some information may be available from current processes
  - pp->tT j may tell us something about pp->tTbB?
    - j=g->bB
  - CKKW may tell us something about higher multiplicity final states

can we develop rules-of-thumb about size of HO corrections?

1. pp->WW jet
2. pp->H + 2 jets now complete
   1. background to VBF production of Higgs
3. pp->tT bB
   1. background to tTH
4. pp->tT + 2 jets
   1. background to tTH
5. pp->WWbB
6. pp->V V + 2 jets
   1. background to WW->H->WW
7. pp->V + 3 jets
   1. beneral background to new physics
8. pp->V V V
   1. background to SUSY trilepton

Are there any other cross sections that should be on this list?

More of benchmark webpages

- what are the uncertainties? what are the limitations of the theoretical predictions?
  - indicate scale dependence of cross sections as well as pdf uncertainties
  - how do NLO predictions differ from LO ones?

Table 1. K-factors for various processes at the Tevatron and the LHC, calculated using a selection of input parameters. In all cases, the CTEQ6M PDF set is used at NLO. $\mathcal{K}$ uses the CTEQ6L1 set at leading order, whilst $\mathcal{K}'$ uses the same set, CTEQ6M, as at NLO. Jets satisfy the requirements $p_T > 15$ GeV and $|\eta| < 2.5$ (5.0) at the Tevatron (LHC). In the $W + 2$ jet process the jets are separated by $\Delta R > 0.52$, whilst the weak boson fusion (WBF) calculations are performed for a Higgs of mass 120 GeV.

	Typical scales		Tevatron K-factor			LHC K-factor		
Process	μ_0	μ_1	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$
W	m_W	$2m_W$	1.33	1.31	1.21	1.15	1.05	1.15
$W + 1$ jet	m_W	$\langle p_T^{\text{jet}} \rangle$	1.42	1.20	1.43	1.21	1.32	1.42
$W + 2$ jets	m_W	$\langle p_T^{\text{jet}} \rangle$	1.16	0.91	1.29	0.89	0.88	1.10
$t\bar{t}$	m_t	$2m_t$	1.08	1.31	1.24	1.40	1.59	1.48
$b\bar{b}$	m_b	$2m_b$	1.20	1.21	2.10	0.98	0.84	2.51
Higgs via WBF	m_H	$\langle p_T^{\text{jet}} \rangle$	1.07	0.97	1.07	1.23	1.34	1.09

from review paper; in process of adding more processes; any favorites missing?

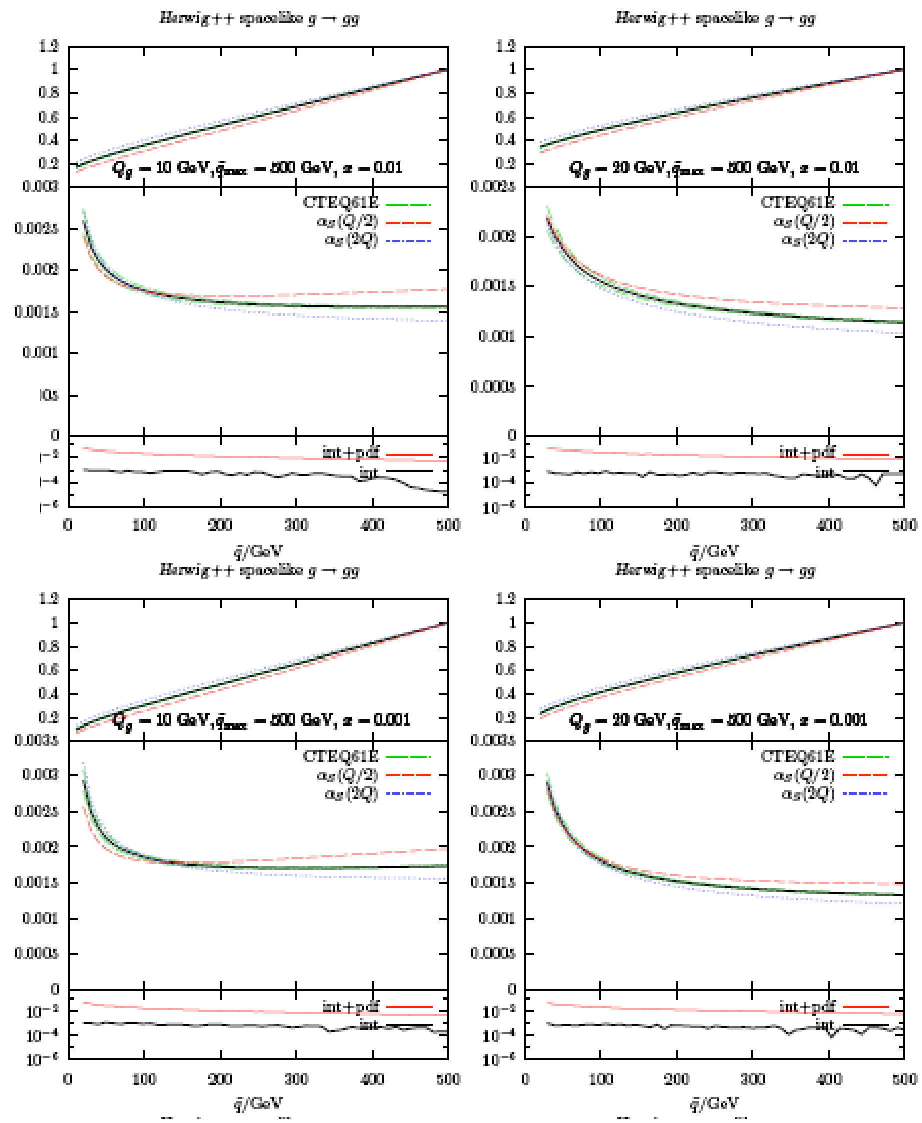
- to what extent are the predictions validated by current data?
- what measurements could be made at the Tevatron and HERA before then to add further information?

Back to Sudakov form factors

- The Sudakov form factor gives the probability for a parton not to radiate, with a given resolution scale, when evolving from a large scale down to a small scale

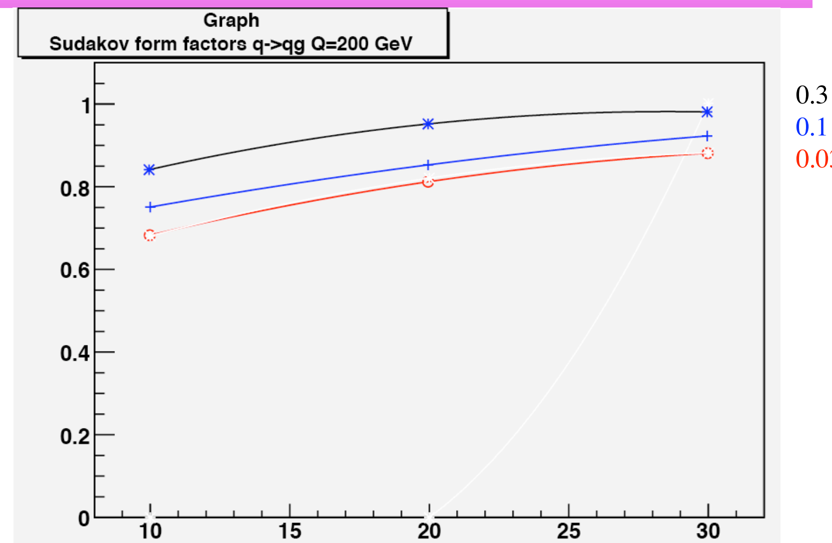
$$\Delta(t) \equiv \exp\left[-\int_{t_0}^{t}\frac{dt'}{t'}\int\frac{dz}{z}\frac{\alpha_s}{2\pi}P(z)\,\frac{f(x/z,t)}{f(x,t)}\right]$$

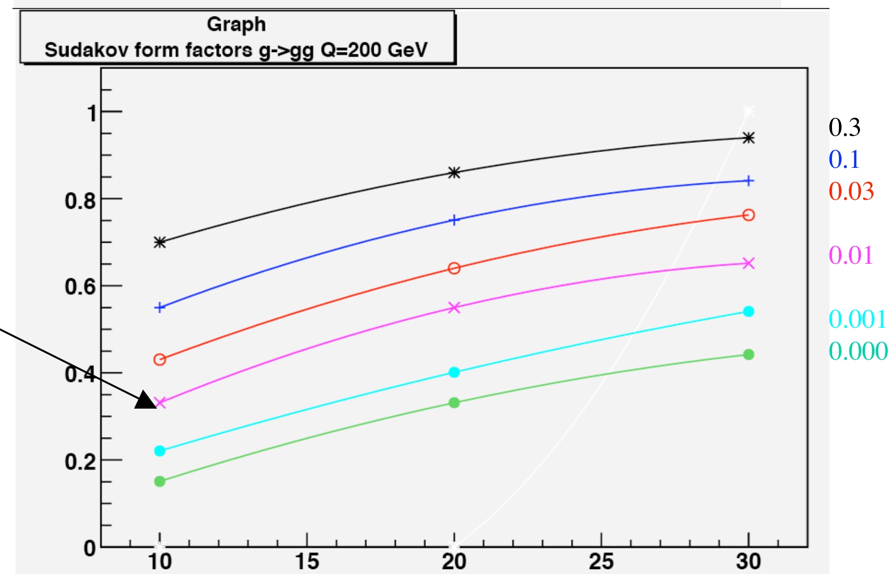
- Probability of emission increases with color charge (gluon vs quark), with larger max scale, with decreasing scale for a resolvable emission and with decreasing parton x



Sudakov form factors

- Curves from top to bottom correspond to x values of 0.3,0.1, 0.03, 0.01, 0.001, 0.0001
- Sudakov form factors for q->qg for x<0.03 are similar to form factor for x=0.03 (and so are not shown)
- Sudakov form factors for g->gg continue to drop with decreasing x
  - g->gg splitting function P(z) has singularities both as z->0 and as z->1
  - q->qg has only z->1 singularity
- For example, probability for an initial state gluon of x=0.01 not to emit a gluon of >=10 GeV when starting from an initial scale of 200 GeV is ~35%, i.e. there is a 65% probability for such an emission

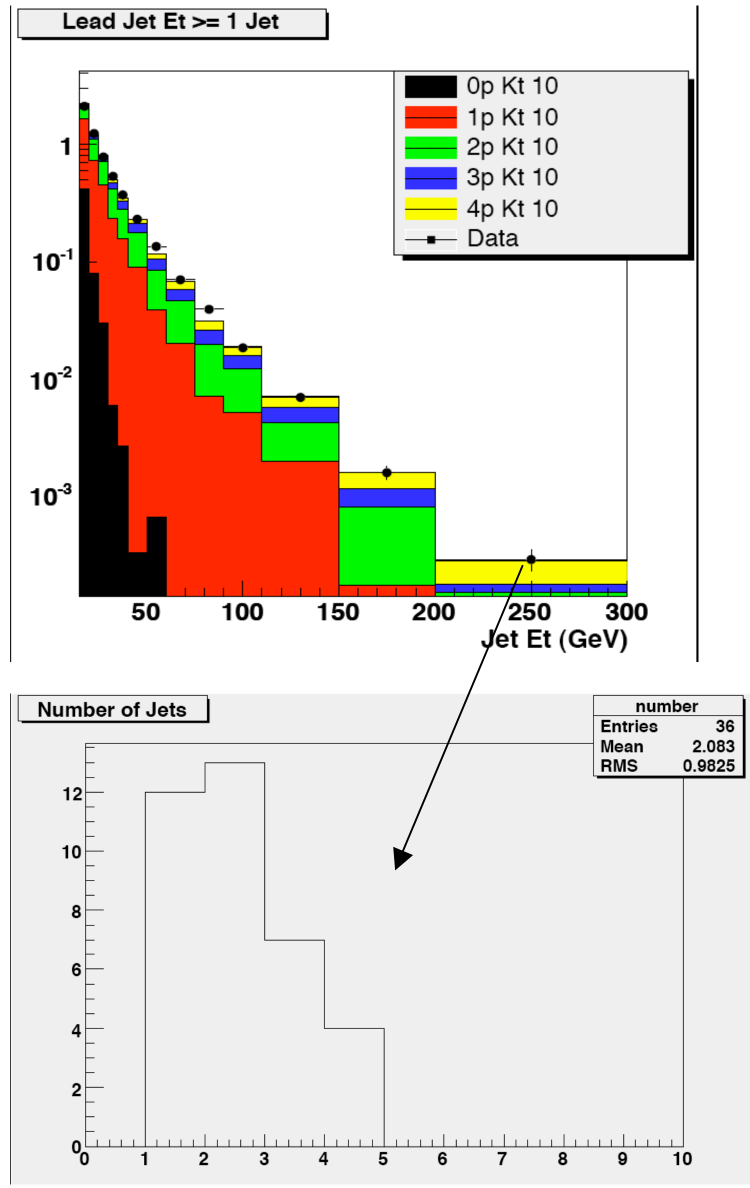




Resolution scale -> ~p_T of gluon

W + jet(s)

- Consider W + jet at the Tevatron where the jet has a high transverse momentum
- In the CKKW formalism, most of these events will have been produced by W + n parton configurations where n>1
- …or in other words, there is a Sudakov suppression of final states with just the lead jet and no additional (softer) jets
    - I can use the types of curves on the previous page to estimate the rate for ISR jets
    - note I can also get extra jets from final state radiation



Sudakov form factors

- If I go to small x, or high scale or a gluon initial state, then probability of a ISR gluon emission approaches unity
- The above sentence basically describes the LHC

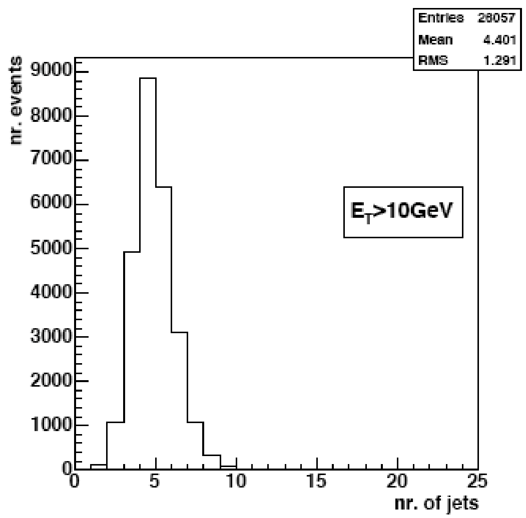
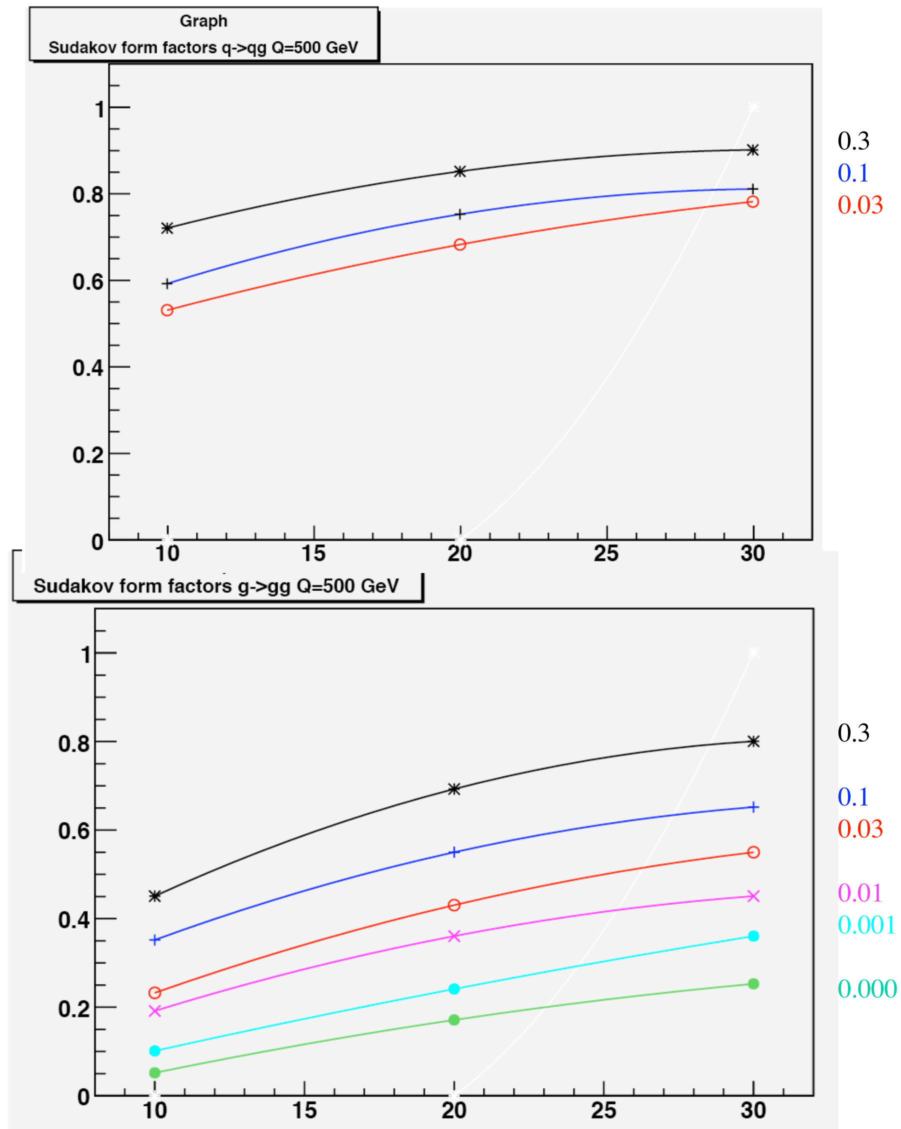


Figure 75. The jet multiplicity in $t\bar{t}$ events with a lepton + jets final state at the LHC. A cut of 10 GeV has been applied to the jets.



More of benchmark webpages

- what are the uncertainties? what are the limitations of the theoretical predictions?
  - indicate scale dependence of cross sections as well as pdf uncertainties
  - how do NLO predictions differ from LO ones?

Table 1. K-factors for various processes at the Tevatron and the LHC, calculated using a selection of input parameters. In all cases, the CTEQ6M PDF set is used at NLO. $\mathcal{K}$ uses the CTEQ6L1 set at leading order, whilst $\mathcal{K}'$ uses the same set, CTEQ6M, as at NLO. Jets satisfy the requirements $p_T > 15$ GeV and $|\eta| < 2.5$ (5.0) at the Tevatron (LHC). In the $W+2$ jet process the jets are separated by $\Delta R > 0.52$, whilst the weak boson fusion (WBF) calculations are performed for a Higgs of mass 120 GeV.

	Typical scales		Tevatron K-factor			LHC K-factor		
Process	μ_0	μ_1	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$	$\mathcal{K}(\mu_0)$	$\mathcal{K}(\mu_1)$	$\mathcal{K}'(\mu_0)$
W	m_W	$2m_W$	1.33	1.31	1.21	1.15	1.05	1.15
$W+1$ jet	m_W	$\langle p_T^{\text{jet}} \rangle$	1.42	1.20	1.43	1.21	1.32	1.42
$W+2$ jets	m_W	$\langle p_T^{\text{jet}} \rangle$	1.16	0.91	1.29	0.89	0.88	1.10
$t\bar{t}$	m_t	$2m_t$	1.08	1.31	1.24	1.40	1.59	1.48
$b\bar{b}$	m_b	$2m_b$	1.20	1.21	2.10	0.98	0.84	2.51
Higgs via WBF	m_H	$\langle p_T^{\text{jet}} \rangle$	1.07	0.97	1.07	1.23	1.34	1.09

from review paper; in process of adding more processes; any favorites missing?

- to what extent are the predictions validated by current data?
- what measurements could be made at the Tevatron and HERA before then to add further information?

More…

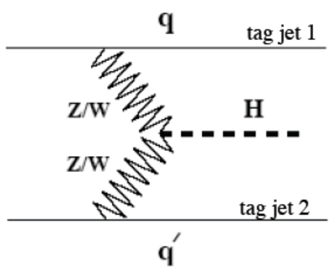
- technical benchmarks
  - jet algorithm comparisons
    - midpoint vs simple iterative cone vs kT
      - top studies at the LHC
      - an interesting data event at the Tevatron that examines different algorithms
    - Building Better Cone Jet Algorithms
      - one of the key aspects for a jet algorithm 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?)

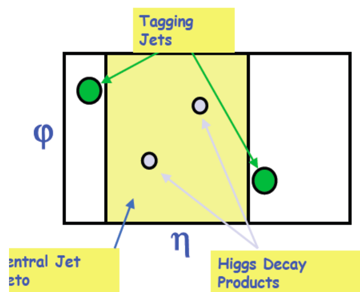
…plus more benchmarks that I have no time to discuss

- variation of ISR/FSR a la CDF (study performed by Un-Ki Yang)
  - low ISR/high ISR
  - FSR
  - power showers versus wimpy showers a la Peter Skands
  - number of additional jets expected due to ISR effects (see also Sudakov form factors)
  - impact on top analyses
  - effect on benchmarks such as Drell-Yan and diphoton production
    - goal is to produce a range for ISR predictions that can then be compared at the LHC to Drell-Yan and to diphoton data
- Sudakov form factor compilation
  - probability for emission of 10, 20, 30 GeV gluon in initial state for hard scales of 100, 200, 500, 1000, 5000 GeV for quark and gluon initial legs
  - see for example, similar plots for quarks and gluons for the Tevatron from Stefan Gieseke
- predictions for W/Z/Higgs p_T and rapidity at the LHC
  - compare ResBos(-A), joint-resummation and Berger-Qiu for W and Z

W + jets at the Tevatron and LHC

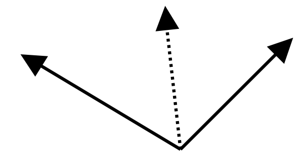
- One of the most promising channels for Higgs production at the LHC is through WW fusion



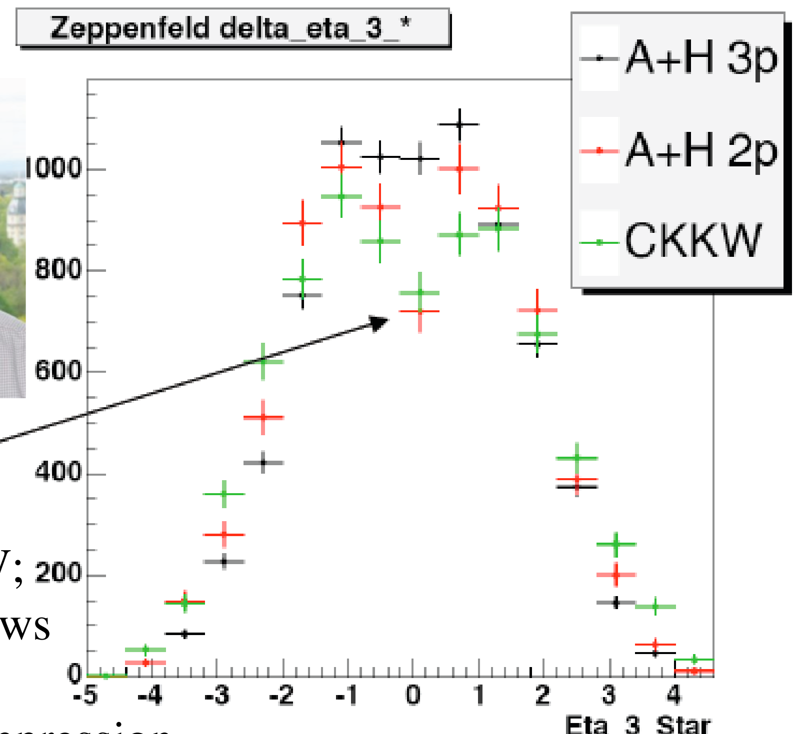


- Plan is to veto on backgrounds from Zjj by requiring no central jets (between tagging jets)
- Look at W + jets at the Tevatron as a way of testing central jet rate and distribution
  - analysis in progress; result will be absolute cross sections
- Extrapolate to LHC using MCFM and CKKW
  - study in progress with Bruce Mellado and Steve Mrenna

2 tagging jets F/B, $\Delta\eta>2$; look at relative rapidity of 3rd jet



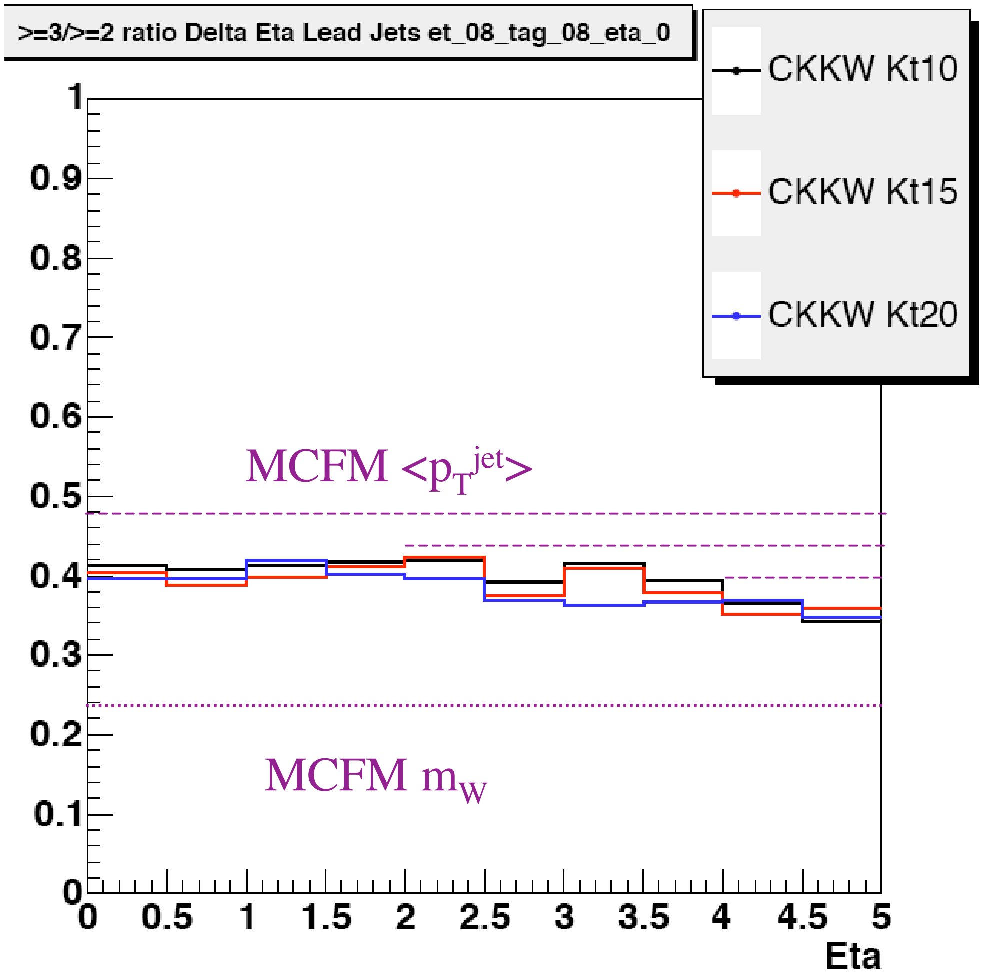
Tag jets > 8 GeV/c; 3rd jet > 8 GeV/c



note central dip with CKKW; CKKW knows about Sudakov suppression for central jet emission (so does data)

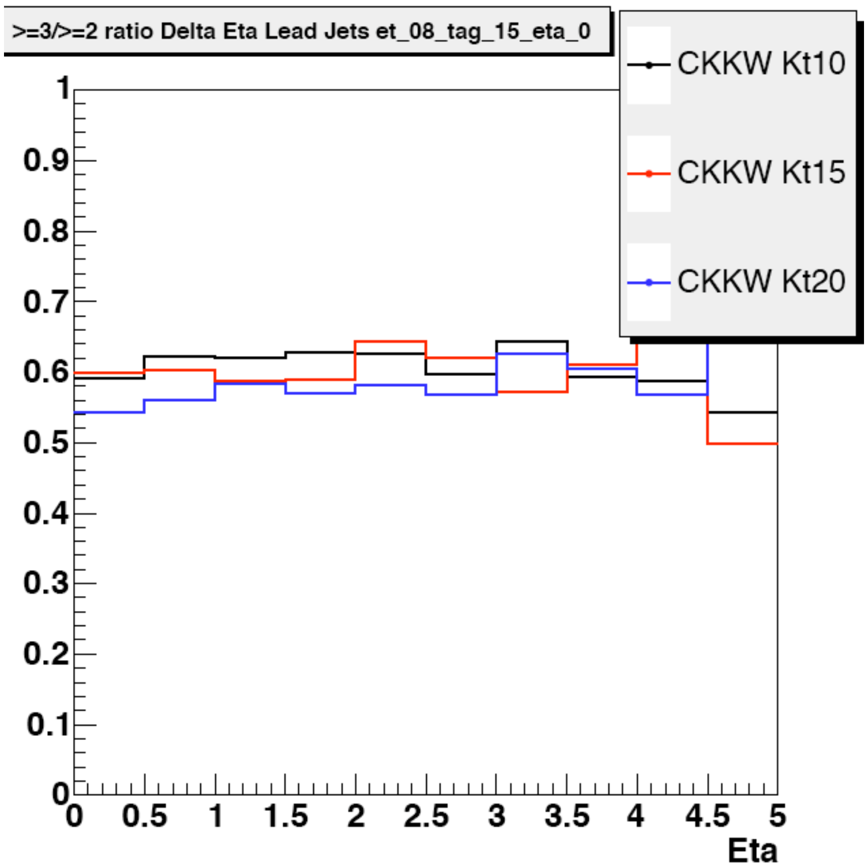
CKKW matching variation

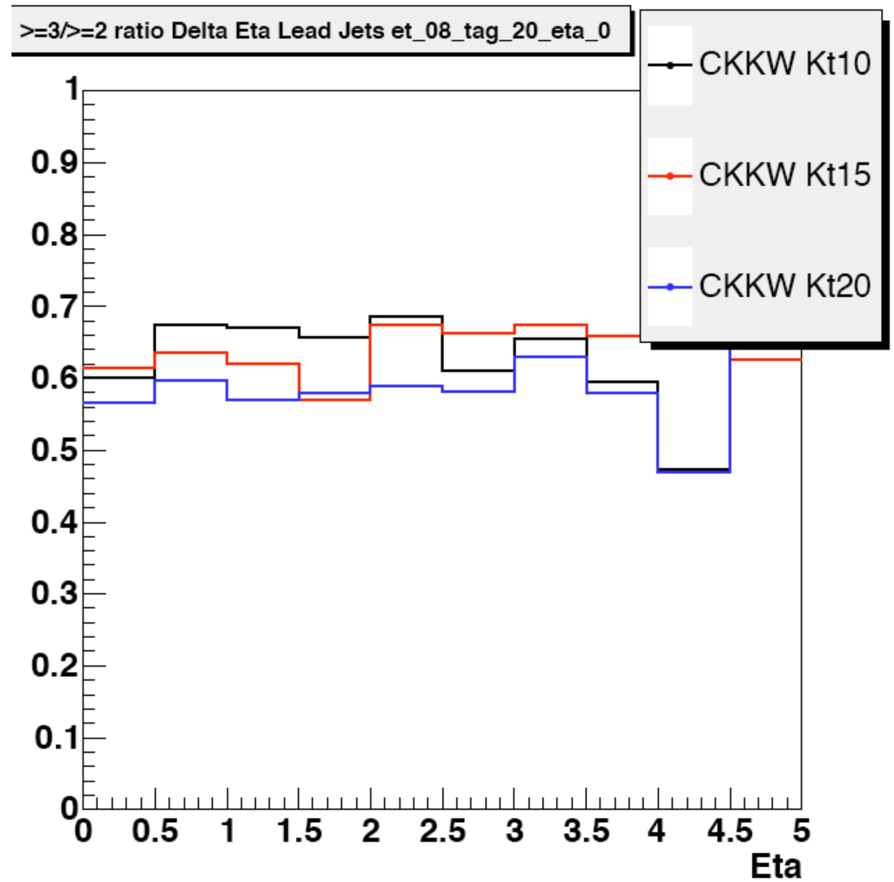
- Look at probability for 3rd jet to be emitted as a function of the rapidity separation of the tagging jets
- Relatively flat probability (although slightly decreasing at low $\Delta\eta$ due to kinematic suppression), stable with CKKW scale
- Bracketed by two predictions for MCFM using m_W and $\langle p_T^{jet} \rangle$ as scales
- Data to be blessed soon



CKKW matching variation

- Increase cut on tagging jet to 15/20 GeV/c
- Probability of jet emission increases





W + jets at LHC

- Look at probability for 3rd jet to be emitted as a function of the rapidity separation of the tagging jets
- At LHC, ratio (p_T^{jet}>15 GeV/c) much higher than at Tevatron
- CKKW comparison underway