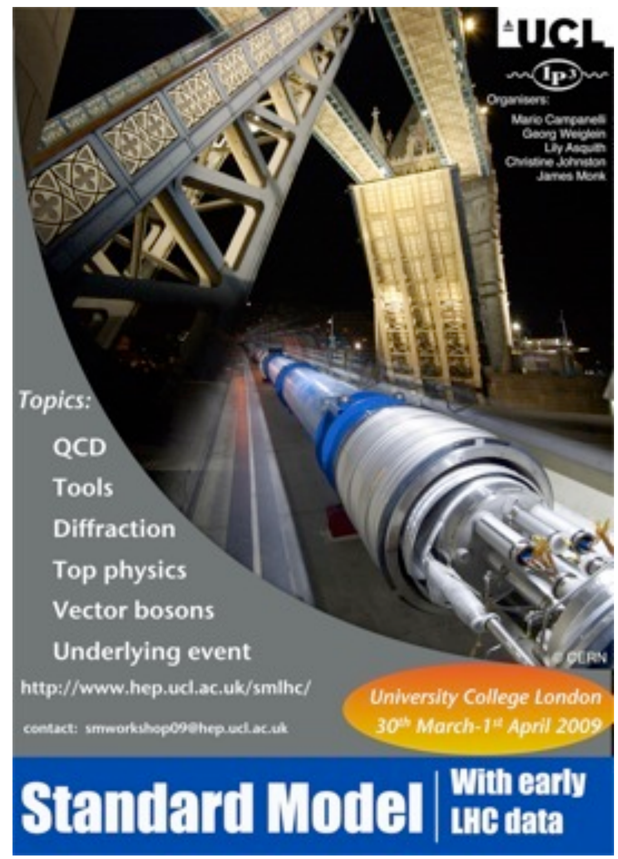
CTEQ update and discussion

J. Huston

Michigan State University *and*

IPPP, University of Durham



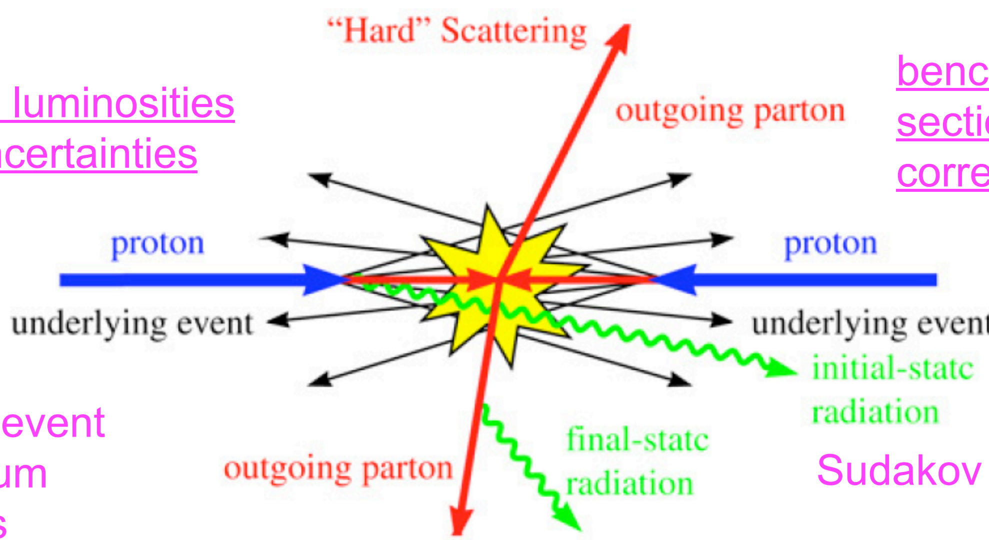
Understanding SM predictions at the LHC

…and re-discovering the Standard Model

LO, NLO and NNLO calculations
K-factors

PDF’s, PDF luminosities
and PDF uncertainties

benchmark cross
sections and pdf
correlations



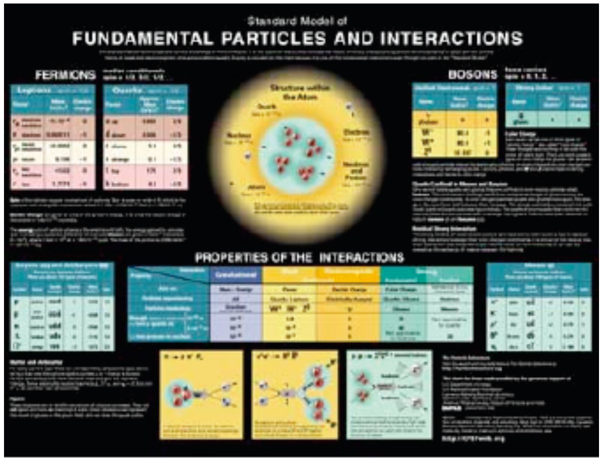
underlying event
and minimum
bias events

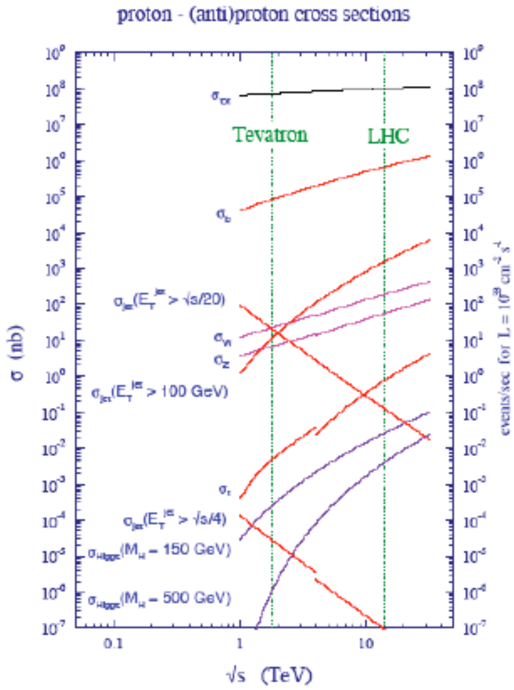
Sudakov form factors

jet algorithms and jet reconstruction

Understanding cross sections at the LHC

- We're all looking for BSM physics at the LHC
- Before we publish BSM discoveries from the early running of the LHC, we want to make sure that we measure/understand SM cross sections
  - detector and reconstruction algorithms operating properly
  - SM physics understood properly
  - SM backgrounds to BSM physics correctly taken into account





Parton distribution functions and global fits

- Calculation of production cross sections at the LHC relies upon knowledge of pdf's in the relevant kinematic region
- Pdf's are determined by global analyses of data from DIS, DY and jet production
- Two major groups that provide semi-regular updates to parton distributions when new data/theory becomes available
  - MRS->MRST98->MRST99 ->MRST2001->MRST2002 ->MRST2003->MRST2004 ->MSTW2008
  - CTEQ->CTEQ5->CTEQ6 ->CTEQ6.1->CTEQ6.5 ->CTEQ6.6->CT09G

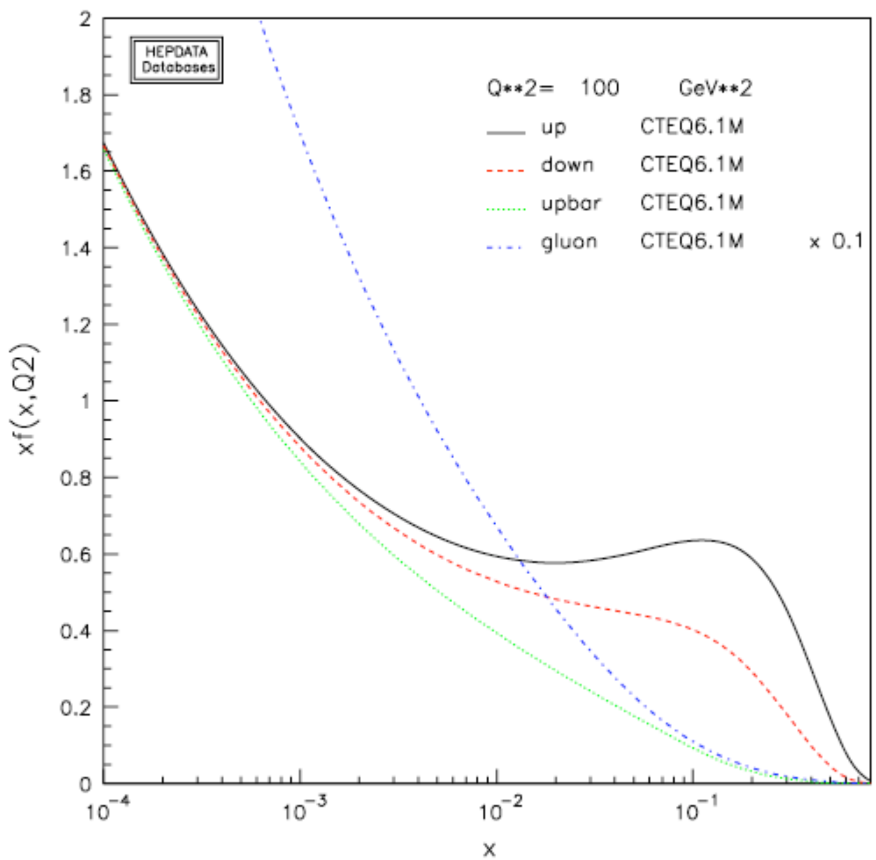
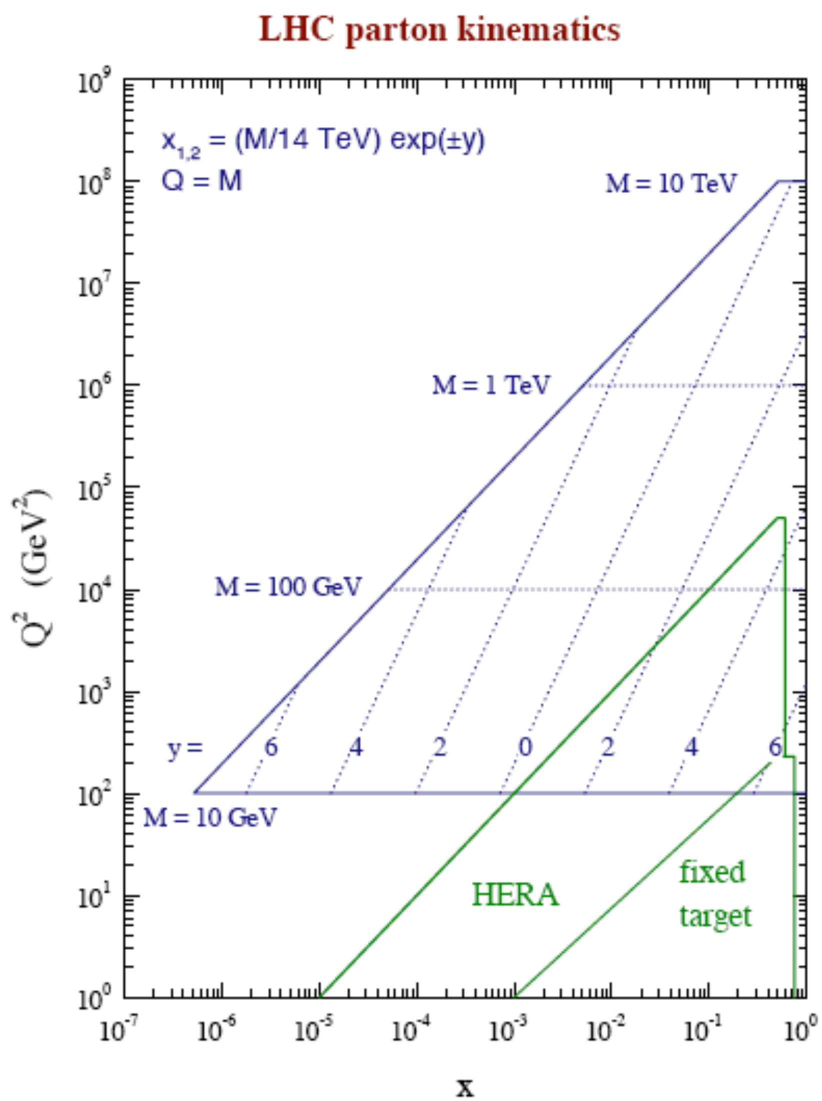


Figure 27. The CTEQ6.1 parton distribution functions evaluated at a Q of 10 GeV.

Cross sections at the LHC

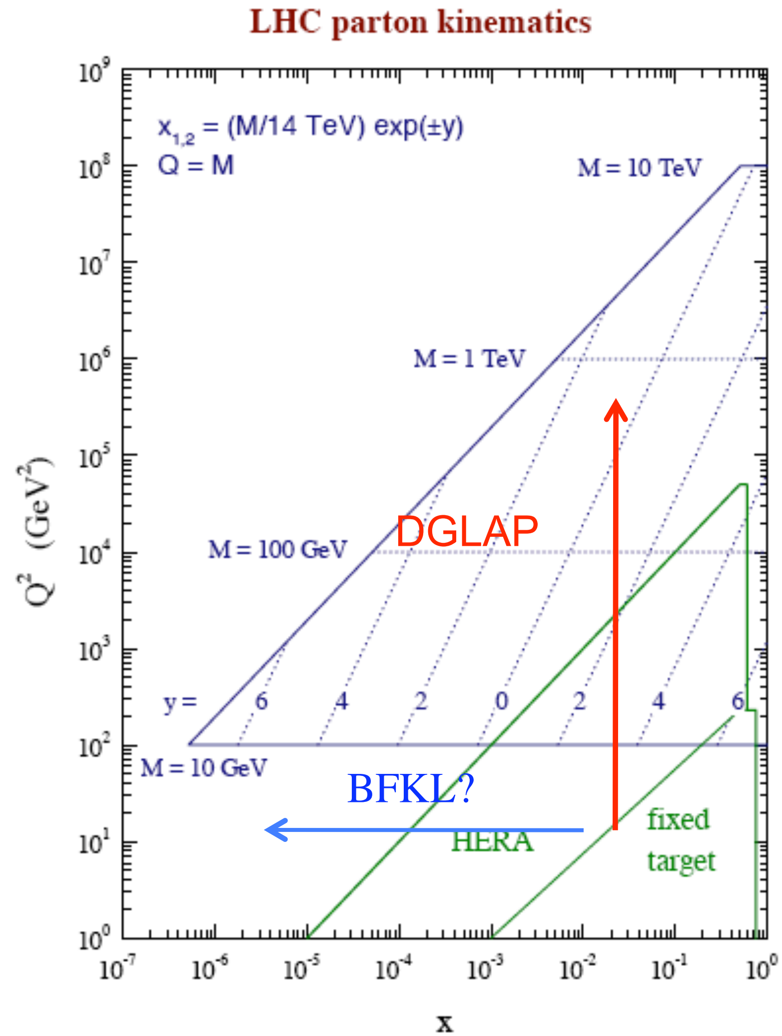
- Experience at the Tevatron is very useful, but scattering at the LHC is not necessarily just “rescaled” scattering at the Tevatron
- Small typical momentum fractions x in many key searches
  - dominance of gluon and sea quark scattering
  - large phase space for gluon emission and thus for production of extra jets
  - intensive QCD backgrounds
  - or to summarize,…lots of Standard Model to wade through to find the BSM pony



Cross sections at the LHC

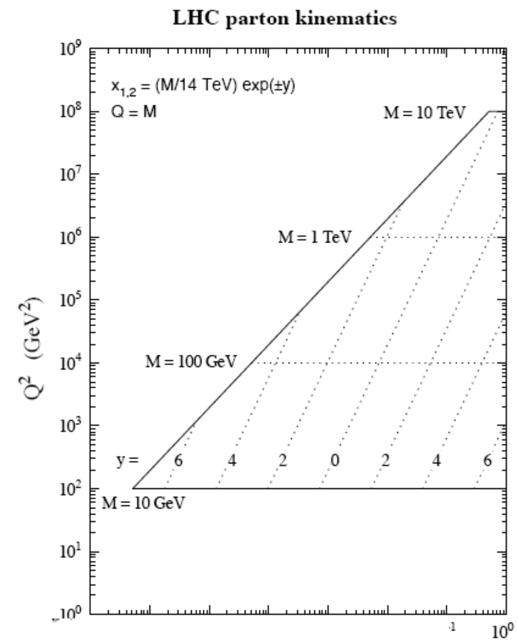
- Note that the data from HERA and fixed target cover only part of kinematic range accessible at the LHC
- We will access pdf's down to 1E-6 (crucial for the underlying event) and Q^2 up to 100 TeV^2
- We can use the DGLAP equations to evolve to the relevant x and Q^2 range, but…
  - we're somewhat blind in extrapolating to lower x values than present in the HERA data, so uncertainty may be larger than currently estimated
  - we're assuming that DGLAP is all there is; at low x BFKL type of logarithms may become important

$$\frac{d\sigma}{dM^2 dy} = \frac{\hat{\sigma}_0}{Ns}\left[\sum_k Q_k^2\left(q_k(x_1, M^2)\bar{q}_k(x_2, M^2) + [1 \leftrightarrow 2]\right)\right]$$



Parton kinematics at the LHC

- To serve as a handy "look-up" table, it's useful to define a parton-parton luminosity (a la EHLQ)
- Equation 3 can be used to estimate the production rate for a hard scattering at the LHC as the product of a differential parton luminosity and a scaled hard scatter matrix element



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

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

this is from the CHS review paper

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

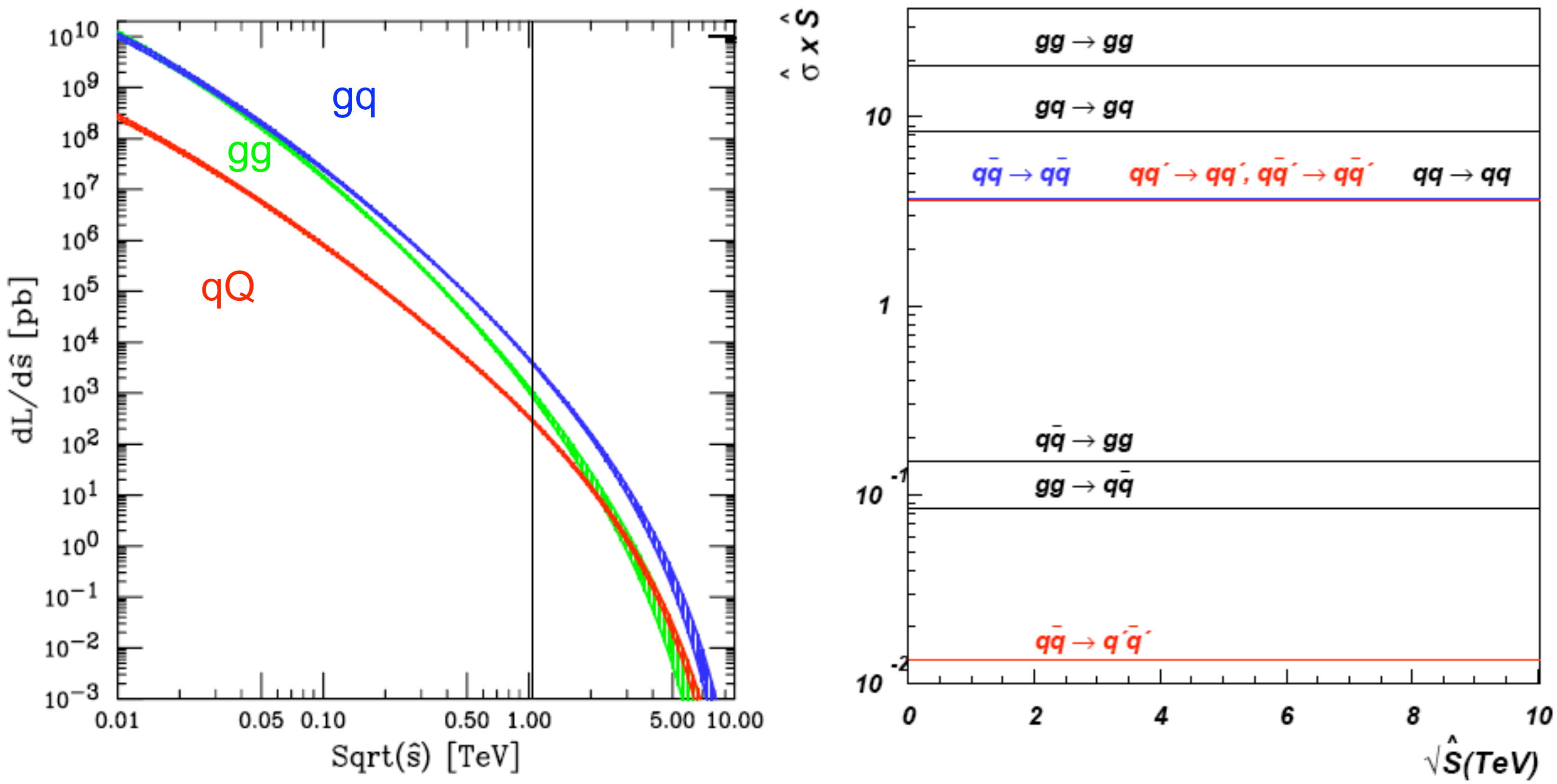
can then be written as

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

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

PDF uncertainties at the LHC

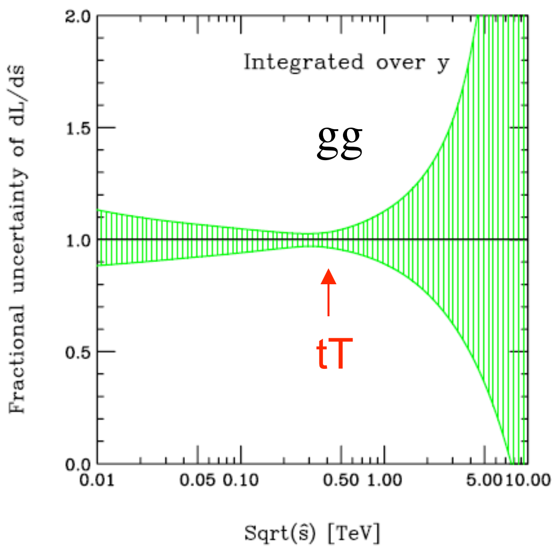


Fig. 4: Fractional uncertainty of gg luminosity integrated over y.

Note that for much of the SM/discovery range, the pdf luminosity uncertainty is small

Need similar level of precision in theory calculations

It will be a while, i.e. not in the first fb^{-1}, before the LHC data starts to constrain pdf's

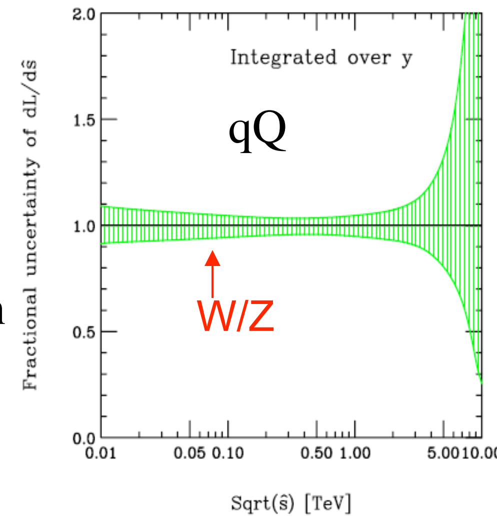


Fig. 7: Fractional uncertainty for Luminosity integrated over y for $d\bar{d} + u\bar{u} + s\bar{s} + c\bar{c} + b\bar{b} + \bar{d}d + \bar{u}u + \bar{s}s + \bar{c}c + \bar{b}b$.

NBIII: tT uncertainty is of the same order as W/Z production

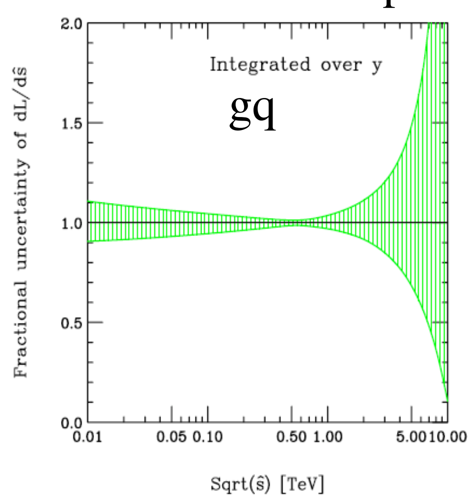


Fig. 6: Fractional uncertainty for Luminosity integrated over y for $g(d+u+s+c+b) + g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b}) + (d+u+s+c+b)g + (\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$,

NB I: the errors are determined using the Hessian method for a Dc^2 of 100 using only experimental uncertainties,i.e. no theory uncertainties

NB II: the pdf uncertainties for W/Z cross sections are not the smallest

Ratios:LHC to Tevatron pdf luminosities

- Processes that depend on qQ initial states (e.g. 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
- W+4 jets is a background to tT production both at the Tevatron and at the LHC
- tT production at the Tevatron is largely through a qQ initial states and so qQ->tT has an enhancement factor at the LHC of ~10
- Luckily tT has a gg initial state as well as qQ so total 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
  - known known: jet cuts have to be higher at LHC than at Tevatron

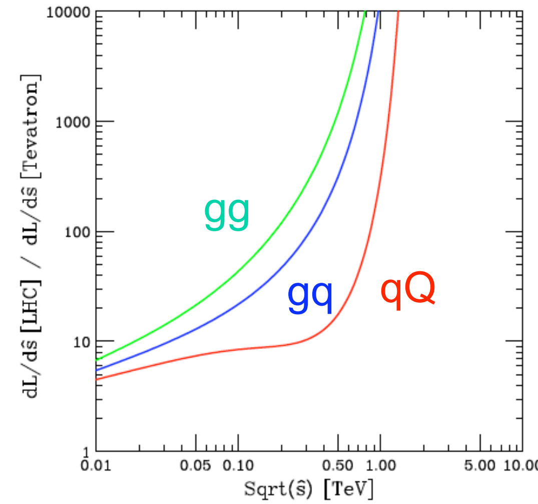


Figure 11. The ratio of parton-parton luminosity $\left[\frac{1}{\hat{s}}\frac{dL_{ij}}{d\tau}\right]$ in pb integrated over y at the LHC and Tevatron. Green=gg (top), Blue=$g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$ (middle), Red=$d\bar{d}+u\bar{u}+s\bar{s}+c\bar{c}+b\bar{b}+\bar{d}d+\bar{u}u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

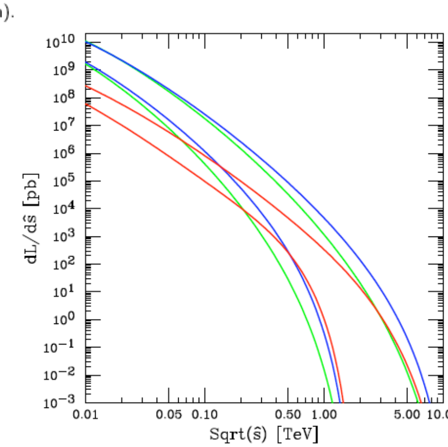


Figure 10. The parton-parton luminosity $\left[\frac{1}{\hat{s}}\frac{dL_{ij}}{d\tau}\right]$ in pb integrated over y. Green=gg, Blue=$g(d+u+s+c+b)+g(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})+(d+u+s+c+b)g+(\bar{d}+\bar{u}+\bar{s}+\bar{c}+\bar{b})g$, Red=$d\bar{d}+u\bar{u}+s\bar{s}+c\bar{c}+b\bar{b}+\bar{d}d+\bar{u}u+\bar{s}s+\bar{c}c+\bar{b}b$.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

…but wait, we're not running at 14 TeV in 2009-2010

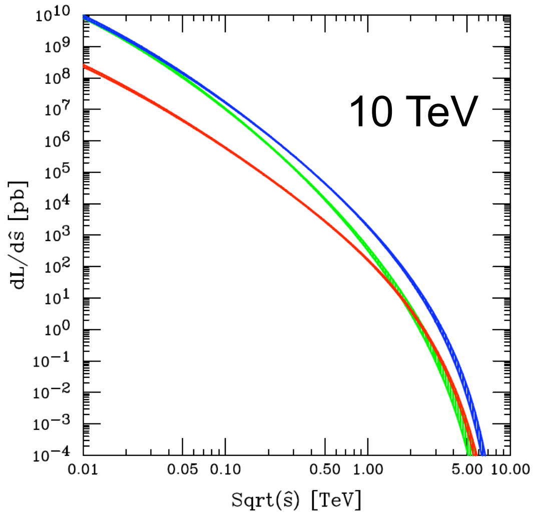


Figure 2: Luminosities integrated over y: LHC(pp) at $\sqrt{s} = 10\,\mathrm{TeV}$.
Green = gg, Blue = $gq + g\bar{q}$, Red = $u\bar{u} + d\bar{d} + s\bar{s} + c\bar{c} + b\bar{b}$.

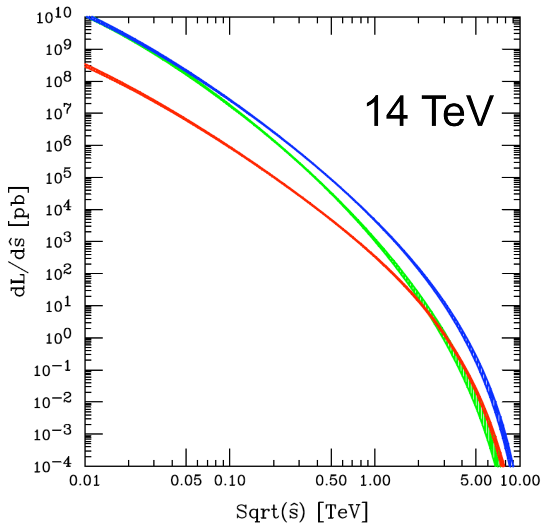


Figure 1: Luminosities integrated over y: LHC(pp) at $\sqrt{s} = 14\,\mathrm{TeV}$.
Green = $g\,g$, Blue = $gq + g\bar{q}$, Red = $u\bar{u} + d\bar{d} + s\bar{s} + c\bar{c} + b\bar{b}$.

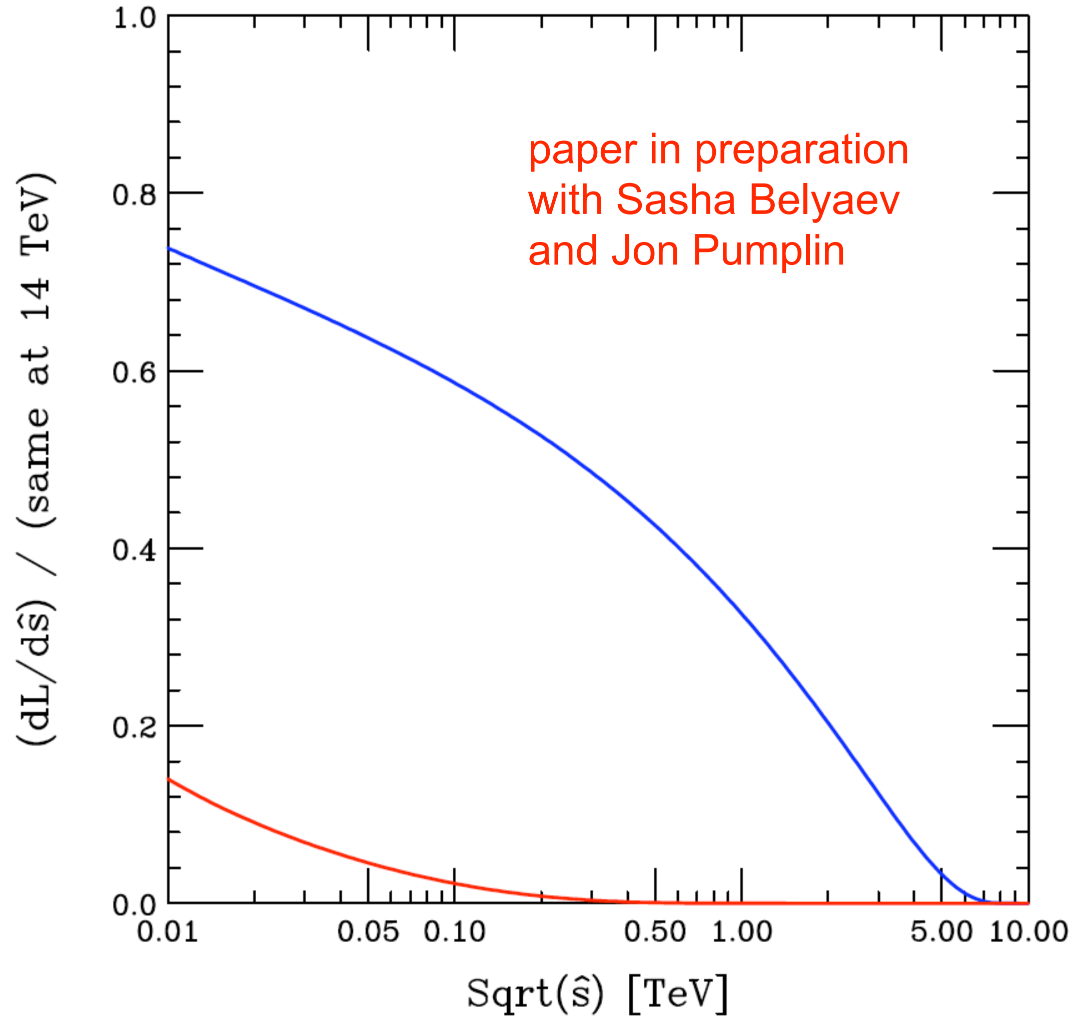


Figure 7: gg luminosity integrated over y:
Blue = (pp at 10 TeV) / (pp at 14 TeV);
Red = ($p\bar{p}$ at 1.96 TeV) / (pp at 14 TeV).

Look at ratios of pdf's at 1.96 and 10 TeV

- The plan is to run the LHC in 2009-2010 accumulating at least 200 pb^{-1}
- Take a discovery region (~1 TeV, say for squark pair production)
- The LHC is a factor of 50 more efficient at producing a 1 TeV object through a qQ initial state…so it would take 10 fb^{-1} at the Tevatron to equal the 200 pb^{-1} at the LHC
- …which the Tevatron will probably get (per expt)
- …with much better understood detectors and much lower backgrounds
- So don't count the Tevatron out just yet for discovery physics

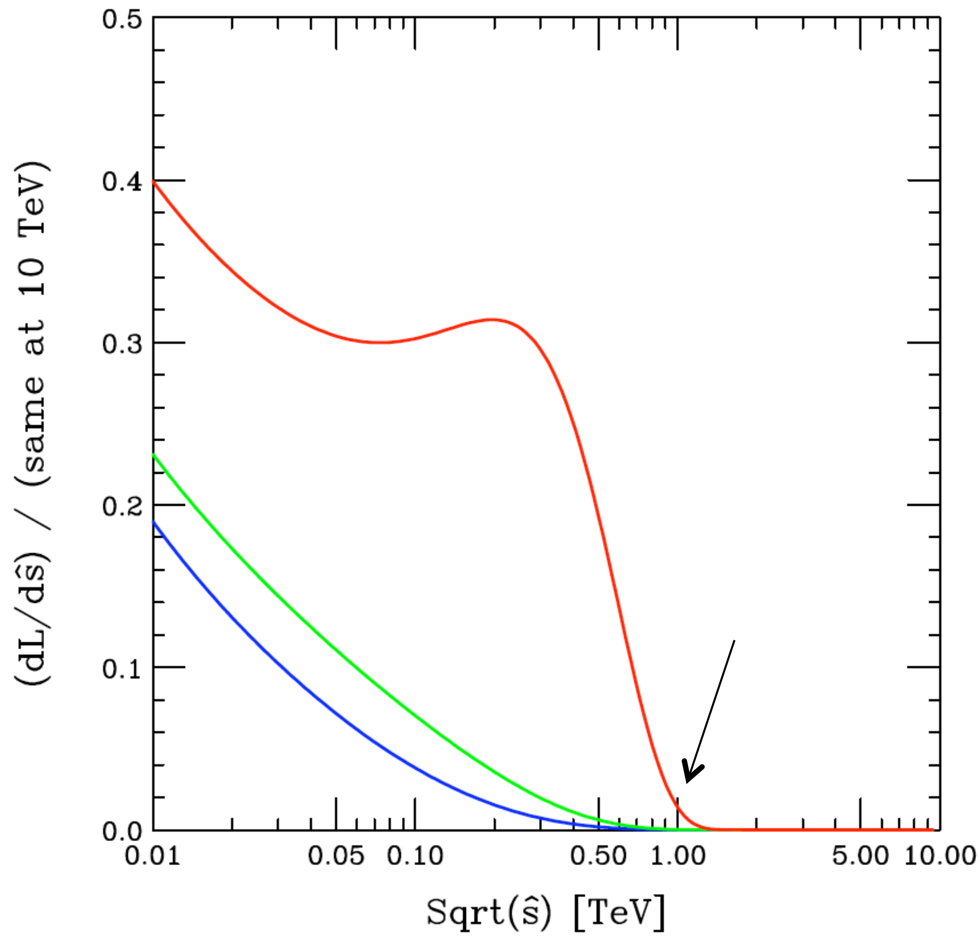


Figure 13: ($p\bar{p}$ at 1.96 TeV) / (pp at 10 TeV). luminosity integrated over y. Blue: gg; Green: $gq + g\bar{q}$; Red: $u\bar{u} + d\bar{d} + s\bar{s} + c\bar{c} + b\bar{b}$.

Precision benchmarks:
W/Z cross sections at the LHC

- CTEQ6.1 and MRST2004 NLO predictions in good agreement with each other
- NNLO corrections are small and negative
- NNLO mostly a K-factor; NLO predictions adequate for most predictions at the LHC

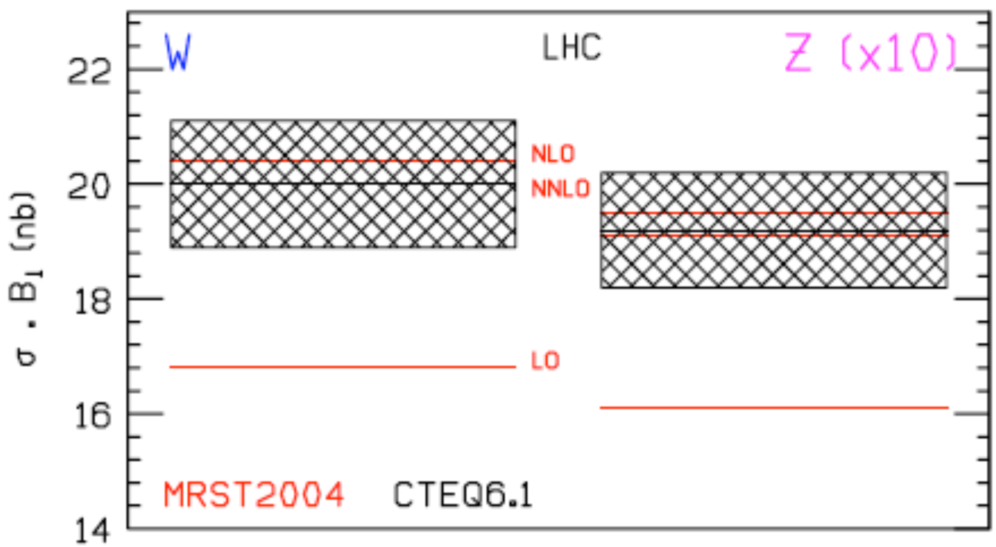


Figure 80. Predicted cross sections for W and Z production at the LHC using MRST2004 and CTEQ6.1 pdfs. The overall pdf uncertainty of the NLO CTEQ6.1 prediction is approximately 5%, consistent with figure 77.

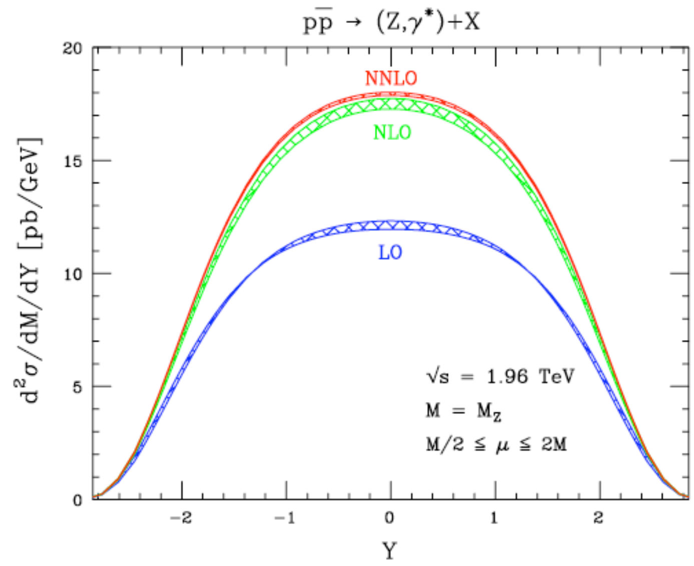
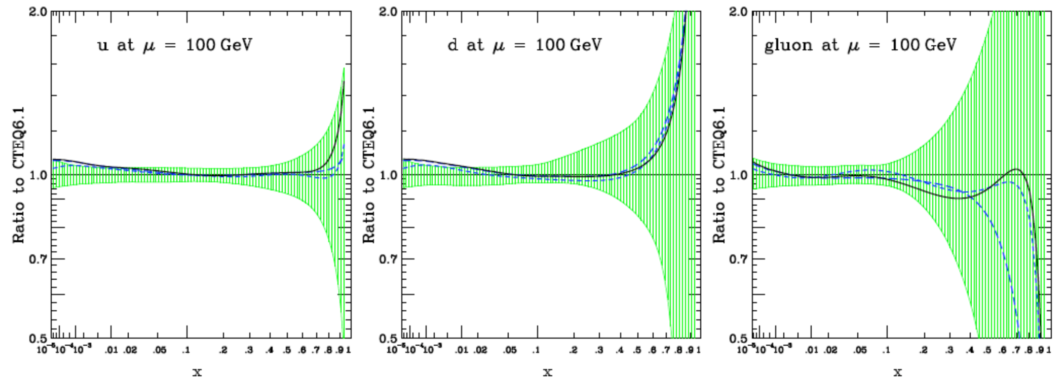


Figure 38. Predictions for the rapidity distribution of an on-shell Z boson in Run 2 at the Tevatron at LO, NLO and NNLO. The bands indicate the variation of the renormalization and factorization scales within the range $M_Z/2$ to $2M_Z$.

Heavy quark mass effects in global fits

- CTEQ6.1 (and previous generations of global fits) used zero-mass VFNS scheme
- With newer sets of pdf's (>=CTEQ6.5), heavy quark mass effects consistently taken into account in global fitting cross sections and in pdf evolution
- In most cases, resulting pdf's are within CTEQ6.1 pdf error bands
- But not at low x (in range of W and Z production at LHC)
- Heavy quark mass effects only appreciable near threshold
  - ex: prediction for F_2 at low x,Q at HERA smaller if mass of c,b quarks taken into account
  - thus, quark pdf's have to be bigger in this region to have an equivalent fit to the HERA data



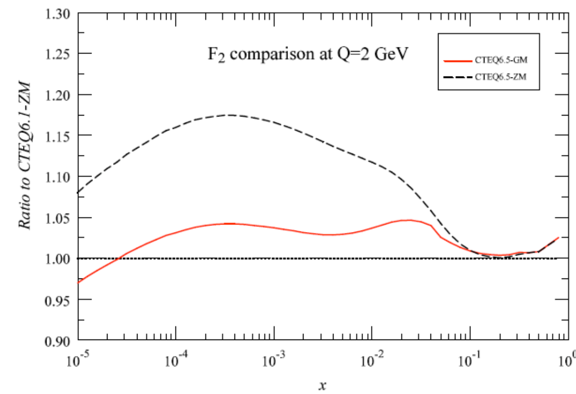
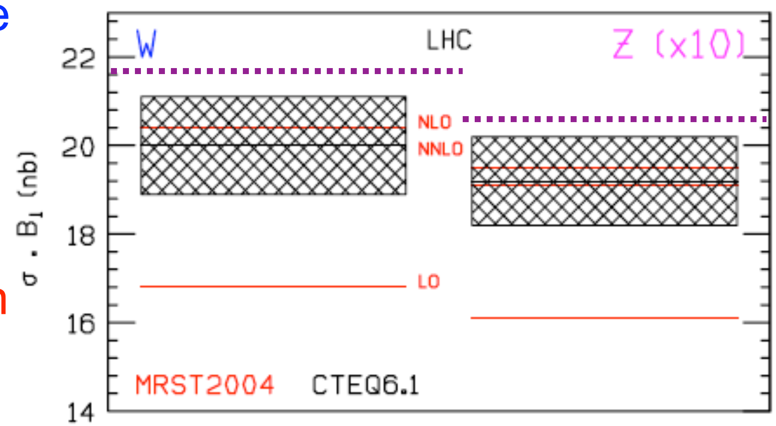


Figure 6: Comparison of theoretical calculations of F_2 using CTEQ6.1M in the ZM formalism (horizontal line of 1.00), CTEQ6.5M in the GM formalism (solid curve), and CTEQ6.5M in the ZM formalism (dashed curve).

implications for LHC phenomenology

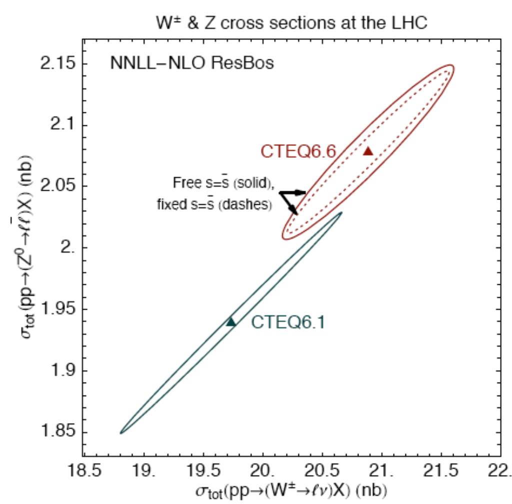
CTEQ6.5(6)

- Inclusion of heavy quark mass effects affects DIS data in x range appropriate for W/Z production at the LHC
- Cross sections for W/Z increase by 7-8%
  - now CTEQ and MRST2004 in disagreement, not a good sign for an important LHC benchmark
  - and relative uncertainties of W/Z increase
  - although individual uncertainties of W and Z decrease somewhat
- Two new free parameters in fit dealing with strangeness degrees of freedom so now have 44 error pdf's rather than 40



CTEQ6.5(6)

Figure 80. Predicted cross sections for W and Z production at the LHC using MRST2004 and CTEQ6.1 pdfs. The overall pdf uncertainty of the NLO CTEQ6.1 prediction is approximately 5%, consistent with figure 77.



Note importance of strange quark uncertainty for ratio

Figure 8: W & Z correlation ellipses at the LHC obtained in the fits with free and fixed strangeness.

…but

- Inclusion of heavy quark mass effects affects DIS data in x range appropriate for W/Z production at the LHC
- …but MSTW2008 also has increased W/Z cross sections at the LHC
  - now CTEQ6.6 and MSTW2008 in better agreement

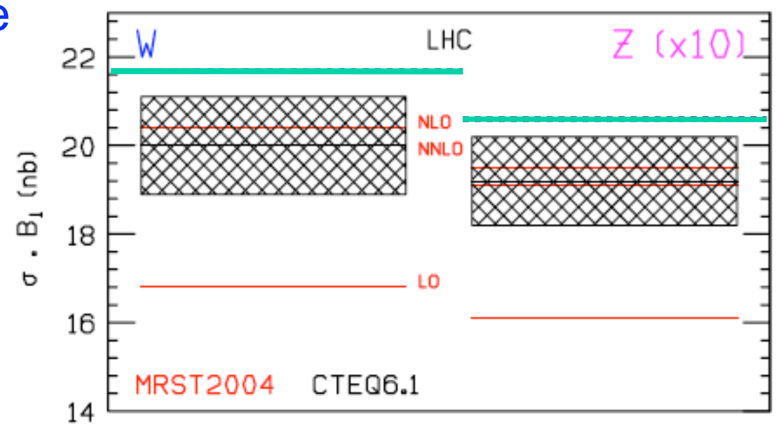


Figure 80. Predicted cross sections for W and Z production at the LHC using MRST2004 and CTEQ6.1 pdfs. The overall pdf uncertainty of the NLO CTEQ6.1 prediction is approximately 5%, consistent with figure 77.

CTEQ6.5(6)

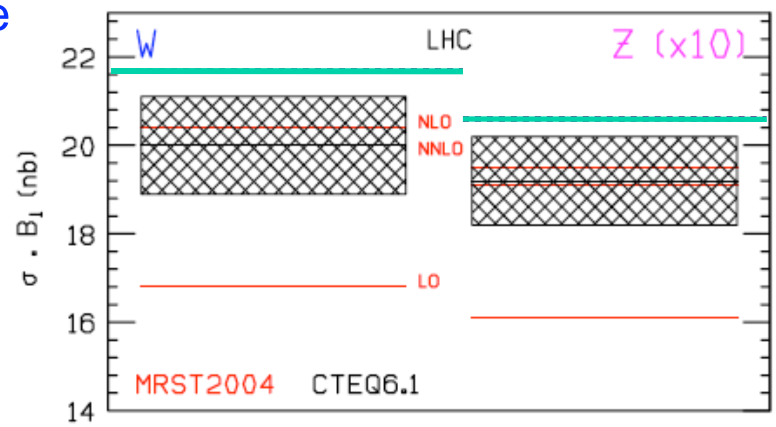
MSTW08

•more discussion at the end of the session
•Robert has prepared a few slides, especially about philosophy, i.e. what terms should be included at what order

PQCD Formulations with Heavy Quark Masses and Global Analysis.
R.S. Thorne (University Coll. London) , W.K. Tung (Michigan State U. & Washington U., Seattle) . Sep 2008. 18pp.
To appear in the proceedings of HERA and the LHC: 4th Workshop on the Implications of HERA for LHC Physics, Geneva, Switzerland, 26-30 May 2008.
e-Print: arXiv:0809.0714 [hep-ph]

…but

- Inclusion of heavy quark mass effects affects DIS data in x range appropriate for W/Z production at the LHC
- …but MSTW2008 with improved heavy quark mass scheme has also lead to increased W/Z cross sections at the LHC
  - now CTEQ6.6 and MSTW2008 in better agreement



CTEQ6.5(6)
MSTW08

Figure 80. Predicted cross sections for W and Z production at the LHC using MRST2004 and CTEQ6.1 pdfs. The overall pdf uncertainty of the NLO CTEQ6.1 prediction is approximately 5%, consistent with figure 77.

CMS expectation for 10 pb^{-1}
$\Delta\sigma/\sigma$(pp-> Z+X -> ee+X) = 1.9 (stat) ± 2.3 (syst) %
$\Delta\sigma/\sigma$(pp->W+X -> e+X) = 1.2 (stat) ± ~5 (syst) %

- Identification/reconstruction efficiency: ~1% from data
- Backgrounds: 5% (e)
- Theory (including acceptance) ~2% (PDFs, ISR)

ATLAS expectation for 50 pb^{-1}
$\Delta\sigma/\sigma$(pp-> Z+X -> μμ+X) = 0.8 (stat) ± 3.8 (syst) %
$\Delta\sigma/\sigma$(pp->W+X -> μ+X) = 0.2 (stat) ± 3.1 (syst) %

- Identification/reconstruction efficiency : 2-3%
- Backgrounds: <1% (muons)
- Theory (including acceptance) ~2% (PDFs, ISR)

F. Petrucchi
Moriond09

On a longer timescale:

CMS expectation for 1 fb^{-1}
$\Delta\sigma/\sigma$(pp-> Z+X -> μμ+X) = 0.13 (stat) ± 2.3 (syst) %
$\Delta\sigma/\sigma$(pp->W+X -> μ+X) = 0.04 (stat) ± 3.3 (syst) %

ATLAS expectation for 1 fb^{-1}
$\Delta\sigma/\sigma$(pp-> Z+X -> ee+X)= 0.20 (stat) ± 2.4 (syst) %
$\Delta\sigma/\sigma$(pp->W+X -> e+X) = 0.04 (stat) ± 2.5 (syst) %

- Eff unc. <1% with data-driven methods
- Background reduced with selections

PDF correlations

- Consider a cross section X(a), a function of the Hessian eigenvectors
- i^{th} component of gradient of X is

$$\frac{\partial X}{\partial a_i} \equiv \partial_i X = \frac{1}{2}(X_i^{(+)} - X_i^{(-)})$$

- Now take 2 cross sections X and Y
  - or one or both can be pdf's
- Consider the projection of gradients of X and Y onto a circle of radius 1 in the plane of the gradients in the parton parameter space
- The circle maps onto an ellipse in the XY plane
- The angle f between the gradients of X and Y is given by

$$\cos\varphi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X\, \Delta Y}\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)\left(Y_i^{(+)} - Y_i^{(-)}\right)$$

- The ellipse itself is given by

$$\left(\frac{\delta X}{\Delta X}\right)^2 + \left(\frac{\delta Y}{\Delta Y}\right)^2 - 2\left(\frac{\delta X}{\Delta X}\right)\left(\frac{\delta Y}{\Delta Y}\right)\cos\varphi = \sin^2\varphi$$

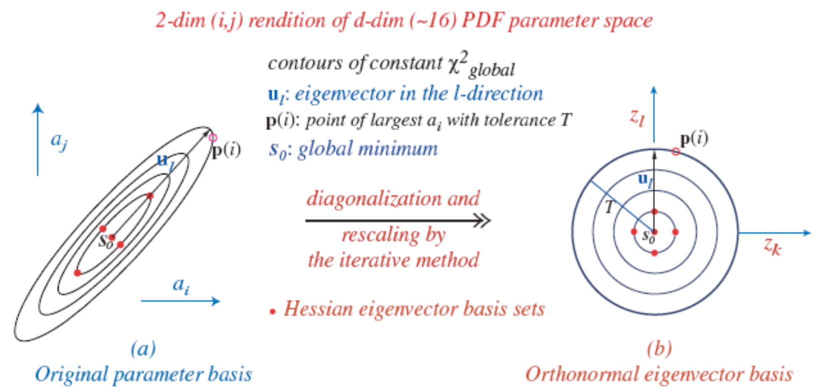


Figure 28. A schematic representation of the transformation from the pdf parameter basis to the orthonormal eigenvector basis.

•If two cross sections are very correlated, then cosf~1
•…uncorrelated, then cosf~0
•…anti-correlated, then cosf~-1

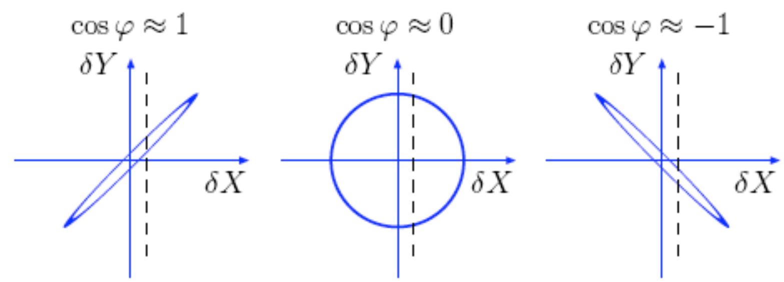


Figure 1: Dependence on the correlation ellipse formed in the $\Delta X - \Delta Y$ plane on the value of the correlation cosine $\cos\varphi$.

Correlations with Z, tT

Define a correlation cosine between two quantities
(see extra slides for more detail)

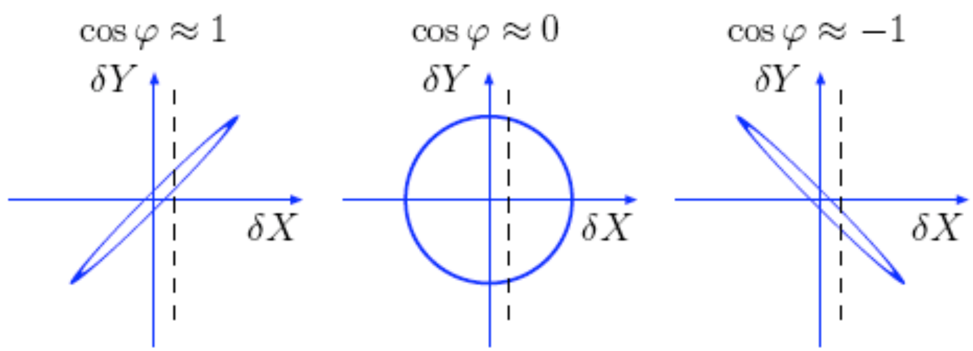
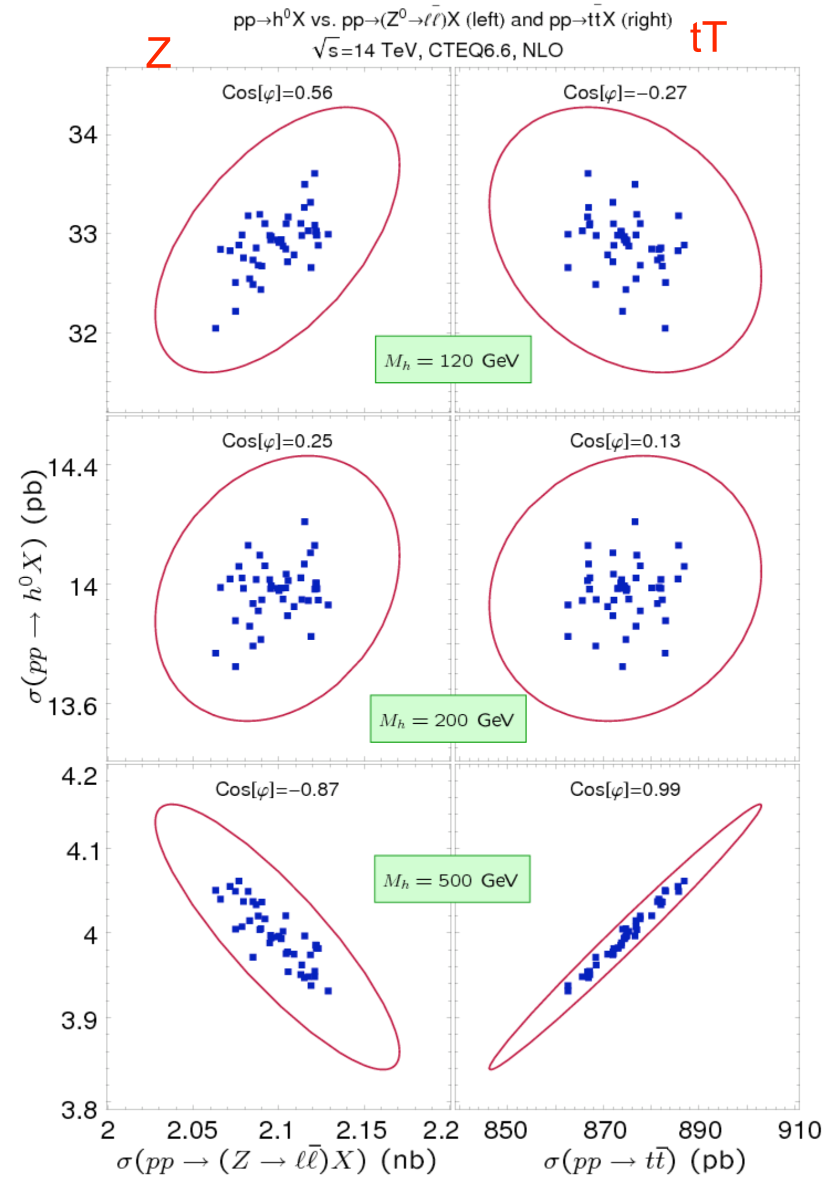


Figure 1: Dependence on the correlation ellipse formed in the $\Delta X - \Delta Y$ plane on the value of the correlation cosine $\cos\varphi$.



- If two cross sections are very correlated, then cosf~1
- …uncorrelated, then cosf~0
- …anti-correlated, then cosf~-1

Correlations with Z, tT

Define a correlation cosine between two quantities

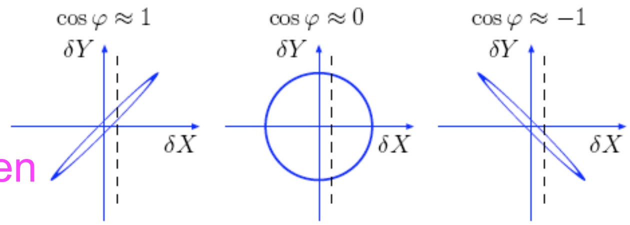
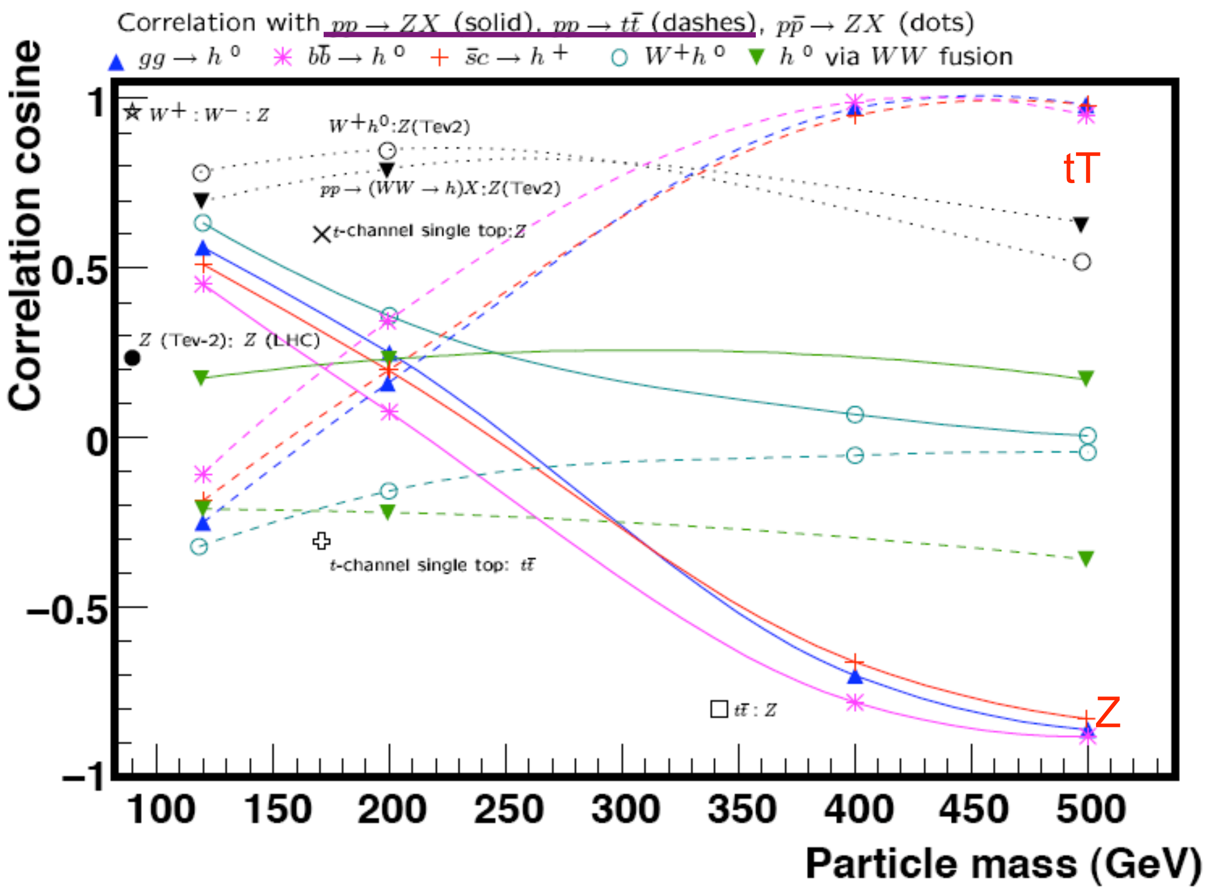


Figure 1: Dependence on the correlation ellipse formed in the $\Delta X - \Delta Y$ plane on the value of the correlation cosine $\cos\varphi$.

- If two cross sections are very correlated, then cosf~1
- …uncorrelated, then cosf~0
- …anti-correlated, then cosf~-1



- Note that correlation curves to Z and to tT are mirror images of each other
- By knowing the pdf correlations, can reduce the uncertainty for a given cross section in ratio to a benchmark cross section **iff** cos f > 0;e.g. $D(s_{W^+}/s_Z)$~1%
- If cos f < 0, pdf uncertainty for one cross section normalized to a benchmark cross section is larger
- So, for gg->H(500 GeV); pdf uncertainty is 4%; $D(s_H/s_Z)$~8%

W/Z summary so far

- We will use W and Z cross sections as luminosity normalizations in early running and perhaps always
  - because integrated luminosity is not going to be known much better than 15-20% at first and maybe never better than 5-10%
- The pdf uncertainty for the ratio of a cross section that proceeds with a qQ initial state to the W/Z cross section is significantly reduced
- The pdf uncertainty for the ratio of a cross section that proceeds with a gg initial state to the W/Z cross section is significantly increased
- Would it be reasonable to use tT production as an *additional* benchmark?
  - yeah, yeah I know it's difficult, and it won't happen early, but just keep it in mind

PDF progress from CTEQ

- NLO updates (CT09G)
  - using new Tevatron Run 2 data, concentrating on jets
  - some new eigenvector tools
  - paper should be out within a week or so
- Combined fits (q_T+x)
  - useful for precision physics such as W mass determination
- Mod LO pdf's
  - for use in parton shower Monte Carlos at the LHC
  - some discussion in extra slides
- NNLO pdf's
  - precision physics at the LHC
  - HOPPET used for evolution

NLO fits

- 37 data sets with 2898 data points
  - chisquare=2756
  - full correlated experimental errors used for all data sets that report such errors
- Gluon parametrization

$$g(x,\mu_0) = a_0\, x^{a_1}\, (1-x)^{a_2}\, \exp(a_3 x \,+\, a_4 x^2 \,+\, a_5 \sqrt{x})$$

  - more general than what was used in CTEQ66
  - crucial to have flexible parametrization to correctly calculate uncertainties
  - have to control instabilities caused by numerical evaluation of second deriviatives of the Hessian
  - now 24 free parameters
- Have added a penalty to chisquare that rises as the 4th power to prevent large contributions from any particular experiment
  - this will be more crucial for eigenvector sets
- CTEQ66 pdf's known to describe Run 2 data well, so don't expect too much change with their inclusion in the fit
  - chisquare decreases to 2740, a reduction of 16
  - only significant change is in the gluon sector

Tension

- Is there a tension between the Run 1 and Run 2 data?
- Have to ask some crucial questions when adding new data sets, in this case the Run 2 jet data

1. Are the new data consistent with theory?
2. Are the new data consistent with the previous experiments?
3. Are the new data sets consistent with each other?
4. Do the new data sets provide significant new constraints?

1: are the data sets consistent with theory?

- Data sets are consistent with theory if they lie (roughly) within sqrt(2N) of expected chisquare
- CDF1 is outside but due to fluctuations in a few data points
- D01 is consistent with theory; in fact, maybe too consistent; systematic errors may be overestimated
- CDFII consistent with theory
- D0II consistent with theory

CDF_I (33 pts) Wt	χ^2	$D0_I$ (90 pts) Wt	χ^2	CDF_{II} (72 pts) Wt	χ^2	$D0_{II}$ (110 pts) Wt	χ^2	$\Delta\chi^2$ non-jet
0	55.4	0	115.3	0	99.5	0	134.0	0.0
1	52.6	1	47.0	0	105.6	0	138.3	11.8
0	56.6	0	82.2	1	85.6	1	124.1	6.2
1	52.1	1	59.4	1	88.5	1	121.5	9.6
0	58.4	0	60.9	10	79.6	10	120.4	39.9
1	54.8	1	58.8	10	80.3	10	120.0	39.4
10	54.1	10	35.6	0	112.9	0	156.7	24.1
10	53.1	10	38.6	1	102.6	1	142.3	21.9
10	51.6	10	49.7	10	82.8	10	120.9	39.6
10	49.5	0	73.5	0	110.4	0	125.3	12.5
50	47.3	0	74.0	0	123.9	0	139.3	80.5
0	58.6	10	32.1	0	122.7	0	172.2	25.2
0	66.8	50	30.6	0	140.0	0	189.1	58.6
1	59.6	1	67.5	10	75.2	1	130.9	32.0
1	63.4	1	70.4	50	71.6	1	140.0	92.9
1	50.6	1	60.0	1	93.0	10	116.5	20.6
1	50.5	1	61.6	1	96.6	50	112.6	113.8

Table 1: χ^2 for jet experiments with various weights

2: are the new data consistent with the previous data?

- Chisquare for non-jet data forced to increase by only 10 to accommodate the 4 jet experiments at weight 1 and by 40 to accommodate them at weight 10

CDF_I (33 pts) Wt	χ^2	$D0_I$ (90 pts) Wt	χ^2	CDF_{II} (72 pts) Wt	χ^2	$D0_{II}$ (110 pts) Wt	χ^2	$\Delta\chi^2$ non-jet
0	55.4	0	115.3	0	99.5	0	134.0	0.0
1	52.6	1	47.0	0	105.6	0	138.3	11.8
0	56.6	0	82.2	1	85.6	1	124.1	6.2
1	52.1	1	59.4	1	88.5	1	121.5	9.6
0	58.4	0	60.9	10	79.6	10	120.4	39.9
1	54.8	1	58.8	10	80.3	10	120.0	39.4
10	54.1	10	35.6	0	112.9	0	156.7	24.1
10	53.1	10	38.6	1	102.6	1	142.3	21.9
10	51.6	10	49.7	10	82.8	10	120.9	39.6
10	49.5	0	73.5	0	110.4	0	125.3	12.5
50	47.3	0	74.0	0	123.9	0	139.3	80.5
0	58.6	10	32.1	0	122.7	0	172.2	25.2
0	66.8	50	30.6	0	140.0	0	189.1	58.6
1	59.6	1	67.5	10	75.2	1	130.9	32.0
1	63.4	1	70.4	50	71.6	1	140.0	92.9
1	50.6	1	60.0	1	93.0	10	116.5	20.6
1	50.5	1	61.6	1	96.6	50	112.6	113.8

Table 1: χ^2 for jet experiments with various weights

3: are the four jet experiments consistent with each other?

- The two Run II expts are consistent with each other; the chisquares for CDFII and D0II are not strongly dependent on the weight given to the other experiment
- The Run I and II experiments are consistent with each other; raising the weights of the Run I experiments improves the Run I fits, while resulting in small changes to the Run II jet data and non-jet data

$\mathrm{CDF_I}$ (33 pts) Wt	χ^2	$\mathrm{D0_I}$ (90 pts) Wt	χ^2	$\mathrm{CDF_{II}}$ (72 pts) Wt	χ^2	$\mathrm{D0_{II}}$ (110 pts) Wt	χ^2	$\Delta\chi^2$ non-jet
0	55.4	0	115.3	0	99.5	0	134.0	0.0
1	52.6	1	47.0	0	105.6	0	138.3	11.8
0	56.6	0	82.2	1	85.6	1	124.1	6.2
1	52.1	1	59.4	1	88.5	1	121.5	9.6
0	58.4	0	60.9	10	79.6	10	120.4	39.9
1	54.8	1	58.8	10	80.3	10	120.0	39.4
10	54.1	10	35.6	0	112.9	0	156.7	24.1
10	53.1	10	38.6	1	102.6	1	142.3	21.9
10	51.6	10	49.7	10	82.8	10	120.9	39.6
10	49.5	0	73.5	0	110.4	0	125.3	12.5
50	47.3	0	74.0	0	123.9	0	139.3	80.5
0	58.6	10	32.1	0	122.7	0	172.2	25.2
0	66.8	50	30.6	0	140.0	0	189.1	58.6
1	59.6	1	67.5	10	75.2	1	130.9	32.0
1	63.4	1	70.4	50	71.6	1	140.0	92.9
1	50.6	1	60.0	1	93.0	10	116.5	20.6
1	50.5	1	61.6	1	96.6	50	112.6	113.8

Table 1: χ^2 for jet experiments with various weights

4: will the Run II jet data reduce the gluon uncertainty?

- Error eigenvector predictions from CTEQ6.6 lie outside results with weight 0 for the Run II jet data, so expect only small reduction in the uncertainty as a result of including the new data

CDF_I (33 pts) Wt	χ^2	$D0_I$ (90 pts) Wt	χ^2	CDF_{II} (72 pts) Wt	χ^2	$D0_{II}$ (110 pts) Wt	χ^2	$\Delta\chi^2$ non-jet
0	55.4	0	115.3	0	99.5	0	134.0	0.0
1	52.6	1	47.0	0	105.6	0	138.3	11.8
0	56.6	0	82.2	1	85.6	1	124.1	6.2
1	52.1	1	59.4	1	88.5	1	121.5	9.6
0	58.4	0	60.9	10	79.6	10	120.4	39.9
1	54.8	1	58.8	10	80.3	10	120.0	39.4
10	54.1	10	35.6	0	112.9	0	156.7	24.1
10	53.1	10	38.6	1	102.6	1	142.3	21.9
10	51.6	10	49.7	10	82.8	10	120.9	39.6
10	49.5	0	73.5	0	110.4	0	125.3	12.5
50	47.3	0	74.0	0	123.9	0	139.3	80.5
0	58.6	10	32.1	0	122.7	0	172.2	25.2
0	66.8	50	30.6	0	140.0	0	189.1	58.6
1	59.6	1	67.5	10	75.2	1	130.9	32.0
1	63.4	1	70.4	50	71.6	1	140.0	92.9
1	50.6	1	60.0	1	93.0	10	116.5	20.6
1	50.5	1	61.6	1	96.6	50	112.6	113.8

Table 1: χ^2 for jet experiments with various weights

Restricted parametrization

- Set a_4=a_5=0
- The chisquare for the non-jet data becomes bad for weights on the jet data of 10
- The jet data are in conflict with the non-jet data (using this param)

$$g(x,\mu_0) = a_0\, x^{a_1}\, (1-x)^{a_2}\, \exp(a_3 x + a_4 x^2 + a_5 \sqrt{x})$$

$\mathrm{CDF_I}$ (33 pts) Wt	χ^2	$\mathrm{D0_I}$ (90 pts) Wt	χ^2	$\mathrm{CDF_{II}}$ (72 pts) Wt	χ^2	$\mathrm{D0_{II}}$ (110 pts) Wt	χ^2	$\Delta\chi^2$ non-jet
0	55.8	0	145.9	0	120.6	0	155.2	2.0
1	53.2	1	124.0	0	118.2	0	148.6	7.7
0	58.6	0	121.3	1	98.1	1	137.8	16.8
1	54.5	1	108.8	1	95.5	1	134.2	25.8
10	54.1	10	75.7	0	142.0	0	152.1	184.3
10	51.9	10	74.0	1	101.6	1	134.6	185.5
0	67.1	0	75.3	10	77.3	10	126.1	114.6
1	60.3	1	74.1	10	77.1	10	125.8	119.3
10	51.7	10	64.5	10	76.2	10	126.1	204.3

New pdf's (CT09G)

- Somewhat of a reduction in gluon uncertainty for low Q, but very similar to CTEQ6.6 at high Q
- At large scales, the gluon distributions are very similar

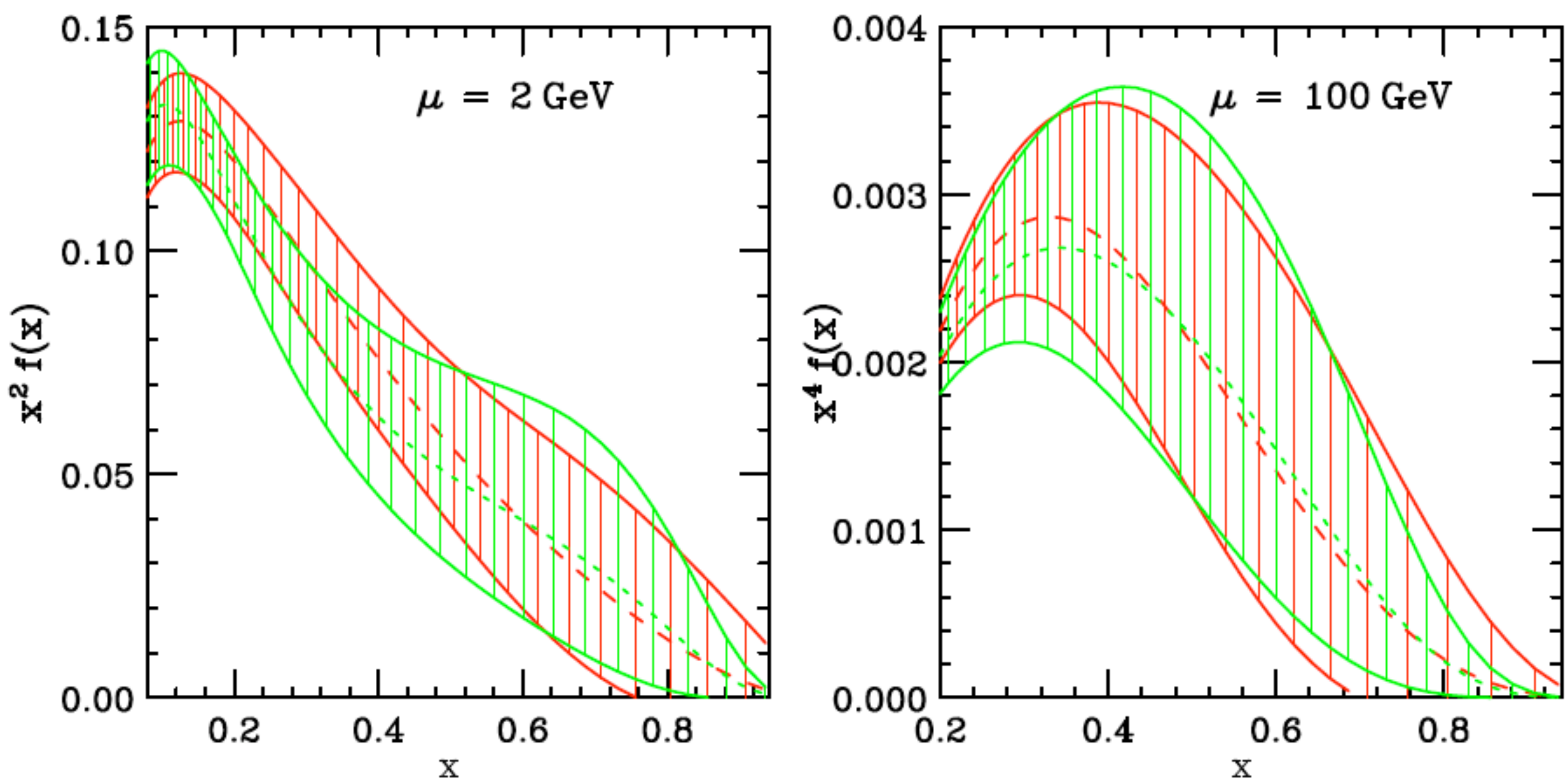


Figure 8: Gluon distributions and uncertainties in CT09G (red) and CTEQ6.6 (green).

New CTEQ technique

- With Hessian method, diagonalize the Hessian matrix to determine orthonormal eigenvector directions; 1 eigenvector for each free parameter in the fit
  - CTEQ6.6 has 22 free parameters, so 22 eigenvectors and 44 error pdf's
  - new NLO pdf's will have 24 free parameters
- Each eigenvector/error pdf has components from each of the free parameters
- Sum over all error pdf's to determine the error for any observable
- But,we are free to make an additional orthogonal transformation that diagonalizes one additional quantity *G*

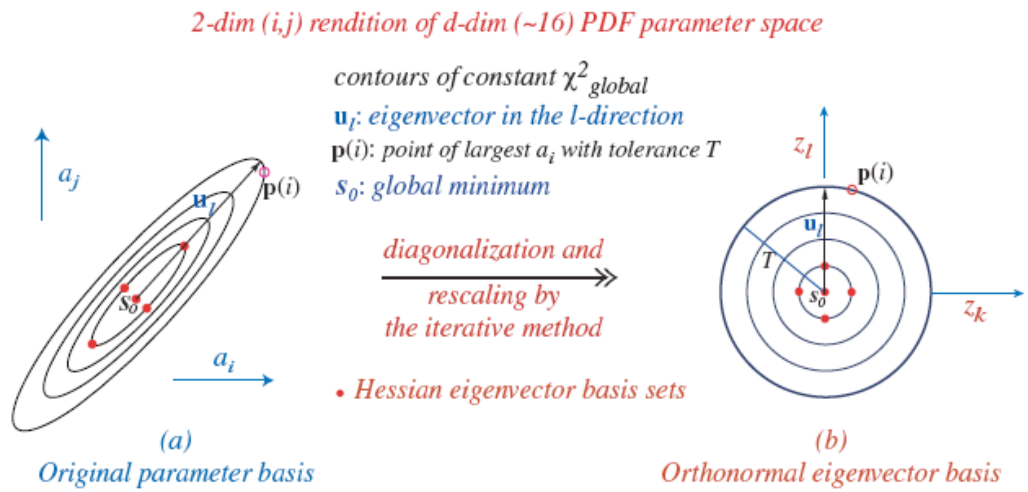


Figure 28. A schematic representation of the transformation from the pdf parameter basis to the orthonormal eigenvector basis.

- In these new coordinates, variation in a given quantity is now given by one or a few eigenvectors, rather than by all 44 (or however many)
- *G* may be the W cross section, or the W rapidity distribution or a Higgs cross section, depending on how clever one wants to be
- In principle these principal error pdf's could be provided as well, for example in CTEQ4LHC ntuples (see later)

Random pdf sets

- Generate random collection of pdf's that lie at the edge of the acceptable chisquare range
- Envelope of random sets covers a much smaller range than the full uncertainty, even though every one of the sets is at the upper limit of acceptable chisquare
- Extreme g(x) at any given x can be thought of as corresponding to a specific direction in the N-dimensional parameter space; and the probability distribution for the component of a random vector along any particular direction z is $dP/dz \sim (1-z^2)^{(N-3)/2}$
- Extremely small probability that a random sampling will find a pdf near the extreme z~1

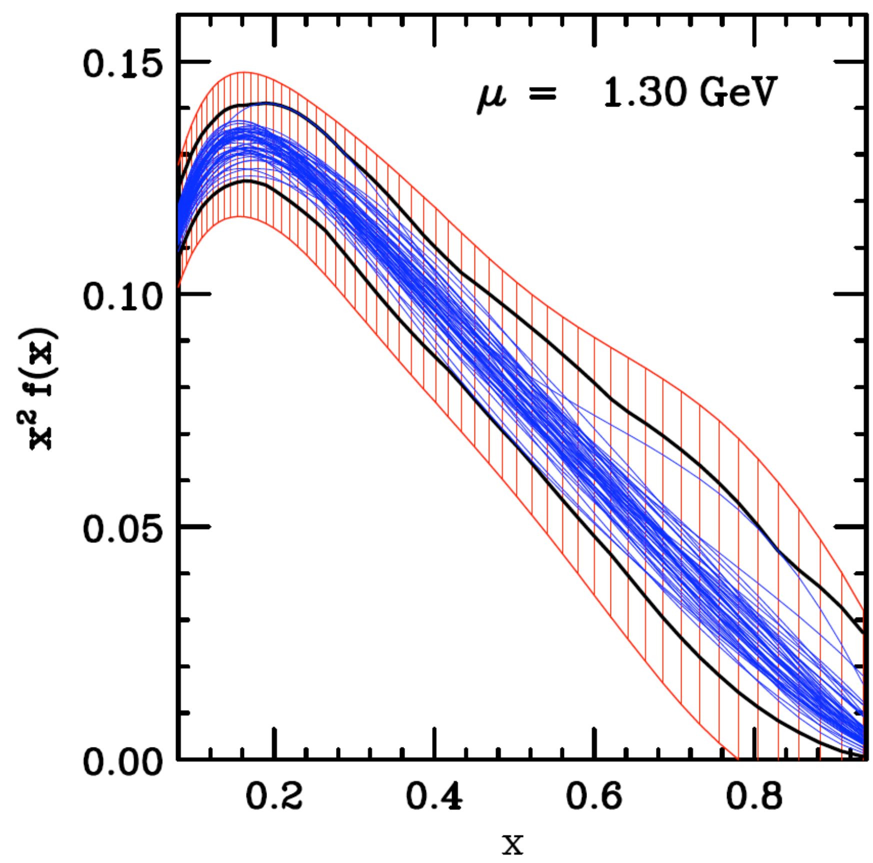


Figure 7: Gluon uncertainty from 50 random PDF sets with $\Delta\chi^2 = 100$ (blue curves); envelope of 500 such sets (black curves); full uncertainty range by Hessian method (red shaded region).

CTEQ4LHC/FROOT

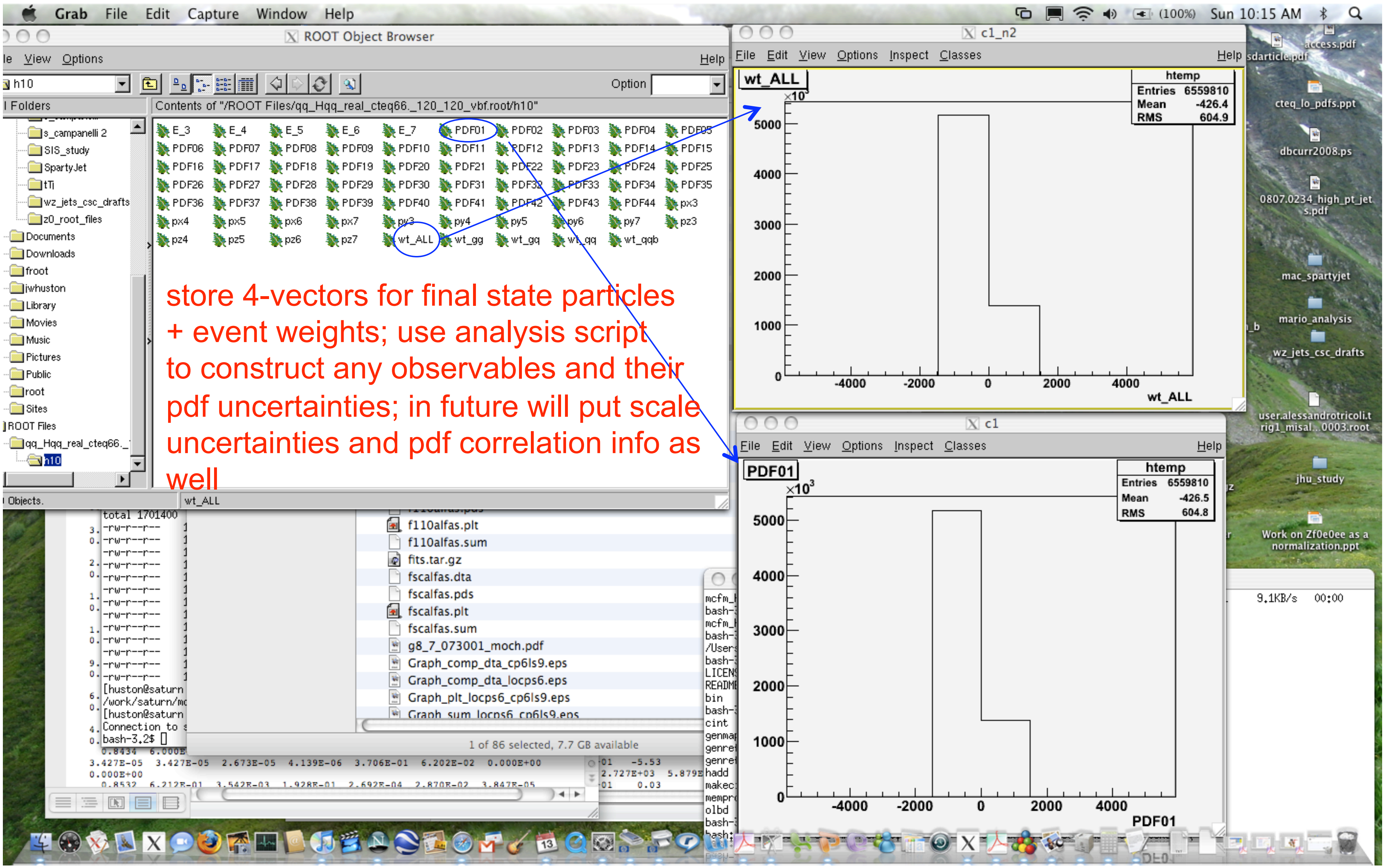
- Collate/create cross section predictions for LHC
  - processes such as W/Z/ Higgs(both SM and BSM)/ diboson/tT/single top/photons/ jets…
  - at LO, NLO, NNLO (where available)
    - new: W/Z production to NNLO QCD and NLO EW
  - pdf uncertainty, scale uncertainty, correlations
  - impacts of resummation (q_T and threshold)
- As prelude towards comparison with actual data
- Using programs such as:
  - MCFM
  - ResBos
  - Pythia/Herwig/Sherpa
  - … private codes with CTEQ
- First on webpage and later as a report

Primary goal: have all theorists (**including you**) write out parton level output into ROOT ntuples
Secondary goal: make libraries of prediction ntuples available

- FROOT: a simple interface for writing Monte-Carlo events into a ROOT ntuple file
- Written by Pavel Nadolsky (nadolsky@physics.smu.edu)

- CONTENTS
- ========
- froot.c -- the C file with FROOT functions
- taste_froot.f -- a sample Fortran program writing 3 events into a ROOT ntuple
- taste_froot0.c -- an alternative top-level C wrapper (see the compilation notes below)
- Makefile

MCFM 5.3 has FROOT built in



store 4-vectors for final state particles + event weights; use analysis script to construct any observables and their pdf uncertainties; in future will put scale uncertainties and pdf correlation info as well prototype webpage

http://www.pa.msu.edu/~huston/cteq4lhc/higgs/cteq4lhc_higgs.html

cteq4lhc_higgs

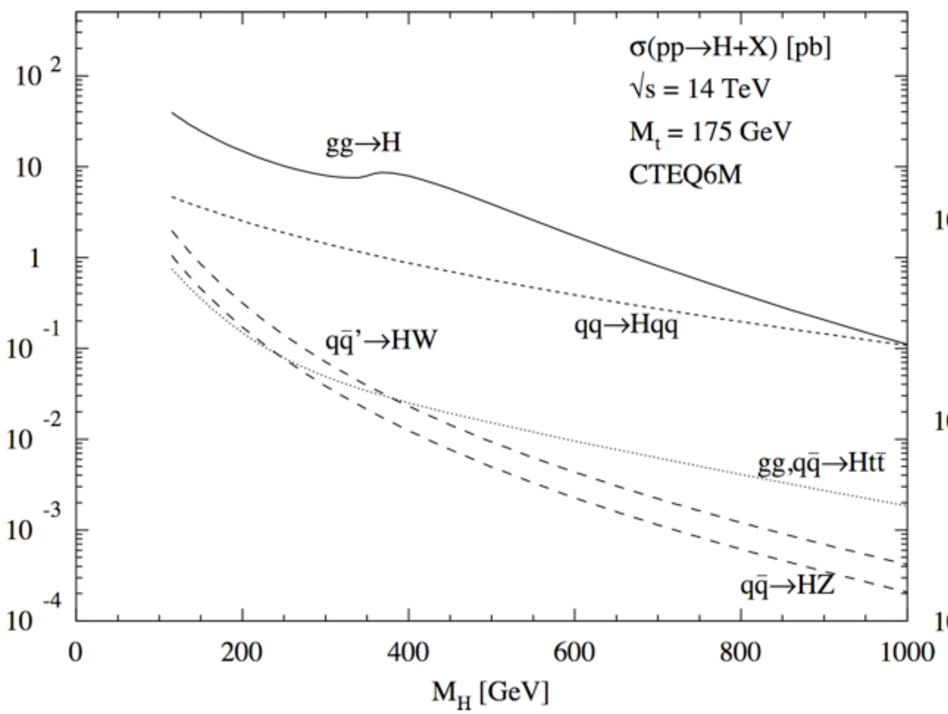
http://www.pa.msu.edu/~huston/cteq4lhc/higgs/cteq4lhc_higgs.html

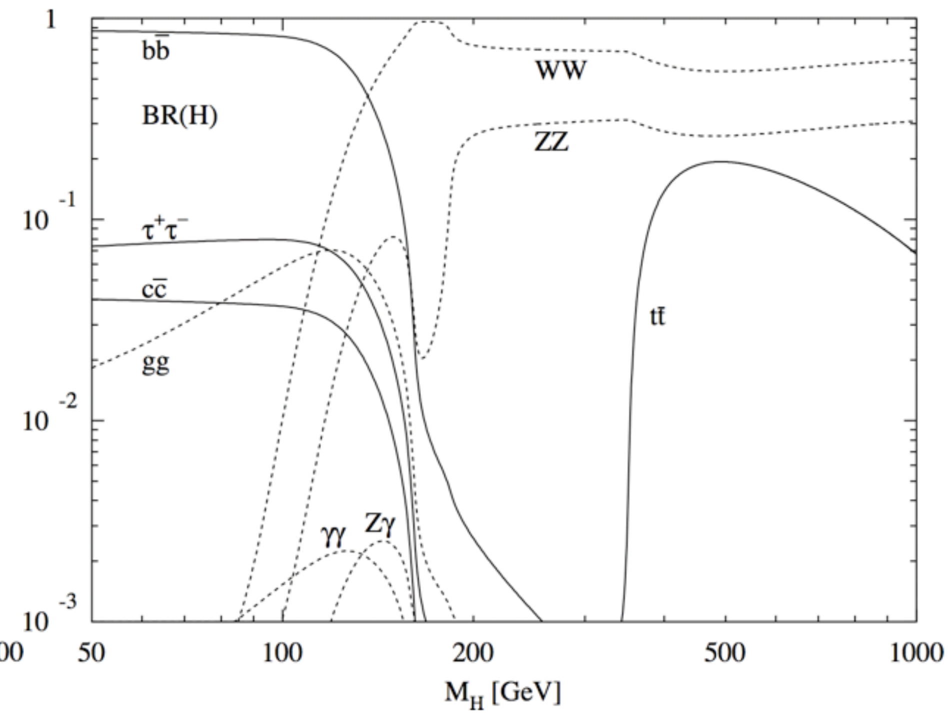
Amazon A Single Thread Les Houches ...V Colliders MTA SZTAKI: ... Dictionary CSCNotesLis...las < TWiki PatVancouve...las < TWiki Quick guide...nda monitor Quick guide...nda monitor

Standard Model Higgs Production at the LHC

The Standard Model inclusive Higgs cross section is known at LO, NLO and NNLO. The Higgs tranverse momentum distribution has been calculated to to NNLO+NNLL. The link to the discussion of Higgs production in CHS can be found here.

The cross section for Higgs production at NLO, using CTEQ6M pdfs, as a function of its mass is shown below. The largest production mechanism is gg fusion, through a top quark loop. The branching ratios for the Standard Model Higgs decay, as a function of its mass are also shown below.





The plan (for Higgs for ATLAS/CMS)

- SM Higgs cross sections with production by
  - gg fusion
  - VBF
  - associated production
- For masses of
  - 115,125,150,175,200,250,300,400,500,600,700,800
- The Higgs decays into a tau pair
- At center-of-mass energies of 10 and 14 TeV
- So far only with MCFM, but there are private codes this can be done for such as tTh, which are in progress
- Ketevi and Aleandro suggest A/H and tbH$^+$ as well; have to get those authors to write the ntuples
- Can also do for NNLO Higgs

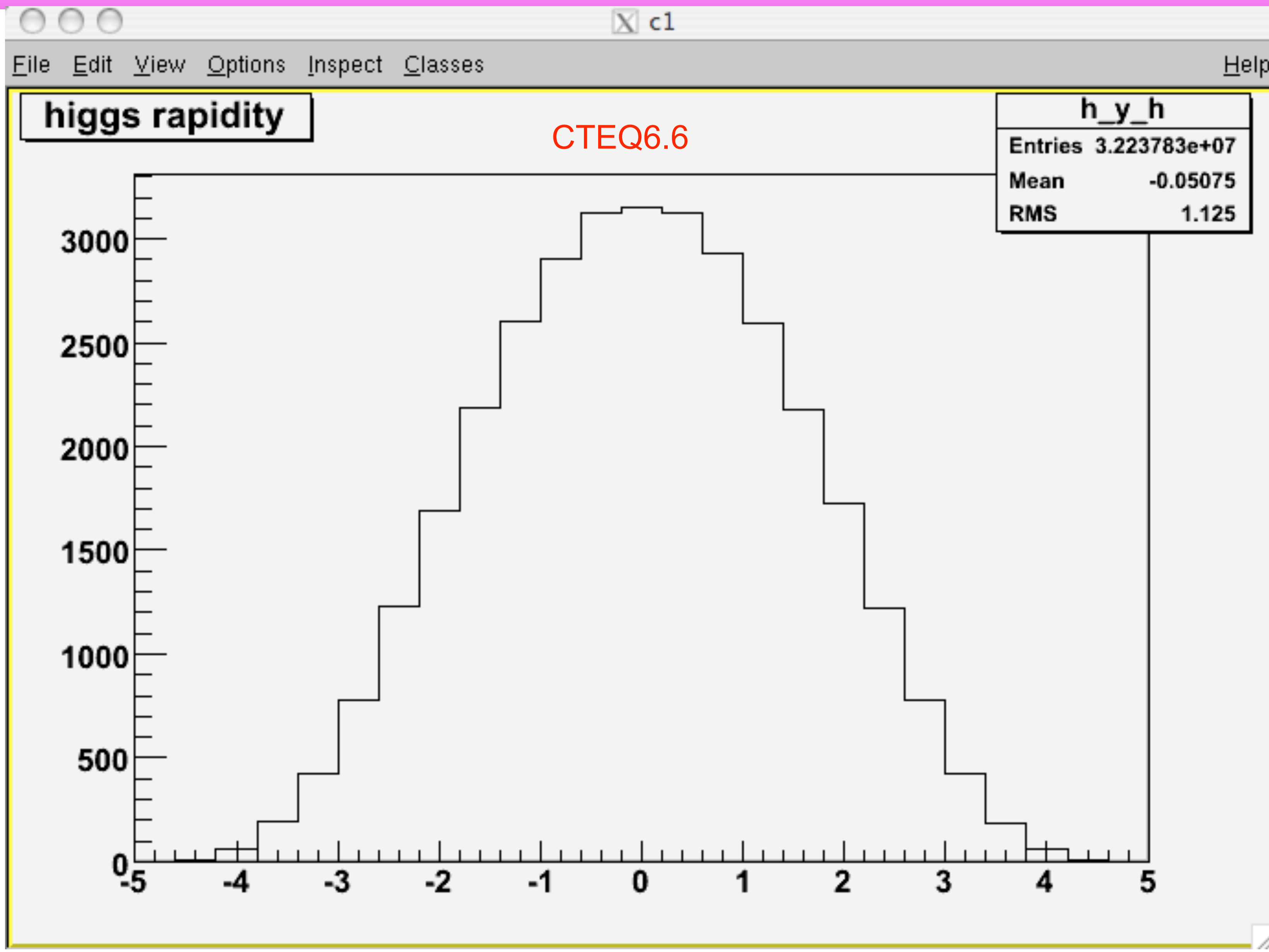
- So far I've done the gg fusion and VBF at 14 TeV
- Is it worthwhile creating for 10 TeV?
- Will we actually be running at 14-2*epsilon TeV in 2011?
- Is this serving a need? Will anyone use it? I have heard the comment that it would be useful to have 1 webpage that contains as much useful Higgs cross section info as possible
- Of course, we will also encourage the private codes to be made public so that users can run themselves…but as you know, making a correct version of the code and making a public version are two different things

Example:gg->Higgs (125 GeV)

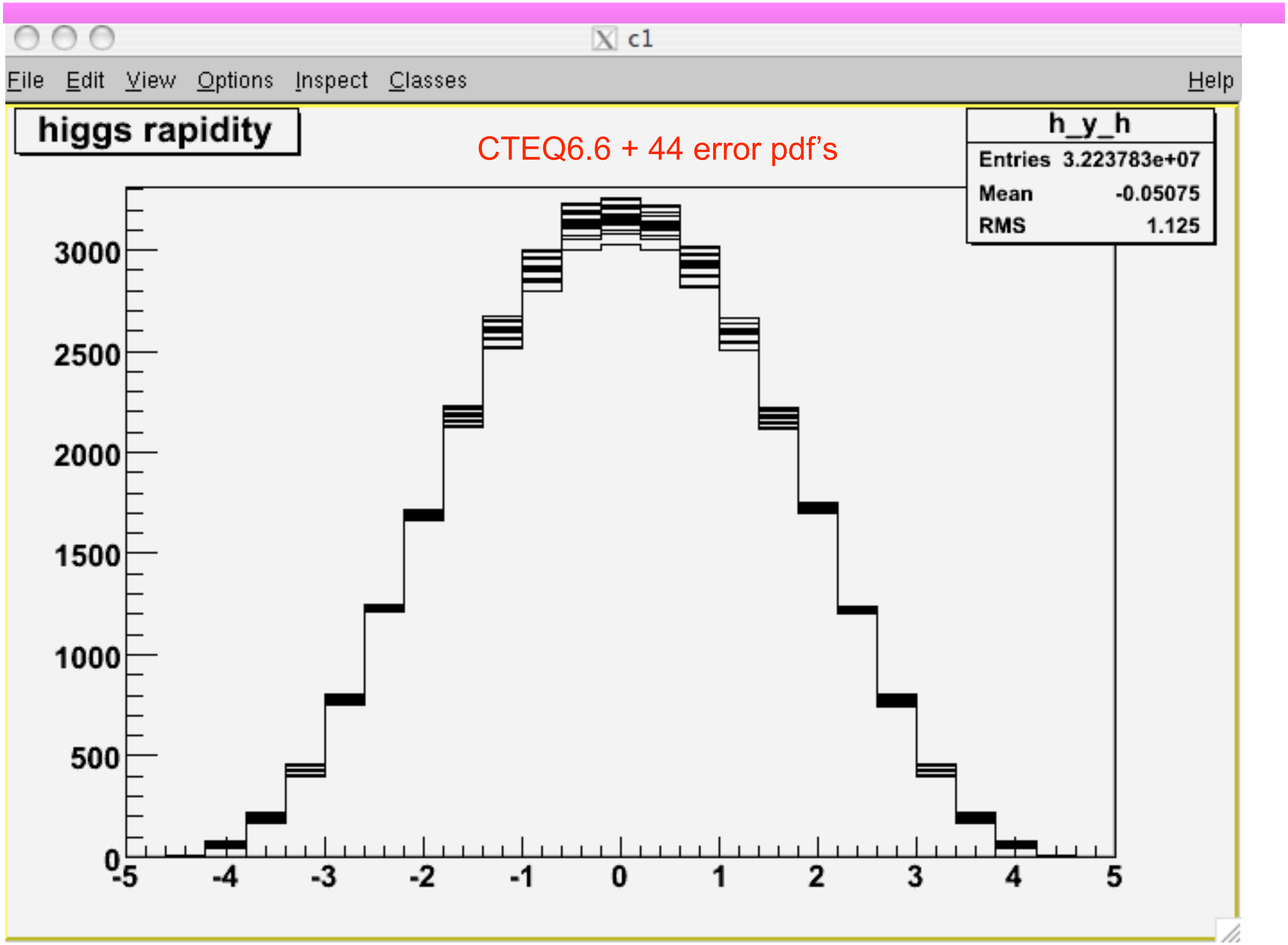
6.6 GB total for real+virtual

```
[huston@saturn 125_E6]$ ls -l
total 6474536
-rw-r--r--    1 huston   users    1900836256 Mar  9 16:00 ggfus0_real_cteq66._125_125_125_
lhc_1.root
-rw-r--r--    1 huston   users    1440740081 Mar  9 19:31 ggfus0_real_cteq66._125_125_125_
lhc_2.root
-rw-r--r--    1 huston   users       38165 Mar  9 19:31 ggfus0_real_cteq66._125_125_125_lh
c.dat
-rw-r--r--    1 huston   users      120912 Mar  9 19:31 ggfus0_real_cteq66._125_125_125_lh
c_error.top
-rw-r--r--    1 huston   users    1902099628 Mar  9 11:46 ggfus0_real_cteq66._125_125_125_
lhc.root
-rw-r--r--    1 huston   users       61186 Mar  9 19:31 ggfus0_real_cteq66._125_125_125_lh
c.top
-rw-r--r--    1 huston   users       27133 Mar 10 11:22 ggfus0_virt_cteq66._125_125_125_lh
c.dat
-rw-r--r--    1 huston   users       19768 Mar 10 11:22 ggfus0_virt_cteq66._125_125_125_lh
c_error.top
-rw-r--r--    1 huston   users    1379302499 Mar 10 11:23 ggfus0_virt_cteq66._125_125_125_
lhc.root
-rw-r--r--    1 huston   users       31156 Mar 10 11:22 ggfus0_virt_cteq66._125_125_125_lh
c.top
-rw-r--r--    1 huston   users       43399 Mar  9 19:31 higgs_125_1E6_real.log
-rw-r--r--    1 huston   users       41181 Mar 10 11:22 higgs_125_1E6_virt.log
-rw-r--r--    1 huston   users       26696 Mar 12 14:07 mcfm_histograms.root
-rw-r--r--    1 huston   users        9360 Mar 12 13:20 read.cc
[huston@saturn 125_E6]$
```

Output plots



Output plots



Summary

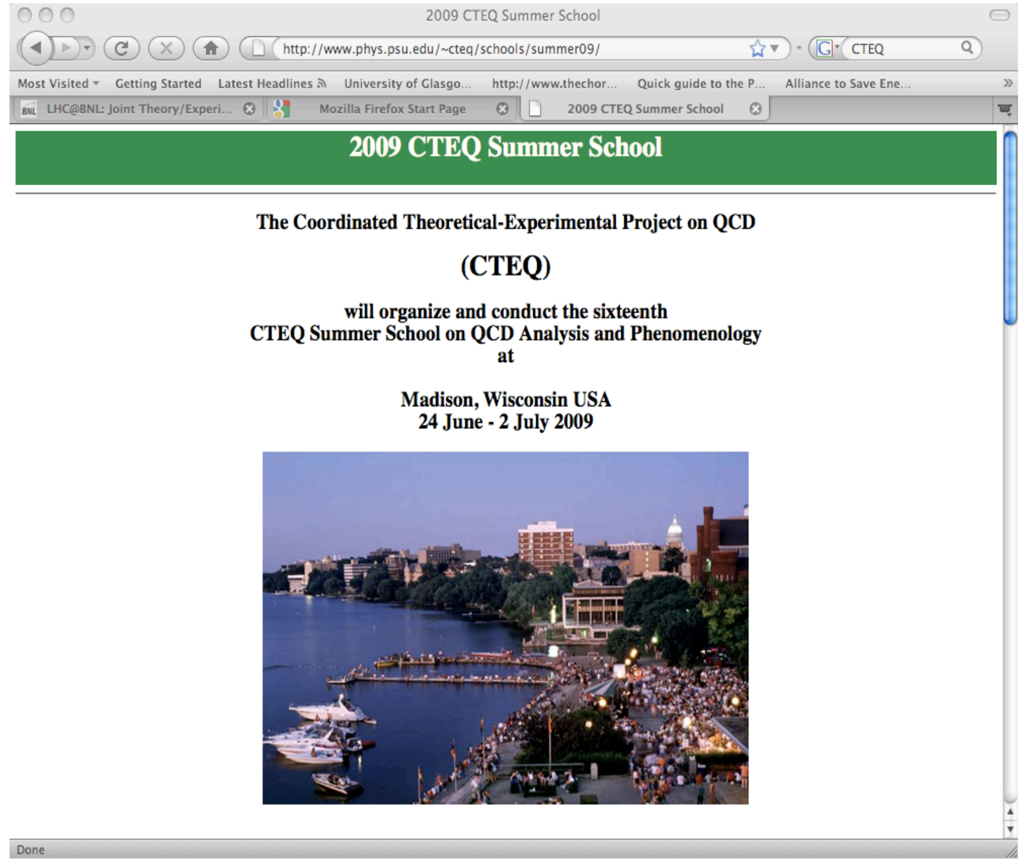
- Physics will come flying hot and heavy when LHC turns on in 2009
- Important to establish both the SM benchmarks and the tools we will need to properly understand this flood of data
- Having (only) 200 pb^{-1} of data at 10 TeV may be the best thing for us…understanding before discovery
- …but perhaps not the most exciting
- Much of the work discussed in this talk will continue at Les Houches
- June 8-26, 2009

http://www.lpthe.jussieu.fr/LesHouches09Wiki/index.php/Main_Page

- Plans for Les Houches
  - collecting results of completed higher order calculations
    - tables, plots and ntuples a la CTEQ4LHC
    - common format for storing parton level information in the ntuples
    - scale variations stored
  - special interest in higher order corrections of Higgs observables
    - any specific ATLAS input?
  - missing processes for wishlist
  - standardization of NLO computations
    - minimal agreement on color and helicity management and on passing IR subtraction terms could lead to transportable modules for virtual corrections
  - new techniques for NLO computations
  - IR safe jet algorithms

Summary-2

- In the near future, CTEQ will also have
  - update to NLO pdf's
    - recent Tevatron data
    - eigenvector tools
  - modified LO pdf's
    - several types
  - combined (x and q_t) pdf fits
    - useful for precision measurements such as W mass
  - NNLO pdf's
    - will then make the relevant Higgs ntuples



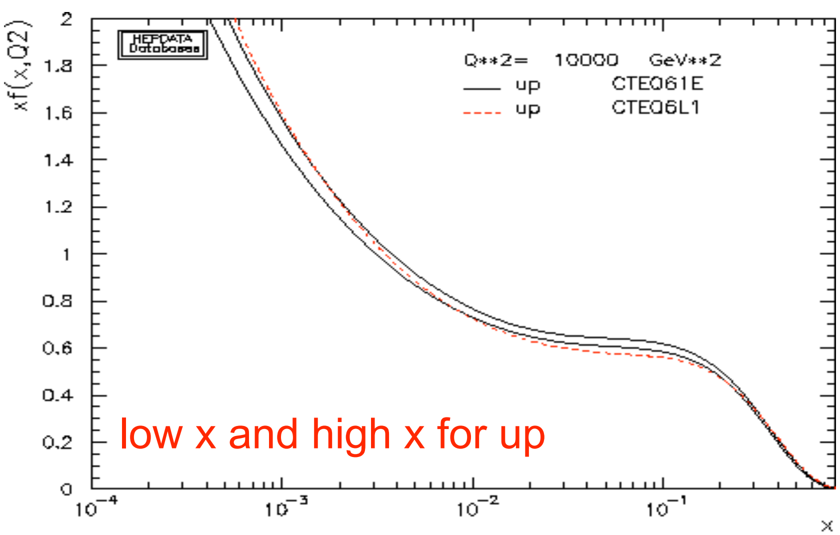
Modified LO pdf's (LO*)

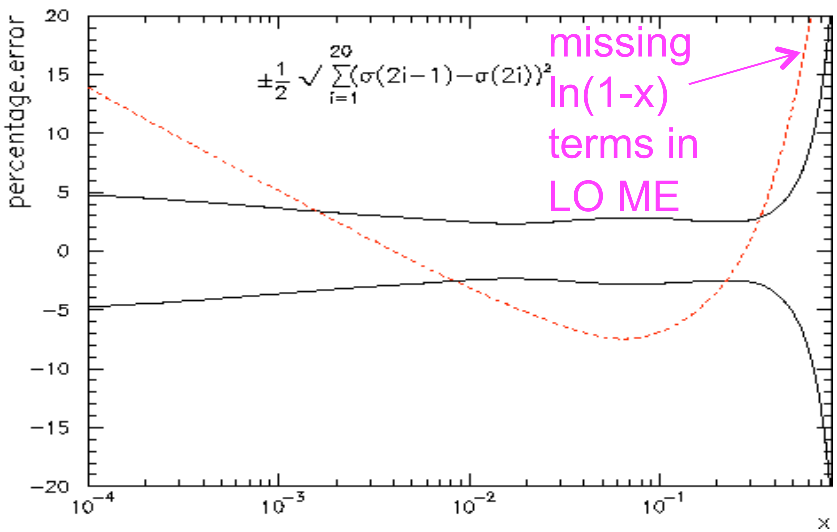
- What about pdf's for parton shower Monte Carlos?
  - standard has been to use LO pdf's, most commonly CTEQ5L/ CTEQ6L, in Pythia, Herwig, Sherpa, ALPGEN/Madgraph+…
- …but
  - LO pdf's can create LHC cross sections/acceptances that differ in both shape and normalization from NLO
    - due to influence of HERA data
    - and lack of ln(1/x) and ln(1-x) terms in leading order pdf's and evolution
  - …and are often outside NLO error bands
  - experimenters use the NLO error pdf's in combination with the central LO pdf even with this mis-match
    - causes an error in pdf re-weighting due to non-matching of Sudakov form factors
  - predictions for inclusive observables from LO matrix elements for many of the collider processes that we want to calculate are not so different from those from NLO matrix elements (aside from a reasonably constant K-factor)

Modified LO pdf's (LO*)

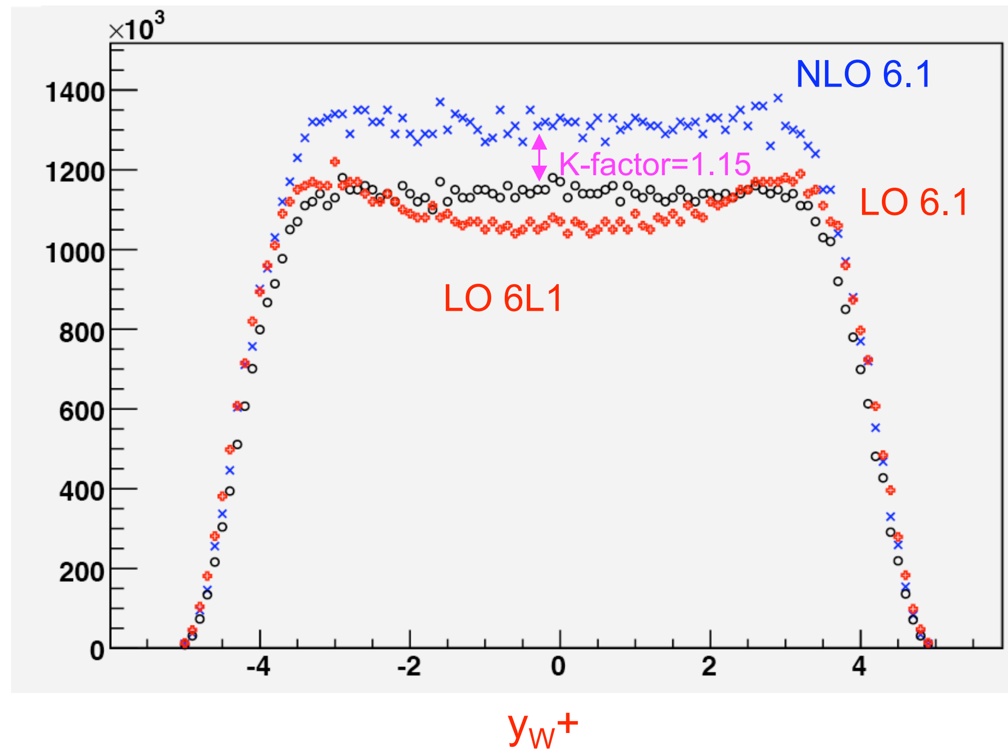
- …but
  - we (and in particular Torbjorn) *like* the low x behavior of LO pdf's and rely upon them for our models of the underlying event at the Tevatron and its extrapolation to the LHC
  - as well as calculating low x cross sections at the LHC
  - and no one listened to me when I urged the use of NLO pdf's
- thus, the need for modified LO pdf's

Where are the differences between LO and NLO partons?





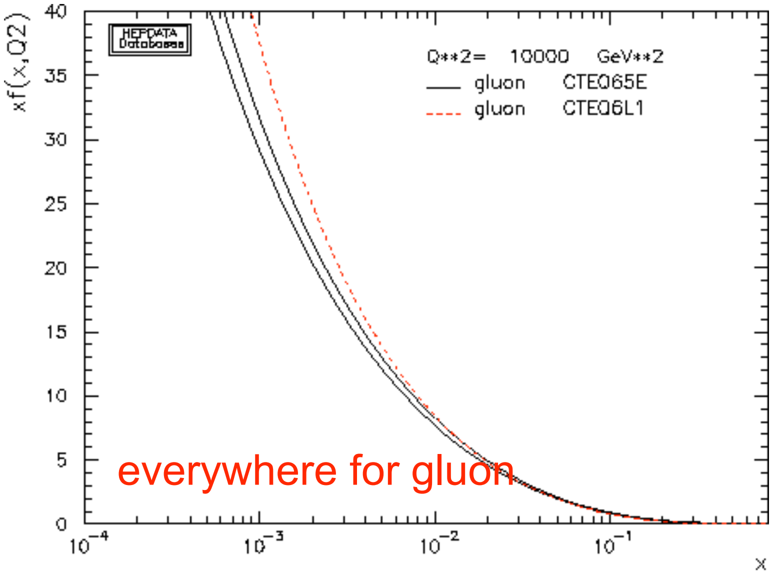
W+ rapidity distribution at LHC

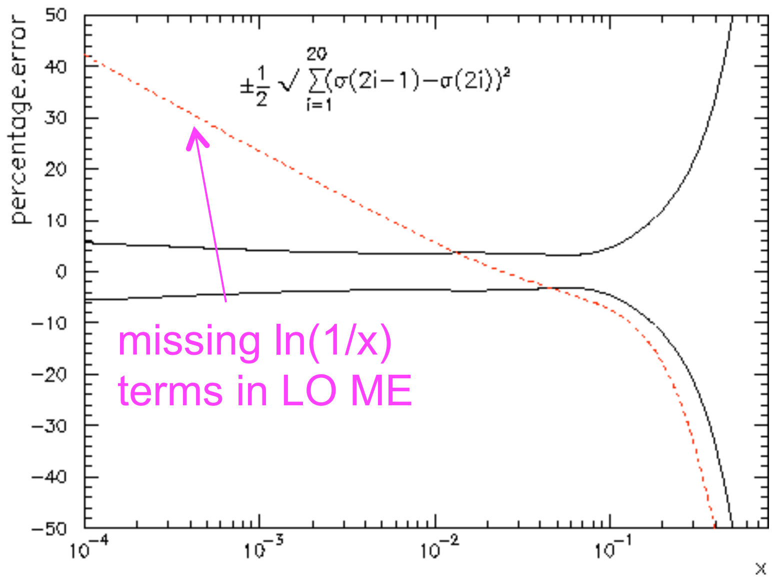


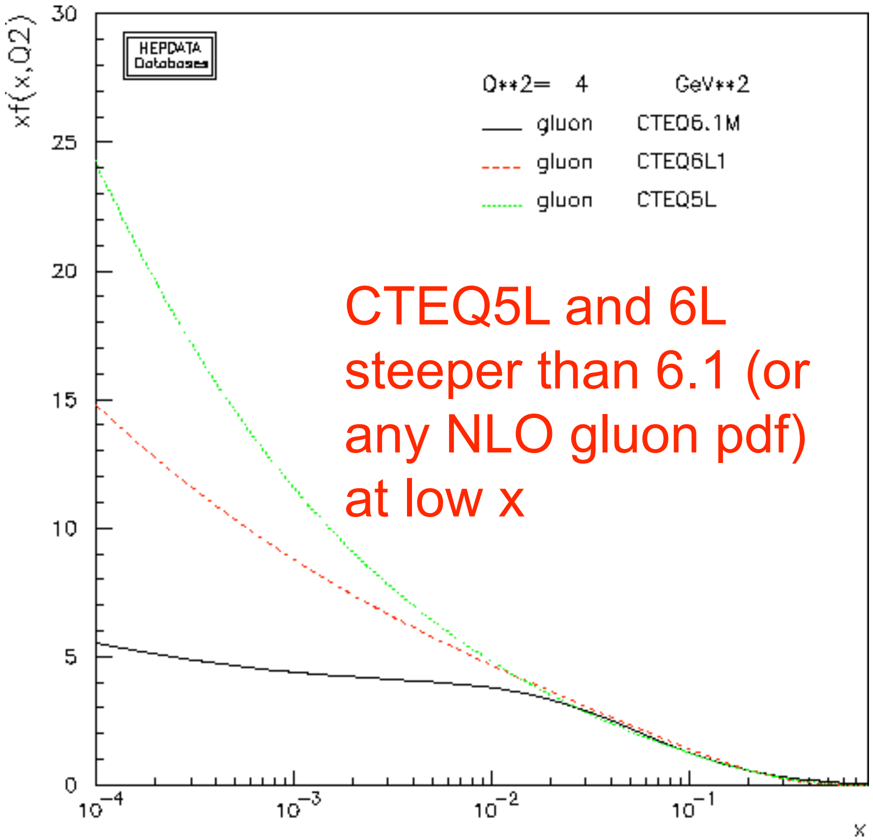
For example, the shape of the W+ rapidity distribution is significantly different than the NLO result if the LO pdf is used, but very similar if the NLO pdf is used.

Where are the differences?

- at low Q





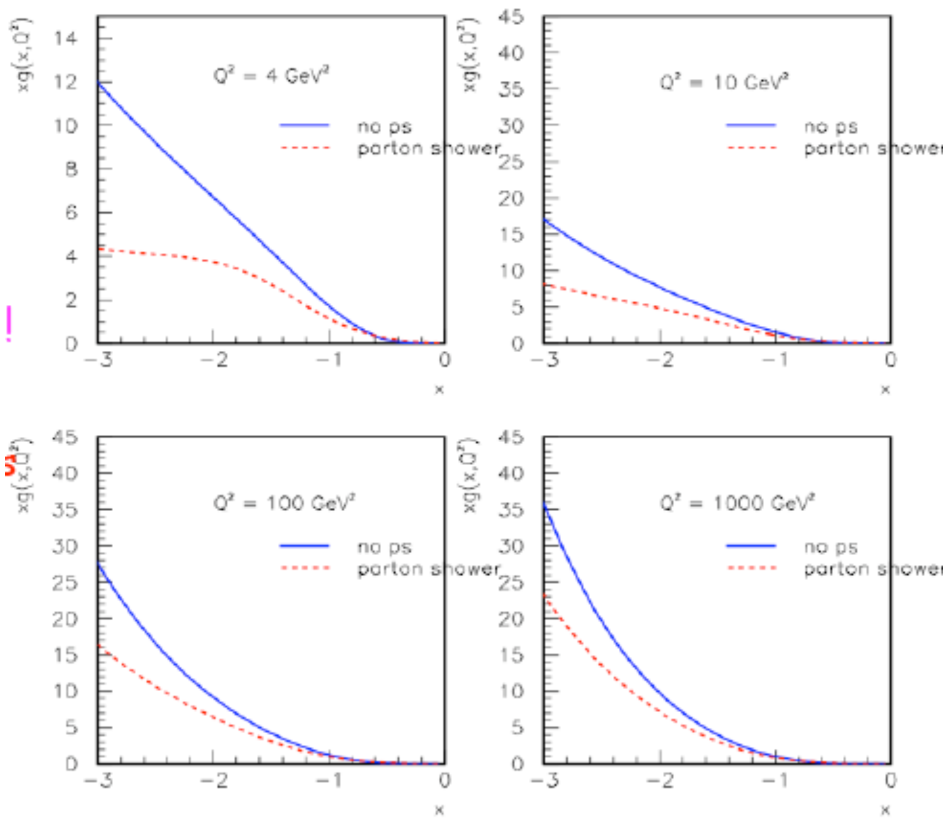


3 techniques from CTEQ

- “1”: since NLO pdf’s give a better description of hard scattering cross sections and LO pdf’s give a better description of the UE, use separate pdf’s for those two functions
  - implicit in Pythia8
  - separate pdf’s for (1) ME evaluation and (2) UE/ISR
- 2 and 3: generate NLO pseudo-data for use in global fit with real data (CTEQ6.6 dataset) to steer desired behavior of modified LO pdf’s
- Would like the modified LO pdf’s to look like LO pdf’s at low x and NLO pdf’s at high x
- This comes about automatically with method 1, and is enforced by the pseudo-data in methods 2 and 3

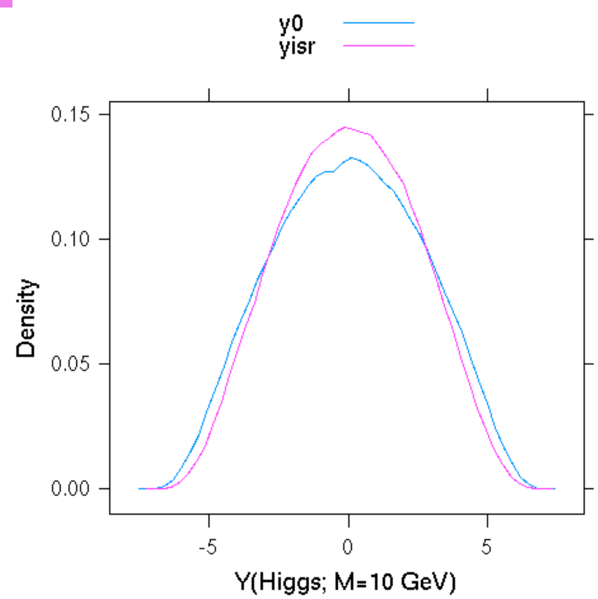
A technical issue

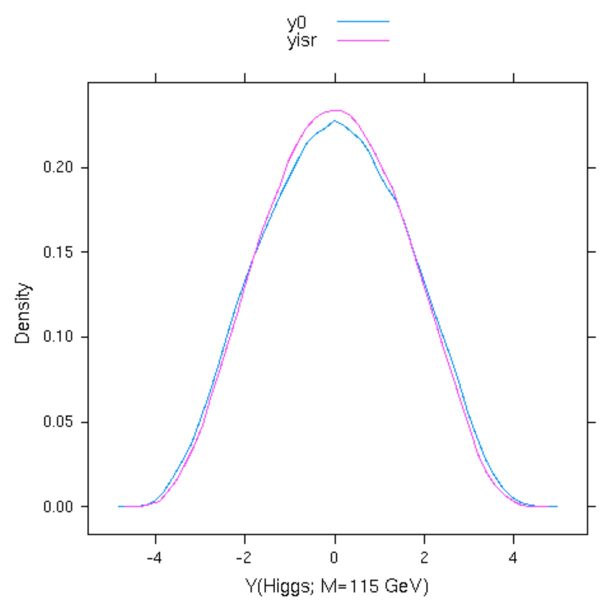
- Pdf's determined by global analysis use DGLAP formalism, where all splittings are collinear
- In parton showering, branches are not collinear, but have a transverse component as well
- Is there a significant suppression for Monte Carlo predictions at the LHC due to the fact that the pdf's are determined using the collinear assumption?
- Do we need pdf's that are determined in a Monte Carlo scheme?
  - perhaps a different scheme for each showering algorithm
  - this point was first brought up by Hannes Jung, I believe, where he saw the effect for HERA predictions



What about at the LHC?

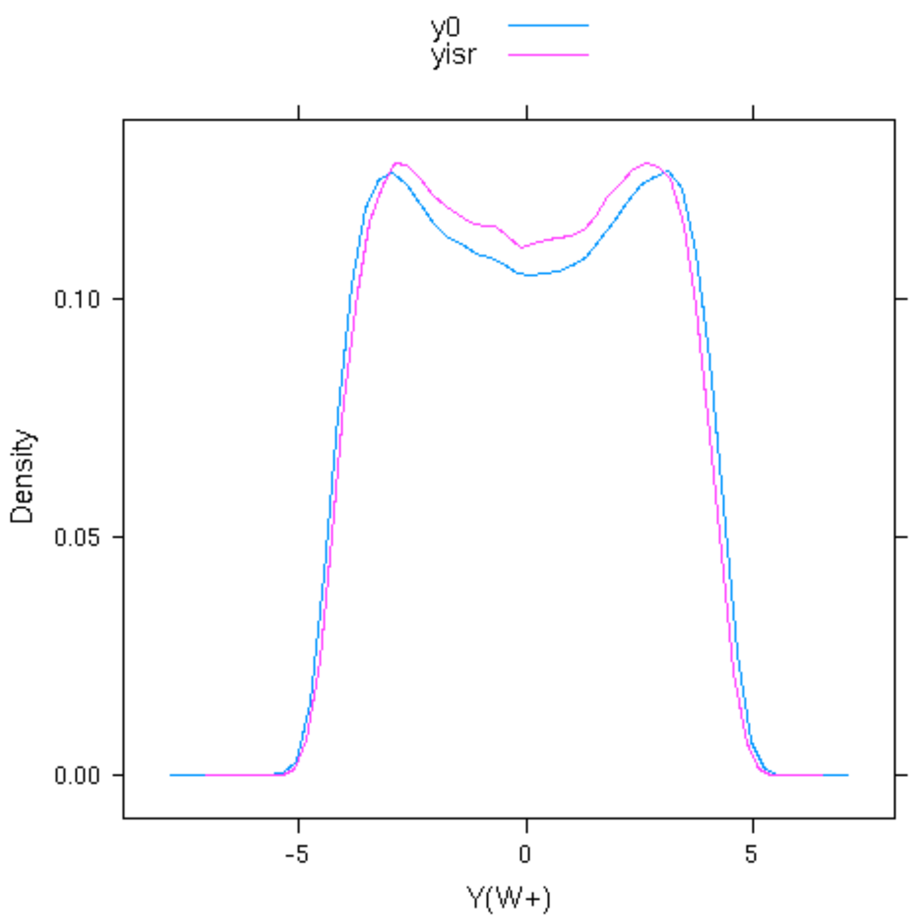
- Consider a 10 GeV "Higgs" produced through gg fusion
- There, a significant impact can be observed by the introduction of parton showering
- But the effect becomes much smaller for the production of a 115 GeV Higgs

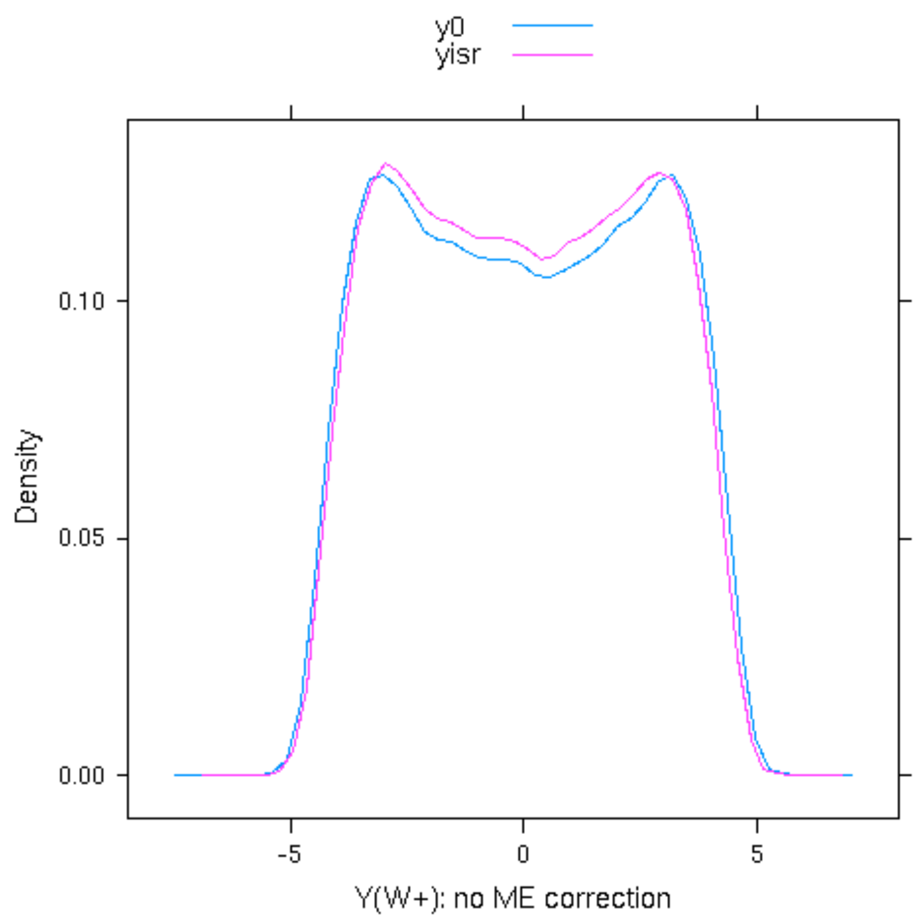




W^+ rapidity

- For W^+ rapidity distribution (either with or without ME correction), parton showering tightens up rapidity distribution, but effect is moderate
- It still may be a good idea to make the mod LO pdf’s a bit more *robust* to counteract this effect





Also

- For rapidly falling inclusive distributions (think jet production at the LHC), there is a difference between the results of a fixed (LO) calculation and a parton shower Monte Carlo
- There is a K-factor implicit in Monte Carlo generation because of the smearing effects of initial state radiation
- Under investigation

CTEQ techniques

- Include in LO* fit (weighted) pseudo-data for characteristic LHC processes produced using CTEQ6.6 NLO pdf's with NLO matrix elements (using MCFM), along with full CTEQ6.6 dataset (2885 points)
  - low mass bB
    - fix low x gluon for UE
  - tT over full mass range
    - higher x gluon
  - W^+,W^-,Z^0 rapidity distributions
    - quark distributions
  - gg->H (120 GeV) rapidity distribution

Choices

- Use of 2-loop or 1-loop a_s
  - Herwig preference for 2-loop
  - Pythia preference for 1-loop
- Fixed momentum sum rule, or not
  - re-arrange momentum within proton and/or add extra momentum
  - extra momentum appreciated by some of pseudo-data sets but not others and may lose some useful correlations
- Fix pseudo-data normalizations to K-factors expected from higher order corrections, or let float
- Scale variation within reasonable range for fine-tuning of agreement with pseudo-data
  - for example, let vector boson scale vary from 0.5 m_B to 2.0 m_B
- Will provide pdf's with several of these options for user

Some observations

- Pseudo-data has conflicts with global data set
  - that's the motivation of the modified pdf's
- Requiring better fit to pseudo-data increases chisquare of LO fit to global data set (although this is not the primary concern; the fit to the pseudo-data is)
  - chisquare better with 2-loop a_s with no pseudo-data in fit
    - no strong preference for 1-loop or 2-loop a_s with inclusion of pseudo-data
  - chisquare improves with momentum sum rule free
    - prefers more momentum (~1.05-1.10); mostly goes into the gluon distribution
    - normalization of pseudo-data (needed K-factor) gets closer to 1 (since the chisquare gets better if that happens)
    - still some conflicts with DIS data that don't prefer more momentum
    - …but we're not making these pdf's for DIS comparisons

Some references

INSTITUTE OF PHYSICS PUBLISHING | REPORTS ON PROGRESS IN PHYSICS
Rep. Prog. Phys. **70** (2007) 89–193 | doi:10.1088/0034-4885/70/1/R02

Hard interactions of quarks and gluons: a primer for LHC physics

J M Campbell[1], J W Huston[2] and W J Stirling[3] CHS

[1] Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
[2] Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48840, USA
[3] Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, UK

E-mail: j.campbell@physics.gla.ac.uk, huston@msu.edu and w.j.stirling@durham.ac.uk

Received 14 July 2006, in final form 6 November 2006
Published 19 December 2006
Online at stacks.iop.org/RoPP/70/89

Abstract

In this paper, we will develop the perturbative framework for the calculation of hard-scattering processes. We will undertake to provide both a reasonably rigorous development of the formalism of hard-scattering of quarks and gluons as well as an intuitive understanding of the physics behind the scattering. We will emphasize the role of logarithmic corrections as well as power counting in α_S in order to understand the behaviour of hard-scattering processes. We will include 'rules of thumb' as well as 'official recommendations', and where possible will seek to dispel some myths. We will also discuss the impact of soft processes on the measurements of hard-scattering processes. Experiences that have been gained at the Fermilab Tevatron will be recounted and, where appropriate, extrapolated to the LHC.

(Some figures in this article are in colour only in the electronic version)

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Progress in Particle and Nuclear Physics 60 (2008) 484–551
Progress in Particle and Nuclear Physics
www.elsevier.com/locate/ppnp

Review

Jets in hadron–hadron collisions

S.D. Ellis[a,*], J. Huston[b], K. Hatakeyama[c], P. Loch[d], M. Tönnesmann[e]

[a] *University of Washington, Seattle, WA 98195, United States*
[b] *Michigan State University, East Lansing, MI 48824, United States*
[c] *Rockefeller University, New York, NY 10021, United States*
[d] *University of Arizona, Tucson, AZ 85721, United States*
[e] *Max Planck Institute fur Physics, Munich, Germany*

arXiv:07122447 Dec 14, 2007

Abstract

In this article, we review some of the complexities of jet algorithms and of the resultant comparisons of data to theory. We review the extensive experience with jet measurements at the Tevatron, the extrapolation of this acquired wisdom to the LHC and the differences between the Tevatron and LHC environments. We also describe a framework (SpartyJet) for the convenient comparison of results using different jet algorithms.

Keywords: Jet; Jet algorithm; LHC; Tevatron; Perturbative QCD; SpartyJet

Contents