# Partons and Jets at the LHC

Davison E. Soper
University of Oregon

Kanpur, November 2002

## Overview

- How to find new physics at the LHC.
  - Direct searches
  - Jet cross sections as a probe
- Jet cross section and its errors
- Higher order calculations
  - $g - 2$ for the muon
  - Progress toward NNLO for generic observables
  - Perturbative summations
- Parton distributions with real error estimates
- Jet definitions

## How will we look for new physics at the LHC?

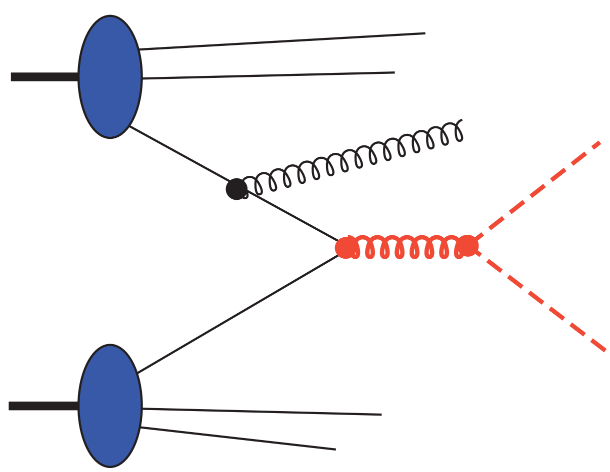
### Look directly.

*E.g.*

SUSY $\Rightarrow$ squarks $\Rightarrow p + p \to$ squark + antisquark + $X$.

We use

$$d\sigma \approx \sum_{a,b} \int_0^1 d\xi_A \int_0^1 d\xi_B \; f_{a/A}(\xi_A, \mu) \; f_{b/B}(\xi_B, \mu) \; d\hat{\sigma}^{ab}(\mu).$$

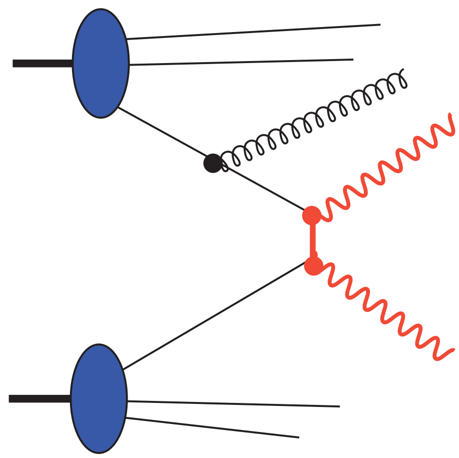


- We need the parton distribution functions $f_{a/A}(\xi, \mu)$.
- We need the hard scattering cross sections $d\hat{\sigma}^{ab}(\mu)$.
  - $d\hat{\sigma}$ has been calculated at next-to-leading order (NLO) for lots of processes of interest.
  - The calculation involves a subtraction in $d\hat{\sigma}$ to allow for the emitted gluon being in $d\hat{\sigma}^{ab}(\mu)$.
  - These are QCD calculations.

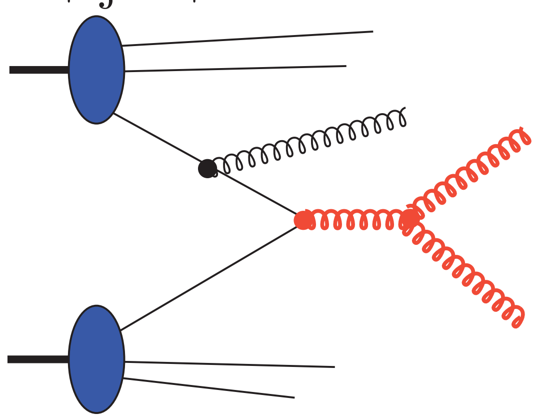
## Look indirectly.

*E.g.*

S.M. $\Rightarrow p + p \to W^+ + W^- + X$.



S.M. $\Rightarrow p + p \to \text{jet} + \text{jet} + X$.



We still use

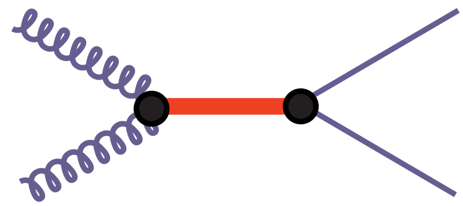
$$d\sigma \approx \sum_{a,b} \int_0^1 d\xi_A \int_0^1 d\xi_B \; f_{a/A}(\xi_A, \mu) \; f_{b/B}(\xi_B, \mu) \; d\hat{\sigma}^{ab}(\mu).$$

If anything goes wrong it must be new physics (if the discrepancy is outside the errors).

## Jet cross sections and new physics signatures

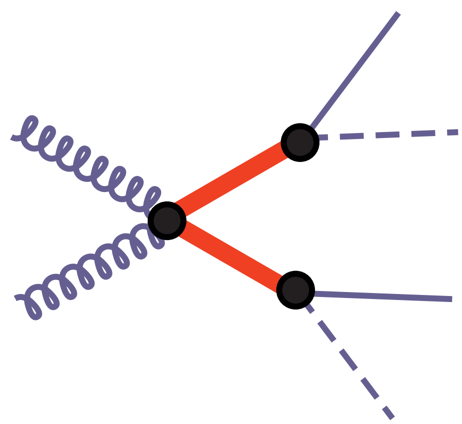
• Suppose there is a new interaction at a scale $\Lambda$.

If $\Lambda < E_T^{\max}$:



a) New particle with mass $M$ that decays to two jets.

* One jet inclusive cross section
  Look for threshold effect at $E_T = M/2$.
* Two jet inclusive cross section
  Look for resonance structure at $M_{\mathrm{Jet-Jet}} = M$.



b) New particle with mass $M$ that decays to 3 jets, 4 jets, 2 jets + invisible particles, 2 jets + leptons; new particles that are produced in pairs...

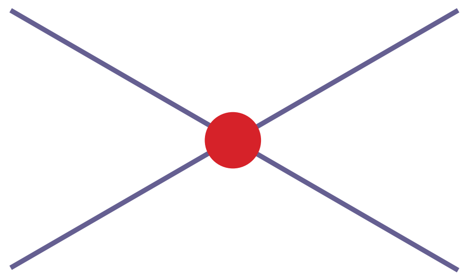
* In principle, this contributes to one and two jet inclusive cross sections, but background $\gg$ signal.

$\Rightarrow$ look for this directly.

If $\Lambda > E_T^{\max}$:

Get new terms in the effective lagrangian like

$$\Delta\mathcal{L} = \frac{g'}{\Lambda^2}\,(\bar{\psi}\psi)^2$$



Then the one jet inclusive cross section is changed:

$$\frac{d\sigma_{\rm Jet}}{dE_T} \approx \left(\frac{d\sigma_{\rm Jet}}{dE_T}\right)_0 \times \left[1 + ({\rm const.})\,\frac{g'}{\alpha_s}\,\frac{E_T^2}{\Lambda^2}\right]$$

Then the two jet inclusive cross section is also changed.

## Extra dimensions

What if space has more than three dimensions, with the extra dimensions rolled into a little ball of size $R$?

Then a quark or gluon is pointlike when viewed by a probe with wavelength $\lambda \gg R$, but not when viewed by a probe with wavelength $\lambda \lesssim R$.

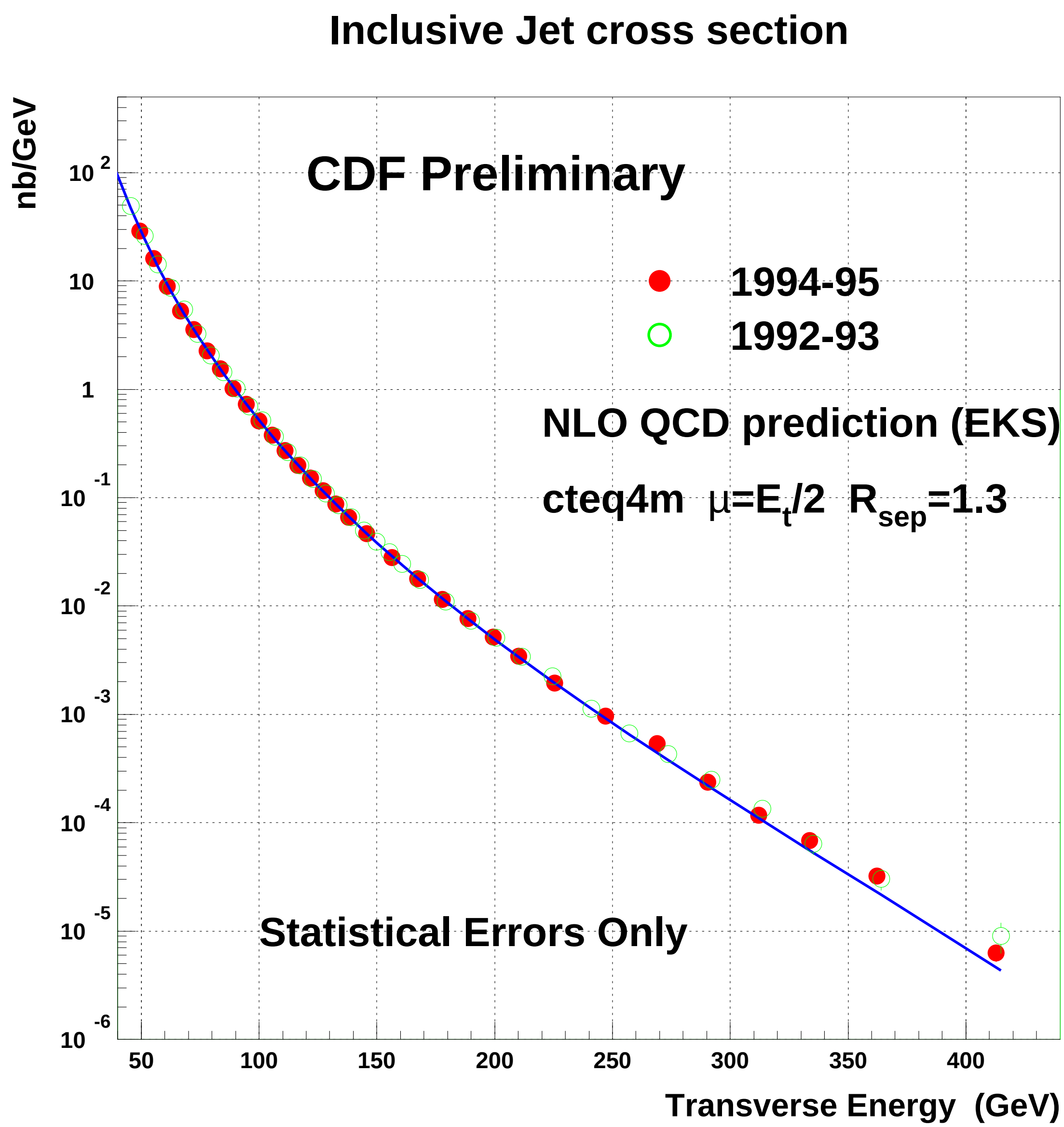
Then the one jet inclusive cross section should be suppressed by a form factor something like:

$$\frac{d\sigma_{\rm Jet}}{dE_T} \approx \left(\frac{d\sigma_{\rm Jet}}{dE_T}\right)_0 \times \exp(-RE_T)$$

(See, for example K. Y. Oda and N. Okada, arXiv:hep-ph/0111298.)

## The evidence so far

- QCD works up to the highest $E_T$ probed by Fermilab

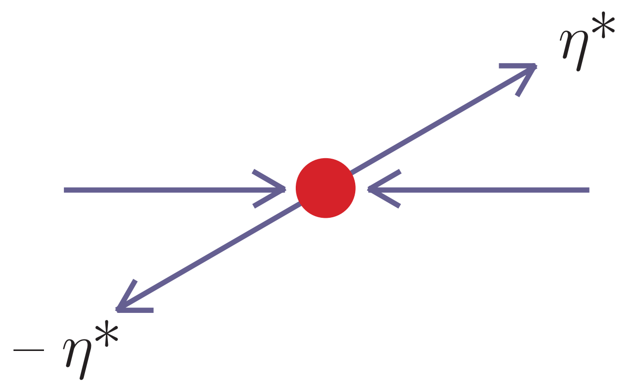


## The two jet inclusive cross section

Find the two jets in each event with the largest $E_T$. Study

$$\frac{d\,\sigma}{d\,M_{JJ}\;d\,\eta_{JJ}\;d\,\eta^*}\;.$$

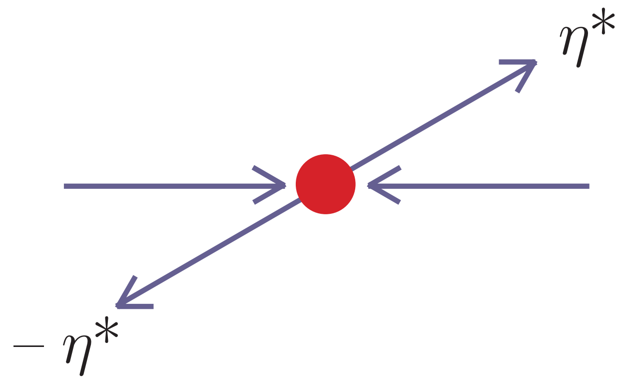
- $\eta_{JJ} = (\eta_1 + \eta_2)/2 =$ rapidity of jet-jet c.m. system
- $\eta^* = (\eta_1 - \eta_2)/2 =$ rapidity of first jet as viewed in the jet-jet c.m. system.
- $\eta^* = -\ln\tan(\Theta^*/2)$.




- $d\sigma/dM_{JJ}$ has essentially the same information as the one jet inclusive cross section.
- However, a resonance that decays to two jets would appear as a bump.

- The two jet angular distribution contains very important information.
- Look at the cross section as a function of $\eta^*$ for a fixed bin of $M_{JJ}$ and $\eta_{JJ}$.

## Vector exchange versus new terms




- Vector boson exchange gives the characteristic behavior

$$\frac{d\,\sigma}{d\,\eta^*} \propto \exp(2\eta^*) \qquad\qquad \eta^* \gg 1\ .$$

- An s-wave distribution gives few events with $\eta^* > 1$.

A convenient angle variable is

$$\chi = \exp(2\eta^*)$$

so

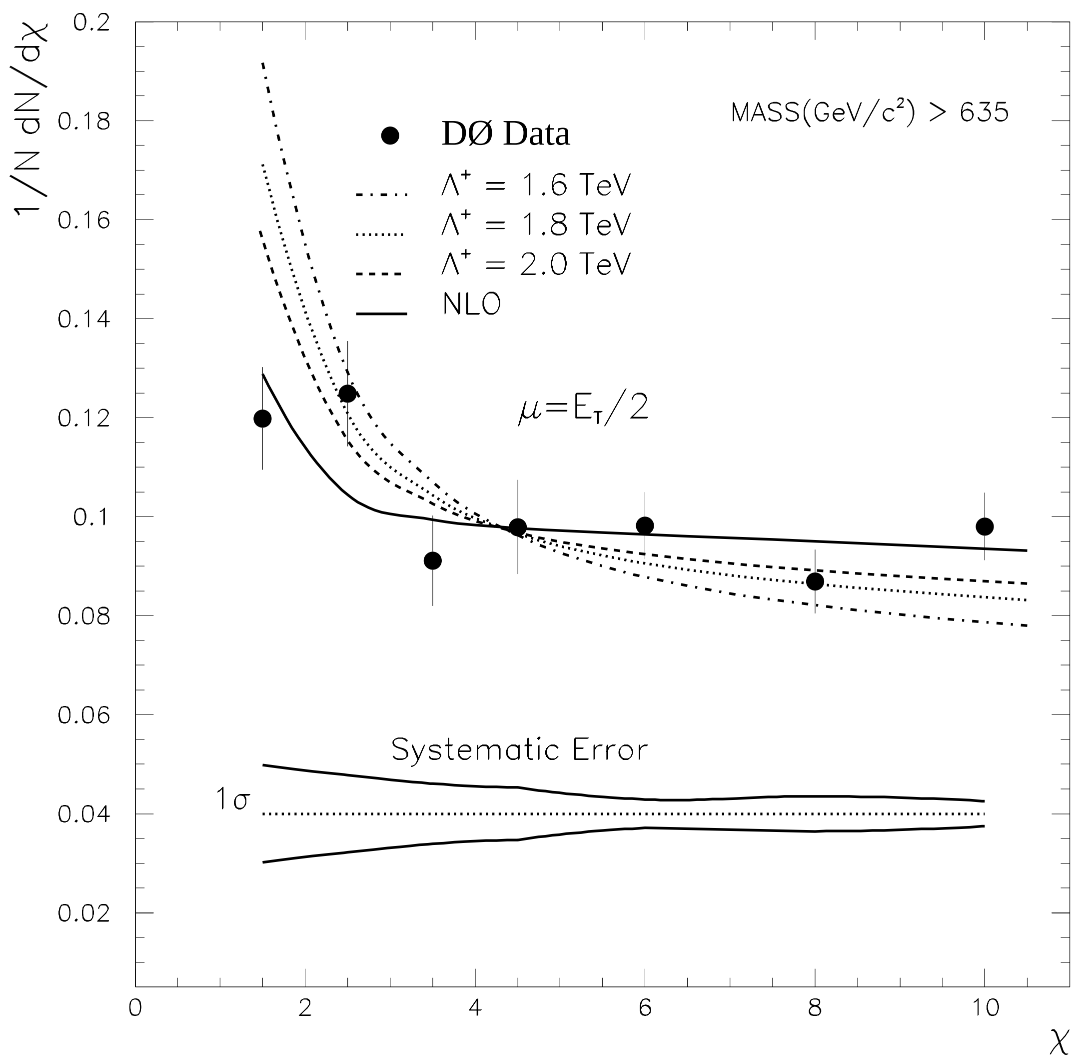
$$\frac{d\,\sigma}{d\,\chi} = \frac{1}{2\,\exp(2\eta^*)}\,\frac{d\,\sigma}{d\,\eta^*}$$

The QCD cross section is quite flat for $\chi \gg 1$.

In contrast, a new physics signal should fall off beyond $\chi \approx 3$.

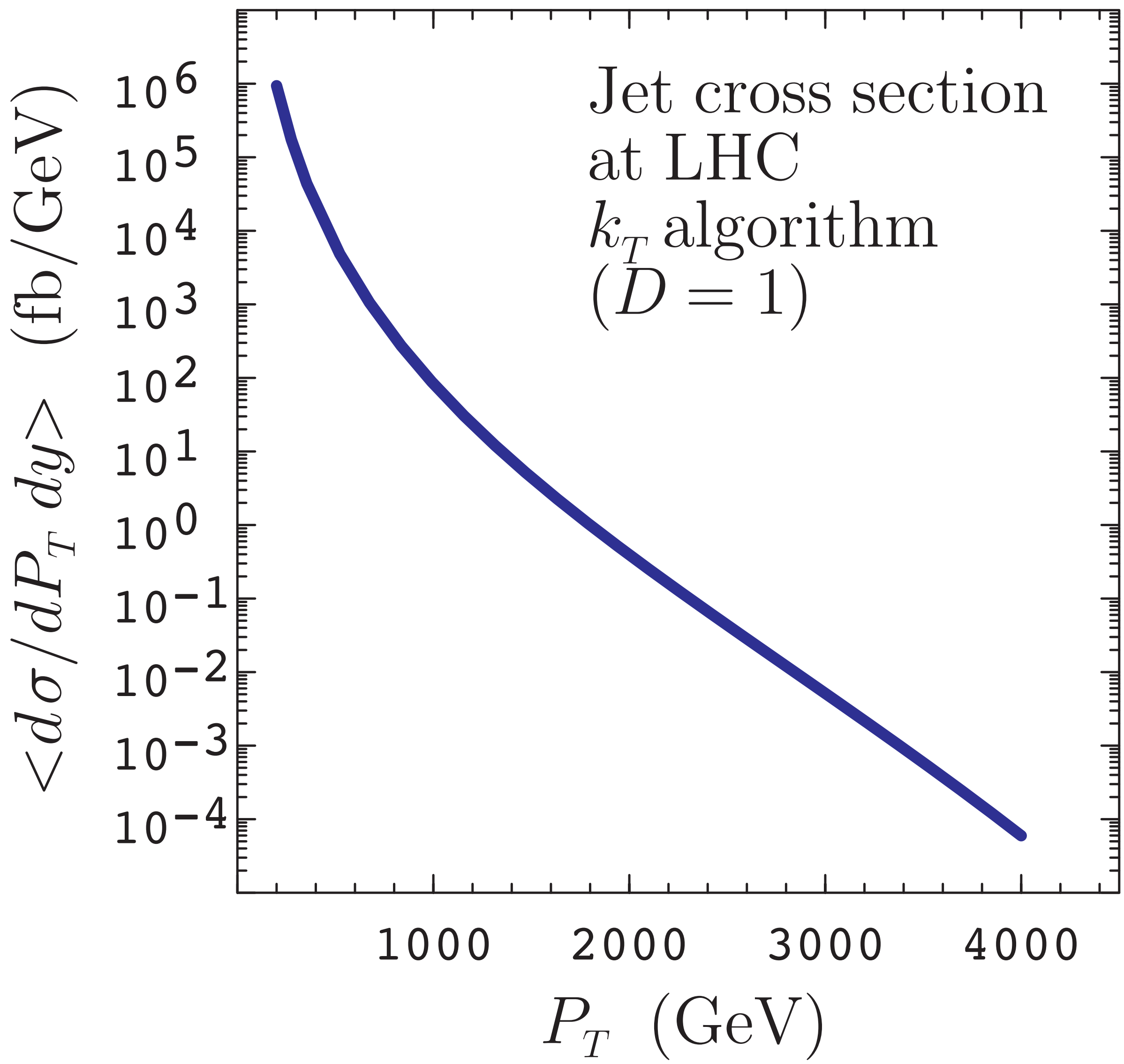
## Comparison with Tevatron data

Here we compare QCD with D0 data for dijets jets with large $M_{JJ}$.




The CDF data for $d\sigma/dE_T$ was showing a distinct excess at large $E_T$ (not seen by D0). But the dijet angular distribution (from both experiments) showed that there was no new physics.

## The prediction for LHC




One jet inclusive cross section $d\sigma/dP_T\,dy$
averaged over $-1 < y < 1$, versus $P_T$.

- Successive combination jet definition, $k_T$ style, with joining parameter $D = 1$.
- $\mu_{\rm UV} = \mu_{\rm coll} = P_T/2$.
- CTEQ5M partons.

## Theory errors (perturbative)

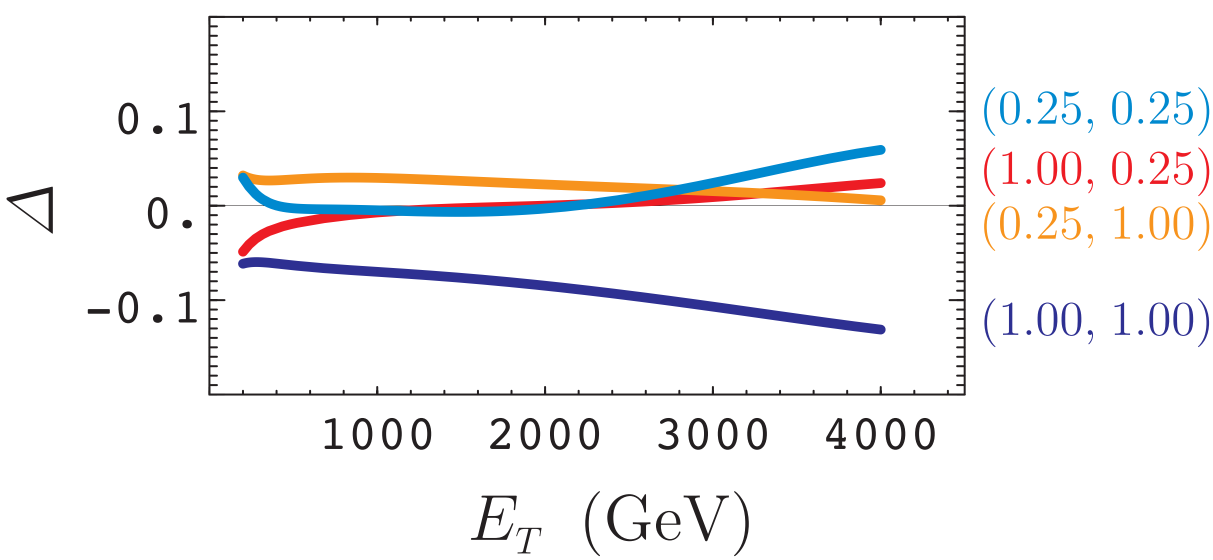
- Calculation includes order $\alpha_s^2$ and $\alpha_s^3$.
- Order $\alpha_s^4$ and higher are left out.
- The missing $\alpha_s^4$ terms are probably not smaller than terms we know about

$$\text{const.} \times \alpha_s^4 \ln(2\,\mu/E_T)$$

where $\mu$ is $\mu_{\rm UV}$ or $\mu_{\rm coll}$.

- Investigate by examining

$$\Delta(\mu_{\rm coll}, \mu_{\rm UV}) = \frac{d\sigma(\mu_{\rm UV}, \mu_{\rm coll})/dE_T}{d\sigma(E_T/2, E_T/2)/dE_T} - 1$$




$\Delta(\mu_{\rm UV}, \mu_{\rm coll})$ for $(\mu_{\rm UV}, \mu_{\rm coll})$ choices $(E_T/4, E_T/4), (E_T, E_T/4), (E_T/4, E_T), (E_T, E_T)$.

## Theory errors (power suppressed)

There are errors of the form

$$\frac{d\sigma}{dE_T} = \left(\frac{d\sigma}{dE_T}\right)_{\rm NLO} \left\{1 + \frac{\Lambda_1}{E_T} + \frac{\Lambda_2^2}{E_T^2} + \cdots\right\}$$

from

- hadronization
- $k_T$ kicks to incoming partons
- splash-in
- splash-out

Rough estimates suggest $\Lambda_i \lesssim 10$ GeV for Fermilab. (Maybe somewhat more for LHC).

This is significant for comparison of jets at $\sqrt{s} = 630$ GeV to jets at $\sqrt{s} = 1800$ GeV at Fermilab.

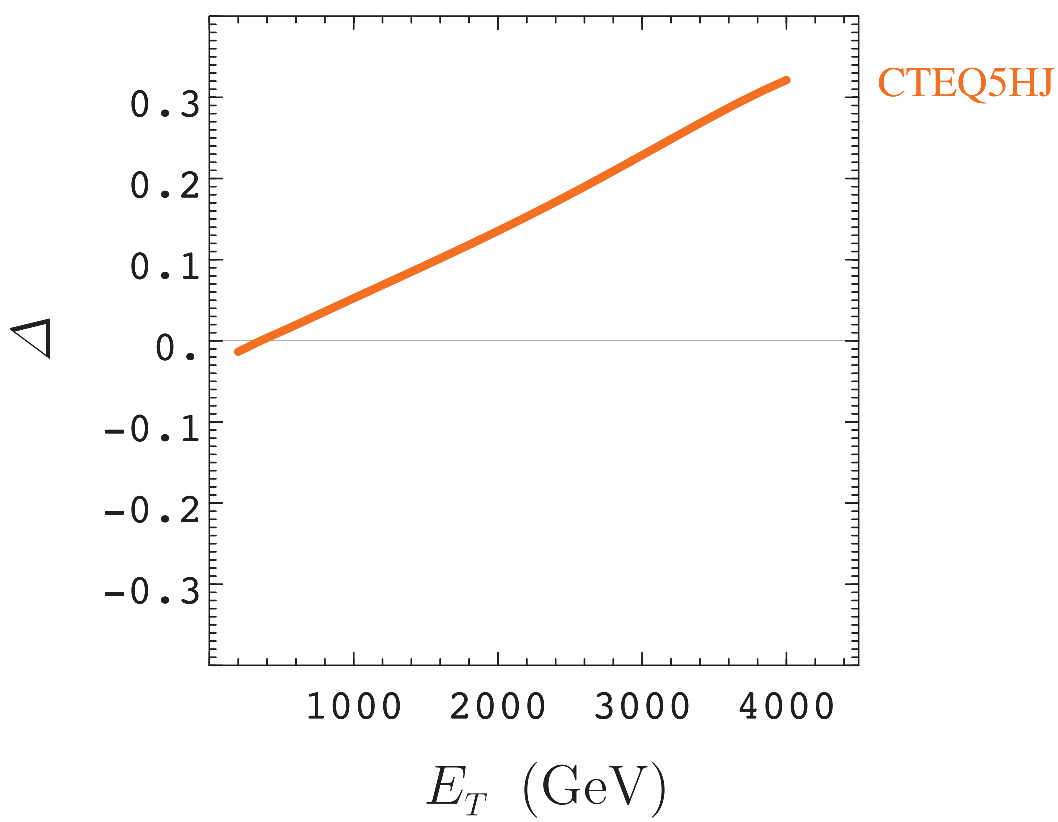
These should be negligible for jets with $E_T > 200$ GeV at LHC.

## Parton distribution errors

- For $x < 0.3$, I suppose we know parton distributions to 10%, so jet cross sections to 20%.
- For larger $x$, knowledge of the gluon distribution is poor.
- To see what can happen, try CTEQ5HJ partons.
  * Enhanced large $x$ gluons to fit average of CDF and D0 jet data at high $E_T$.

Plot

$$\Delta = \frac{d\sigma(CTEQ5HJ)/dE_T}{d\sigma(CTEQ5M)/dE_T} - 1$$




- It would be nice to have parton distributions with errors.
- Giele and Keller have published ideas on this.
- CTEQ has partially accomplished this.

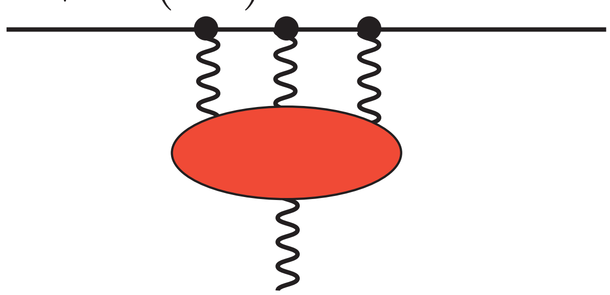
## Reducing the perturbative theory error

We should do the calculation at NNLO.

• The premier example is $g-2$ for the muon.
Experiment E821 at Brookhaven gives

$$(g-2)/2 = 11\,659\,203(8) \times 10^{-10}$$

The corresponding calculation includes QED calculations at $N^4LO$, *i.e.* $\alpha^5$. The calculation also includes two loop graphs with $W$ and $Z$ bosons. There are also QCD contributions, which cannot be purely perturbative because the momentum scale is too low. One contribution had a sign error, fixed by Knecht and Nyffeler, who found that this graph contributes $+8.3(1.2) \times 10^{-10}$.



The theory result is

$$(g-2)/2 = 11\,659\,169(8) \times 10^{-10}.$$

• The revised theoretical contribution helps.

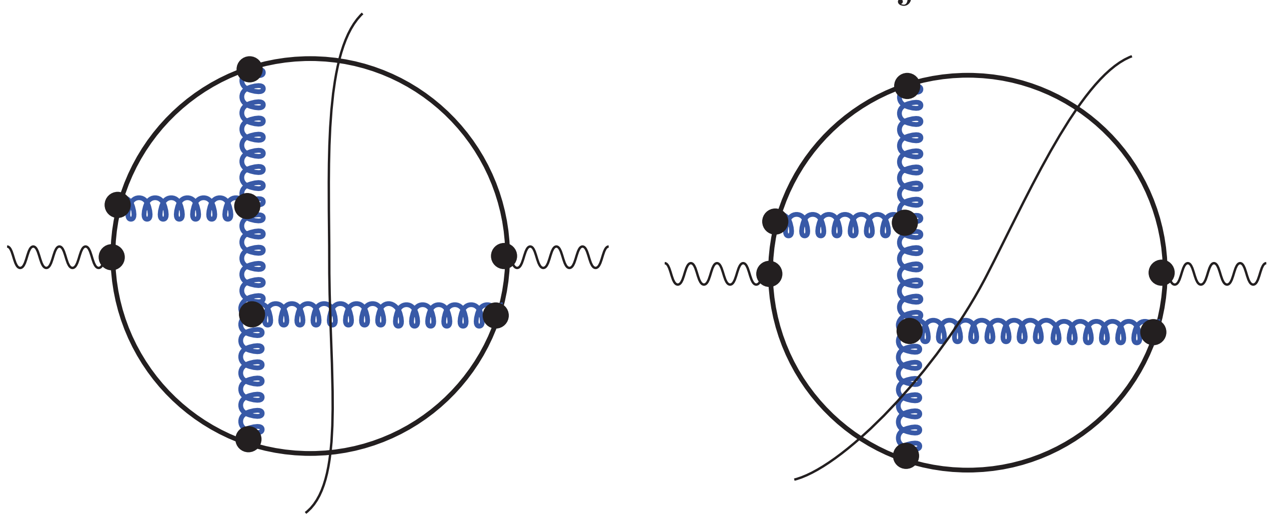
• There is perhaps more theoretical uncertainty than indicated by the (8).

• This has a bearing on LHC physics because it suggests beyond the Standard Model stuff.

• $N^kL0$ calculations matter.

The calculations at NNLO and beyond are successful because they use special tricks based on calculating a simple measurable quantity. It is harder to calculate for *generic* infrared safe observables.

The first results will come for $e^+e^- \to 3$ jets.

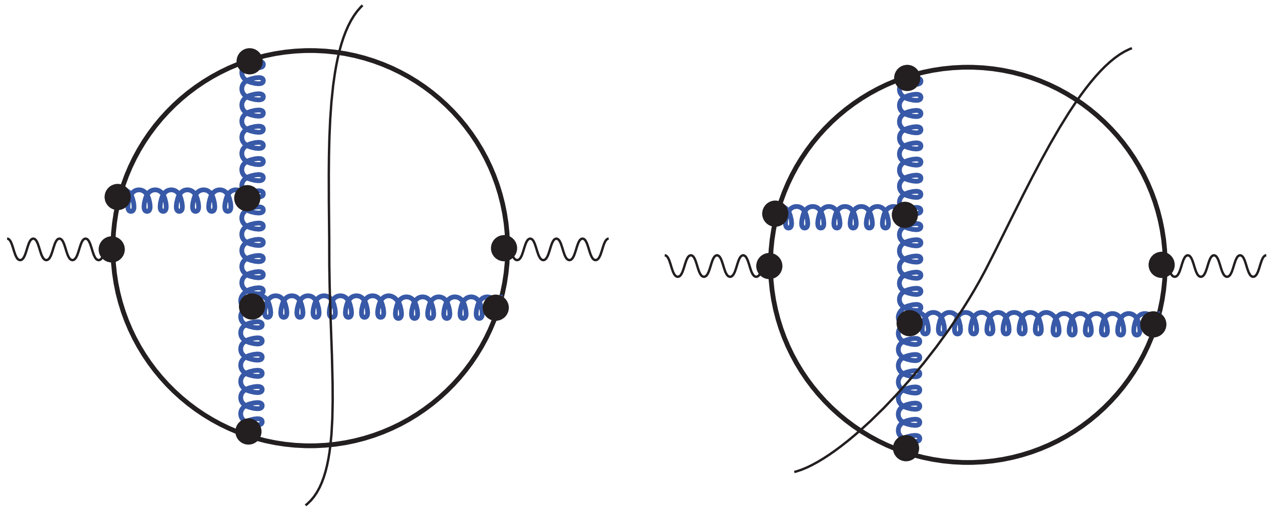


Two cut graphs for $e^+e^- \to 3$ jets

Standard analytical/numerical method:

- Need analytic result for the two loop virtual graph.
  * Regularized using $4 - 2\epsilon$ dimensions.
- Need subtraction scheme for graphs with 4 and 5 final state partons.

## Recent progress



Standard analytical/numerical method:

- Need analytic result for the two loop virtual graphs.
  * Done
- Need subtraction scheme for graphs with 4 and 5 final state partons.
  * In progress

The progress is being made because there are several very good people working on this. For example,

- C. Anastasiou, M. Beneke, Z. Bern, K. G. Chetyrkin, L. Dixon, T. Gehrmann, E. W. Glover, S. Laporta, S. Moch, C. Oleari, E. Remiddi, V. A. Smirnov, J. B. Tausk, P. Uwer, O. L. Veretin, S. Weinzierl.

At NLO, it is possible to do this kind of calculation by a completely numerical method. This offers evident advantages in flexibility. Perhaps it could help at NNLO. See D. E. Soper, Phys. Rev. Lett. **81**, 2638 (1998).

## Beyond fixed order

Simply calculating Feynman diagrams at a fixed order of perturbation theory is not enough.

## Use the factorization property of QCD

$$\frac{d\sigma}{dE_T\,dy} \approx$$
$$\sum_{a,b} \int_{x_A}^{1} d\xi_A \int_{x_B}^{1} d\xi_B \; f_{a/A}(\xi_A,\mu)\; f_{b/B}(\xi_B,\mu)\; \frac{d\hat{\sigma}^{ab}(\mu)}{dE_T\,dy}.$$

## Sum an infinite number of important contributions

- $\sum C_n\,[\alpha_s\,\log(\mu^2/\mu^2_{\rm data})]^n$
- $\sum C_n\,[\alpha_s\,\log^2(k_T^2/Q^2)]^n$
- $\sum C_n\,[\alpha_s\,\log(1/x)]^n$
- $\sum C_n\,[\alpha_s\,\log^2(1-x)]^n$

The first of these is performed by using the renormalization group. The others will be discussed at least briefly in my talk on perturbative summations.

## Parton distribution functions

$$f_{a/p}(x,\mu) \qquad a = g, u, \bar{u}, d, \bar{d}, \ldots .$$

- Important for everything.
- Determined from data for many processes (global fits).
- Produced by CTEQ and MRS.

- Charm and bottom distributions are calculated, based an expansion in powers of $\alpha_s(m_c)$ and $\alpha_s(m_b)$ respectively.
  - Maybe this isn't such a good approximation.

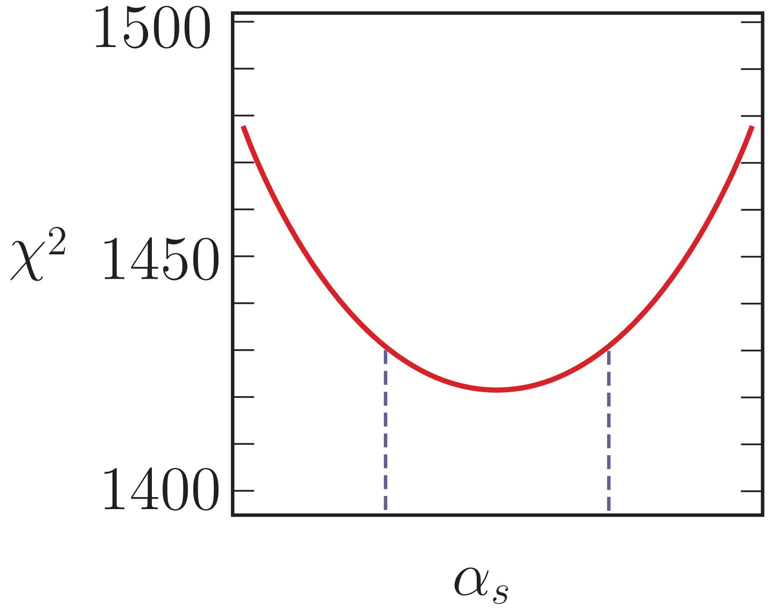
- We don't know the gluon distribution at large values of $x$ (say 0.5) to within a factor of 2.
- So far, error analysis by MRS and CTEQ consists of trying different values of gluon parameters, $\alpha_s$, *etc* and seeing how far one can go before the fit is evidently bad.

Why don't they give us parton distributions with errors?

- It's harder than you think.
- CTEQ has, in part.

## Why parton distributions with errors would seem strange

• With a real error analysis involving roughly 1400 degrees of freedom, we would have a plot of $\chi^2$ for some observable, say $\alpha_s$, that looks like this:




• The difference in $\chi^2$ translates to a likelihood ratio

$$\frac{\mathcal{L}(\alpha_s)}{\mathcal{L}_{\rm best}} = \exp\left\{-\frac{\chi^2(\alpha_s) - \chi^2_{\rm best}}{2}\right\}.$$

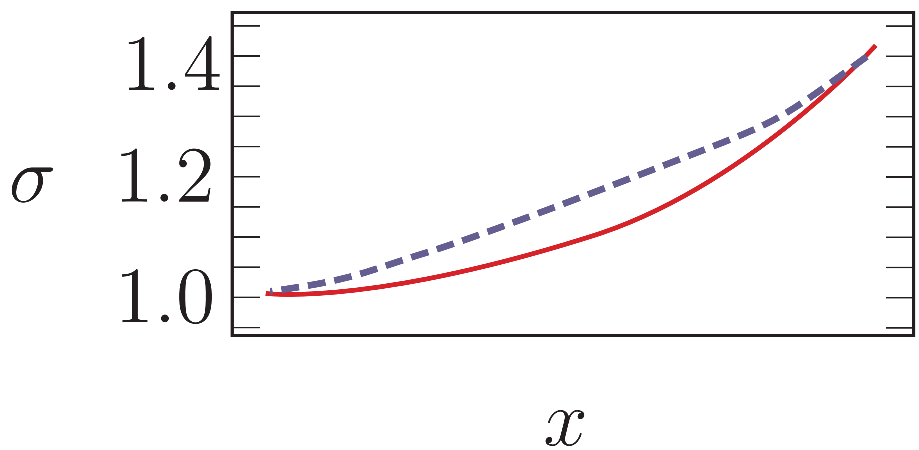
• Thus a $\Delta\chi^2 = 9$ gives $\mathcal{L}(\alpha_s)/\mathcal{L}_{\rm best} = \exp(-4.5) \approx 1/90$.
• Any value of $\alpha_s$ beyond $\Delta\chi^2 = 9$ can be ruled out with high confidence.
• But the fits to the world's data corresponding to much of the disallowed range of $\alpha_s$ would seem perfectly fine: normal fluctuations in $\chi^2$ for 1400 degrees of freedom are about 50.

## Parton error analysis

• In order to take $\chi^2$ or an equivalent statistic seriously, one must be very careful.

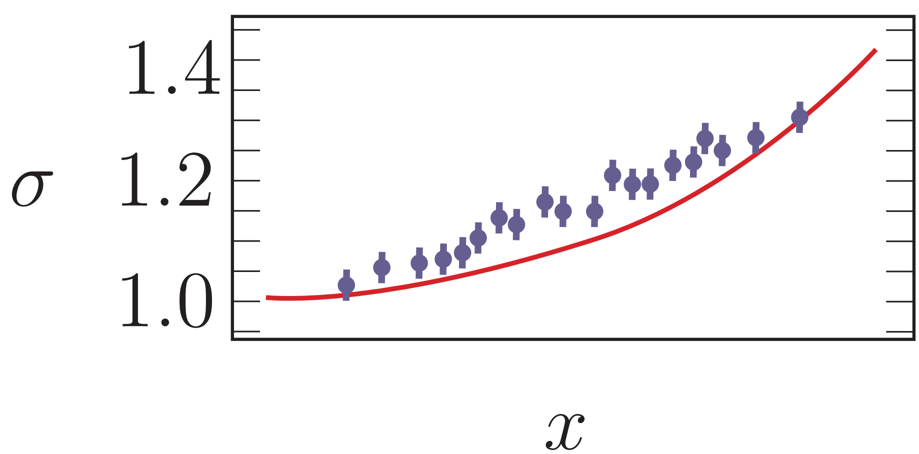
• It is easy to go seriously wrong.

• Suppose that we judge a calculated cross section to have a theoretical error such that a "perfect" Standard Model prediction could differ by 5% in the mid-range of some variable $x$:




• But suppose that we ignore this theoretical error.

• Then if the data look like this




we could erroneously add many units to $\chi^2$ when we should just add one unit of $\chi^2$ for the theory being " $1\ \sigma$" off.

## Where are we?

• Alekhin produced parton distributions with errors based on DIS data in 1996.
• Giele, Keller, and Kosower have developed a method for doing the analysis if the relevant errors are available.
• Their method is flexible so it could handle non-gaussian errors.

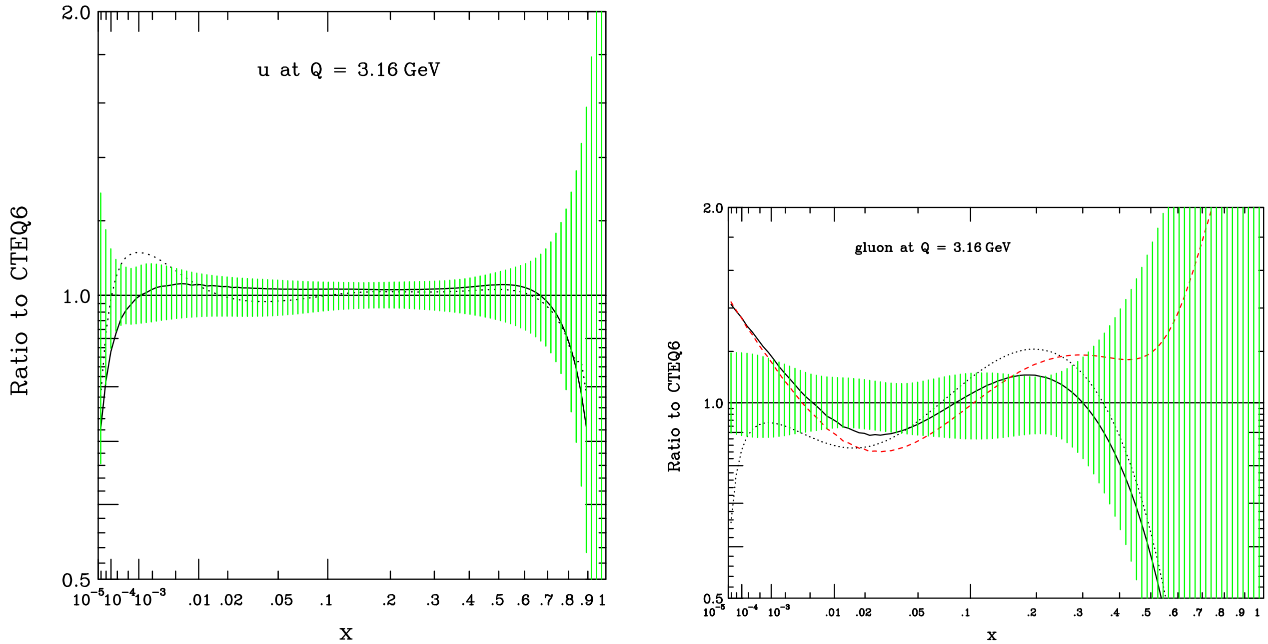
• More than mathematics is needed.
• A realistic treatment of errors would involve a lot of judgment by the fitters as to
  • experimental systematic errors
  • theoretical errors
• The necessary judgment would involve lots of analysis and debate.

• CTEQ's latest parton distribution set, CTEQ6, comes with errors.

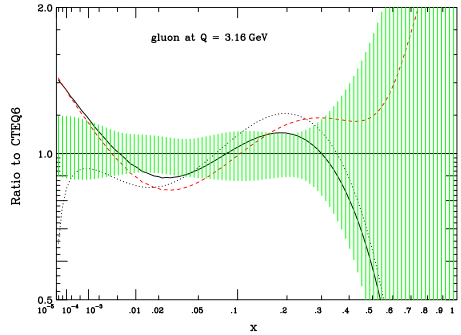
## CTEQ6

J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, JHEP **0207**, 012 (2002)

These parton distribution functions come with errors:




Uncertainty bands for the $u$-quark and gluon distribution functions at $Q^2 = 10\,\mathrm{GeV}^2$. The curves corresponds to CTEQ5M1(solid), CTEQ5HJ (dashed), and MRST2001 (dotted).




What do the error bands mean?

$$f_{a/p}(x, Q_0^2) = f^0_{a/p}(x, Q_0^2) + \sum_{i=1}^{20} \lambda_i \, h^i_{a/p}(x, Q_0^2)$$

so that

$$\chi^2 = \chi^2_{\rm min} + \sum_{i=1}^{20} \lambda_i^2 .$$

Then the error band is the envelope of the forty curves obtained with one $\lambda_i$ set to $+T$ or $-T$ and the others set to 0, where the "tolerance" is $T = 10$. (Actually, what was done was just slightly more complicated than this.)

• This gives some impression of the uncertainty.
• A better estimate would be

$$\Delta f_a = \left\{ \sum_{i=1}^{20} \left[ h^i_{a/p}(x, Q_0^2) \right]^2 \right\}^{1/2}$$

• If $\chi^2$ means what it is supposed to, the error band with $T = 10$ surely overestimates the uncertainty.

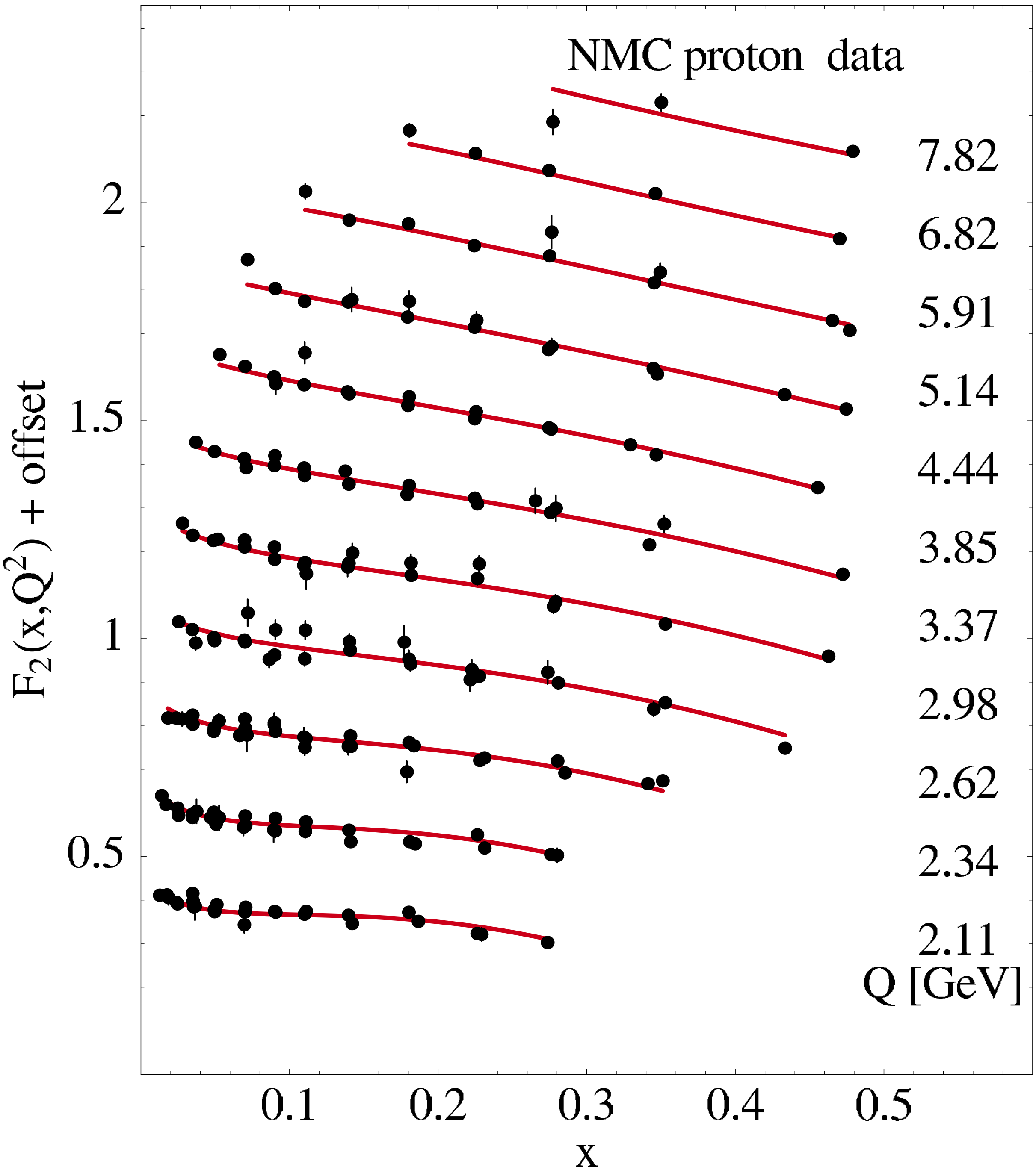
## Is $\chi^2$ the "real" $\chi^2$?

• Experimental systematic errors are included wherever available.
• But theoretical systematic errors are not included.
• Some of the data are not consistent with their errors.

| | $N_e$ | $\chi_e^2$ | $\chi_e^2/N_e$ |
|---|---|---|---|
| BCDMS p | 339 | 377.6 | 1.114 |
| BCDMS d | 251 | 279.7 | 1.114 |
| H1a | 104 | 98.59 | 0.948 |
| H1b | 126 | 129.1 | 1.024 |
| ZEUS | 229 | 262.6 | 1.147 |
| NMC F2p | 201 | 304.9 | 1.517 |
| NMC F2d/p | 123 | 111.8 | 0.909 |
| D0 jet | 90 | 64.86 | 0.721 |
| CDF jet | 33 | 48.57 | 1.472 |

The $\chi^2$ for D0 jet is too low. The $\chi^2$ for CDF jet is somewhat to high. The $\chi^2$ for NMC F2p is way too high. See the figure on the next page to see why $\chi^2$ is too high (Fig NMC.eps).

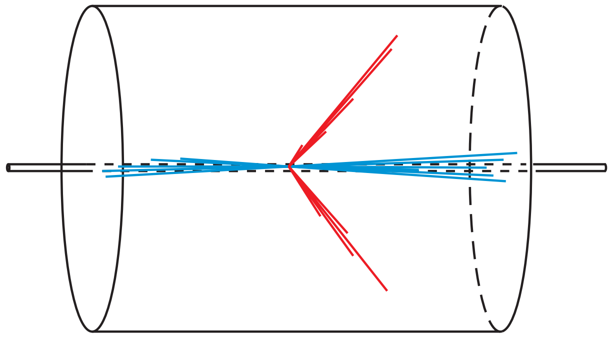
NMC proton data versus CTEC6M fit.




It appears that there is no way that a smooth curve could fit the data much better than CTEQ6 fits the data here.

I conclude that the CTEQ6 approach is OK under the circumstances.

# Jet definitions



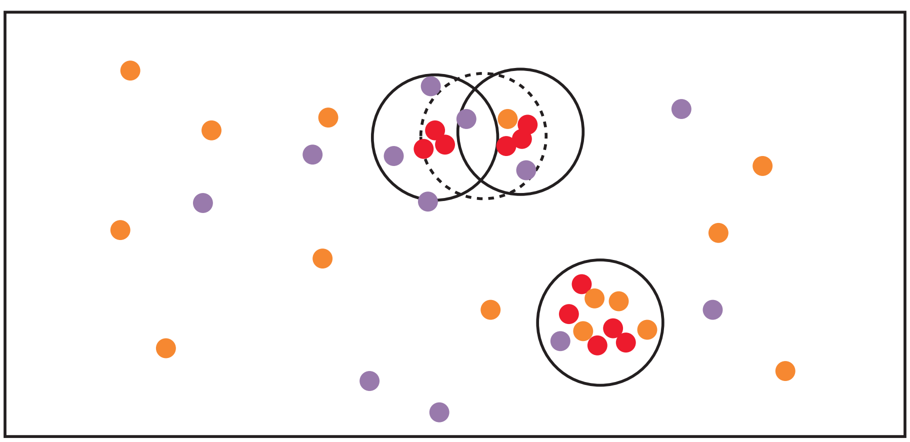
The most common definition in hadron-hadron collisions is based on cones.

The simple starting point.

- Use rapidity $\eta$ and azimuthal angle $\phi$.
- There is a jet axis with angles $\eta_J, \phi_J$.
- Particles with $(\eta - \eta_J)^2 + (\phi - \phi_J)^2 < R^2$ are in the jet.

The fine print.

- Cones can overlap.
- There special rules to say what to do.



- With the special rules, the cone definition is not simple. (See studies of R. Hirosky).
- Unless one is very careful, the algorithm is not infrared safe.

## Infrared safety

To be avoided:

$$\frac{d\sigma}{dE_T} = \left(\frac{d\sigma}{dE_T}\right)_{\rm pert} \times \left\{1 + \alpha_s^N F_{\rm IR}\right\}.$$

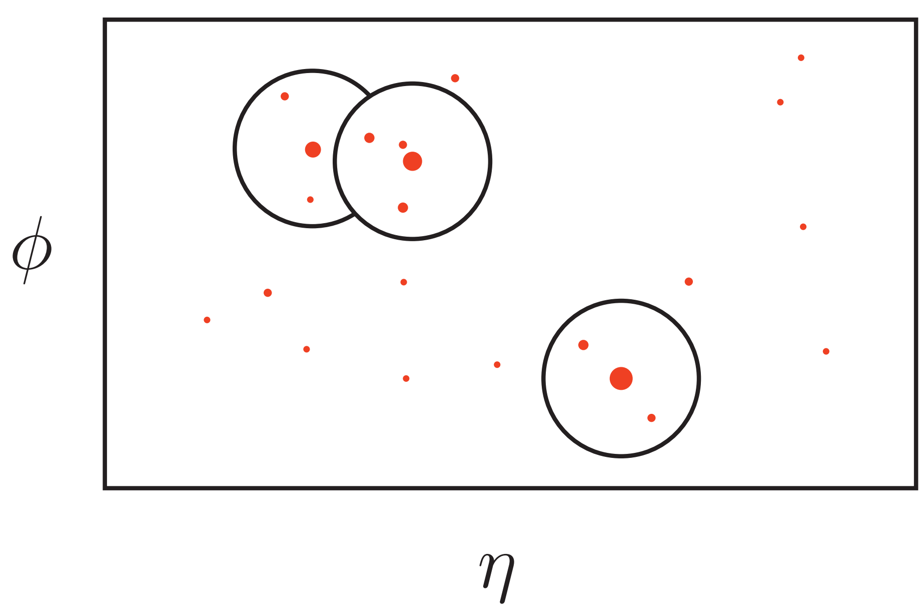
where $F_{\rm IR}$ is an unknown factor of order 1.

The test:

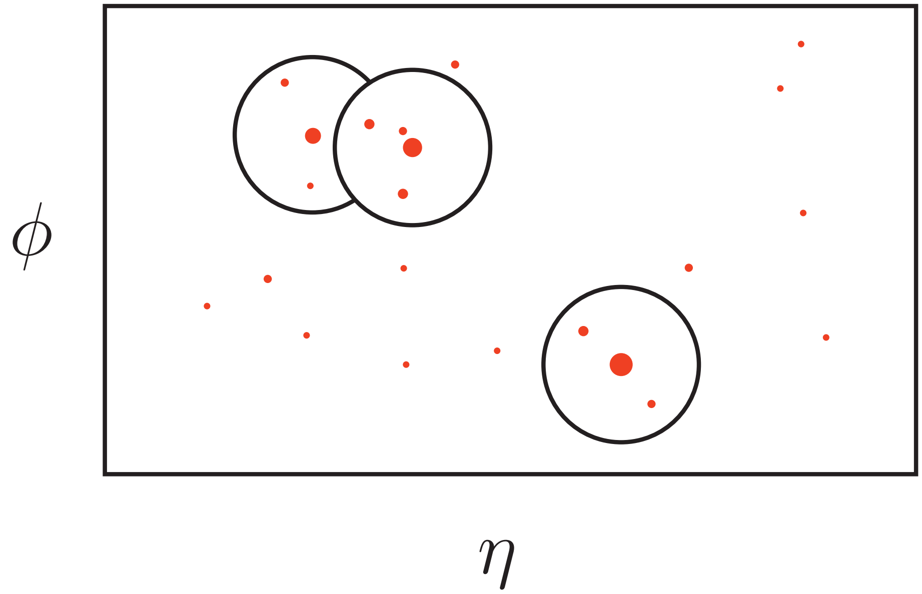
- Collinear splittings can't matter.
- Soft particles can't matter.

Sample algorithm:

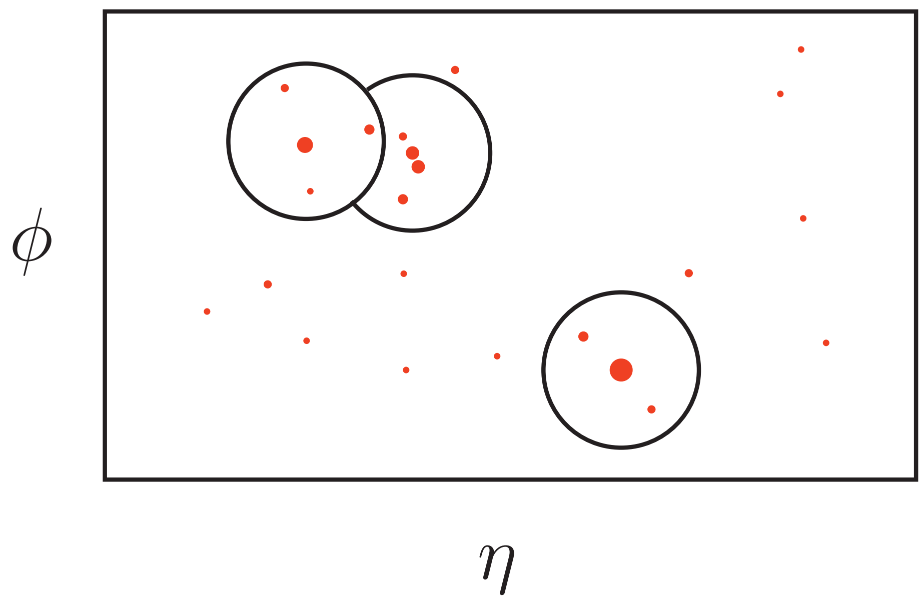
"Pick largest $E_T$ calorimeter call not already included in a jet as a 'seed'; all cells with $(\eta - \eta_{\rm seed})^2 + (\phi - \phi_{\rm seed})^2 < (0.7)^2$ and not already in a jet become the next jet, with $p^\mu_{\rm jet} = \sum_i P_i^\mu$ and $E_{T,{\rm jet}} = |\vec{P}_{T,{\rm jet}}|$."



## The sample algorithm is not IR safe



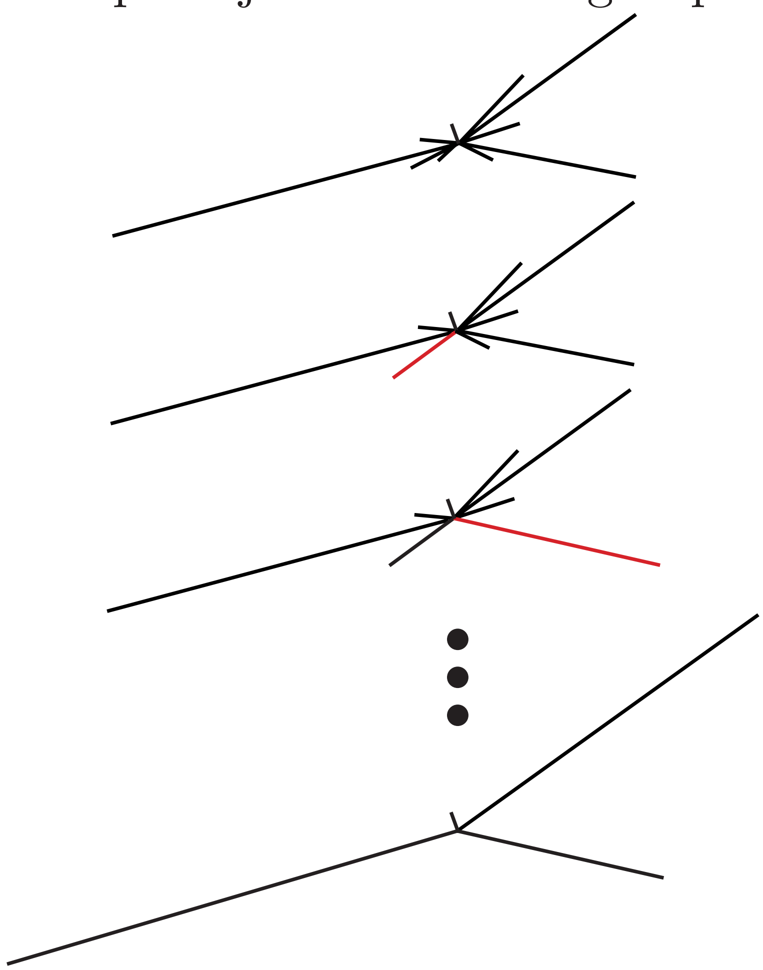

If the parton that made seed 2 splits, it changes the result.



## A simple jet algorithm

The $k_T$ algorithm is based on the successive combination algorithms used in $e^+e^-$ physics but is adapted for hadron collisions.

- At each stage, one has a collection of protojets with variables $(P_{T,i}, y_i, \phi_i)$.
    - $y_i$ is the true rapidity of the protojet.
- To start with the each protojets is an observed particle (or a group of particles in a single calorimeter tower).
- At the end, the protojets have been grouped into jets.



## The rules

- For each protojet $i$, define

$$d_i = P_{T,i}^2$$

and for each pair of protojets $(i, j)$ define

$$d_{i,j} = \min[P_{T,i}^2, P_{T,j}^2][(y_i - y_j)^2 + (\phi_i - \phi_j)^2]/D^2$$

where $D$ is a parameter (say $D = 1$).

(1) Find the smallest of the $d_i$ and $d_{i,j}$.
(2) If the smallest is a $d_i$, remove $i$ from the list of protojets and add it to the list of jets.
(3) If the smallest is a $d_{i,j}$, combine protojets $i$ and $j$.

$$P^\mu = P_i^\mu + P_j^\mu$$

(4) GO TO 1