PDF Error Discussion

J. Huston
Michigan State University

lepton + jets 2/27/04

Global pdf fits

- Calculation of production cross sections at the Tevatron relies upon knowledge of pdfs in relevant kinematic range
- pdfs 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
  - CTEQ->CTEQ5->CTEQ5(1)->CTEQ6->CTEQ6.1
  - also GKK and Alekhin, but not widely used
- CTEQ6(1) is based on series of previous CTEQ distributions, but represents more than an evolutionary advance
  - update to new data sets
  - incorporation of correlated systematic errors for all experiments in the fit
  - new methodology enables full characterization of parton parametrization space in neighborhood of global minimum
    - Hessian method
    - Largrange Multiplier
  - results available both in conventional formalism and in Les Houches accord format (more on this later)

Nuts/bolts of fits

- Functional form used in CTEQ fits is:
  - $xf(x,Q_o) = A_o\, x^{A1}\, (1-x)^{A2}\, e^{A3x}\, (1 + A_4x)^{A5}$
    - Q_o = 1.3 GeV (below any data used in fit)
      - easier to do forward evolution than backward
      - MRST starts at 1 GeV (- gluon distribution)
    - functional form arrived at by adding a 1:1 Pade expansion to quantity d(log xf)/dx
    - more versatile than form used in CTEQ5 or MRST
    - there are 20 free parameters used in the global fit
      - MRST has 15 free parameters
- Light quarks treated as massless; evolution kernels of PDFs are mass-independent
- Zero mass Wilson coefficients used in DIS structure functions
- NB: MRST pdf's not in pure MS-bar scheme; use Roberts-Thorne treatment of heavy quarks at threshold
  - maybe noticeable only at low x

PDF Uncertainties

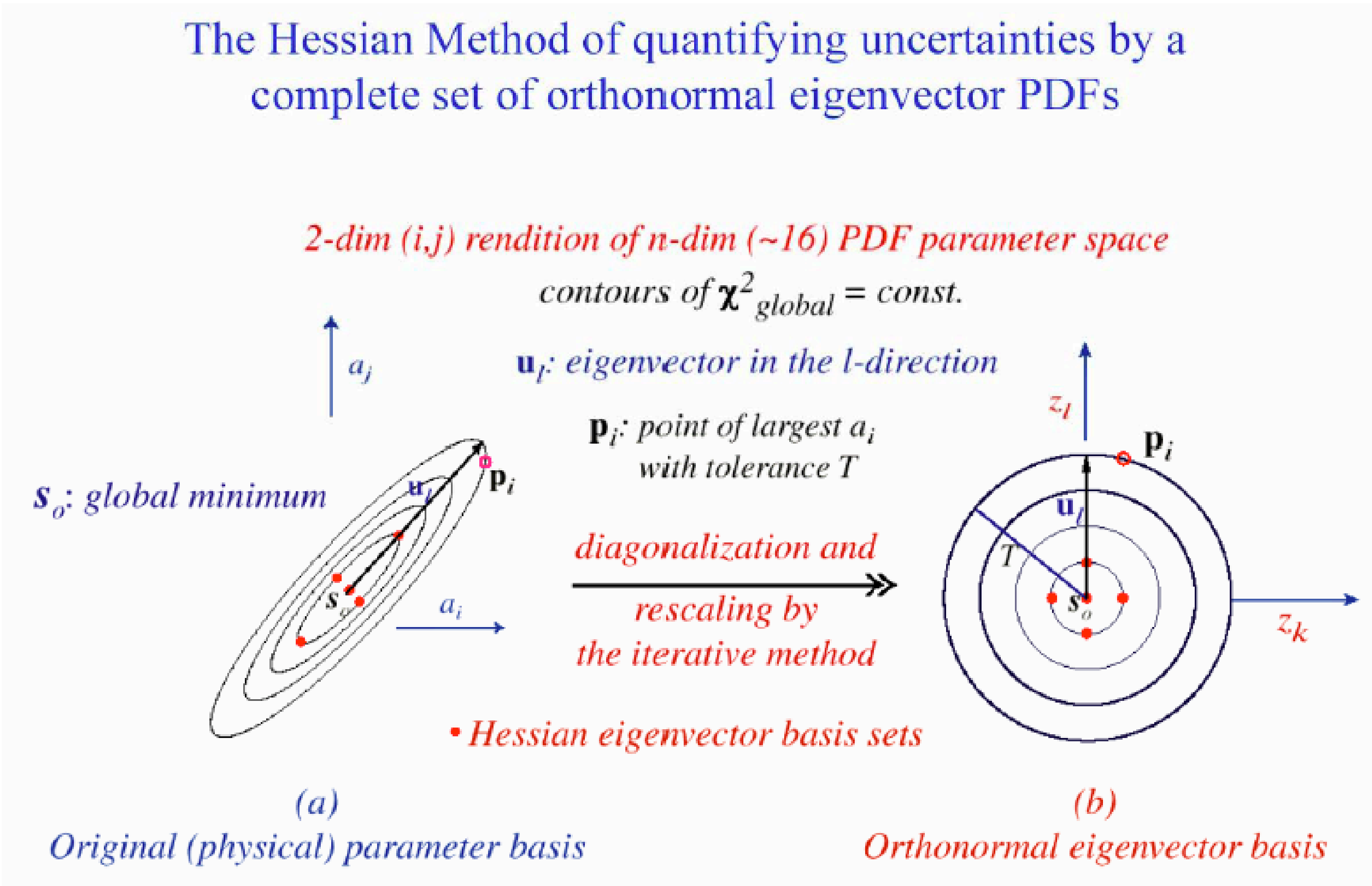
- What's unknown about PDF's
  - the gluon distribution
  - strange and anti-strange quarks
  - details in the {u,d} quark sector; up/down differences and ratios
  - heavy quark distributions

- Σ of quark distributions (q + qbar) is well-determined over wide range of x and Q^2
  - quark distributions primarily determined from DIS and DY data sets which have large statistics and systematic errors in few percent range ($\pm 3\%$ for $10^{-4}<x<0.75$)
  - individual quark flavors, though may have uncertainties larger than that on the sum; important, for example, for W asymmetry

Uncertainties in pdf fits

- Two sources
  - Experimental errors
    - Hessian/Lagrange multiplier techniques designed to address estimate of these effects
      - question is what $\Delta\chi^2$ change best represents estimate of uncertainty (CTEQ uses $\Delta\chi^2$ of 100 (out of 2000) for 90% CL limit; MRST uses $\Delta\chi^2$ of 50); GKK/Alekhin uses 1 (for 1 sigma error)
  - Theoretical
    - higher twist/non-perturbative effects
      - choose Q^2 and W cuts to try to avoid
    - higher order effects
      - is NNLO necessary yet?
    - edge of phase space effects
      - threshold resummation needed?

Hessian method



Improvements in Hessian approach

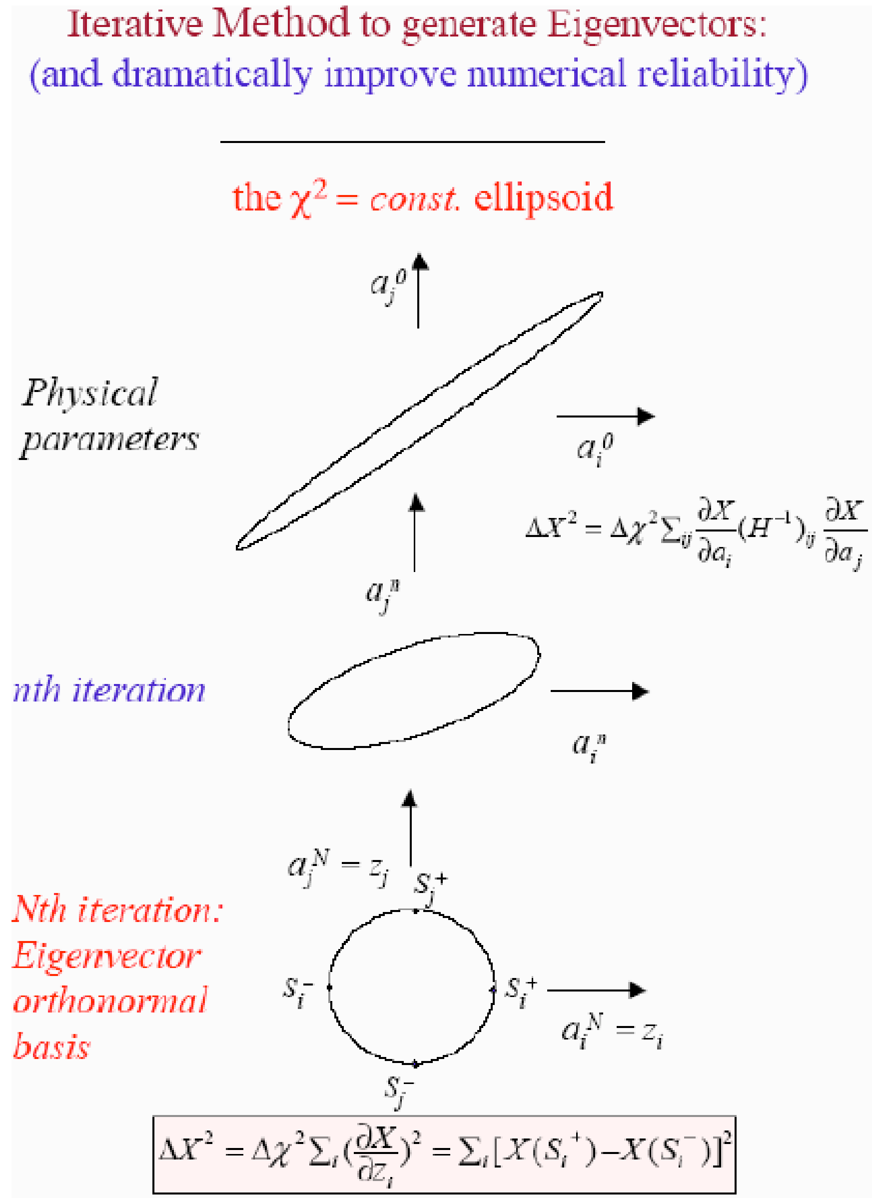
A critical technical advance in the Hessian approach which enabled the CTEQ uncertainty studies

The Hessian method for χ^2 analysis has always been the standard, but uncertainty estimates in global QCD analysis by standard tools had been known to be extremely unreliable due to two practical problems:

- .extreme range of eigenvalues (flat vs. steep)
- numerical fluctuations of theory predictions

An iterative method by Jon Pumplin solved both of these technical difficulties, provided the means to generate reliable eigenvectors in parton parameter space, hence allow the systematic exploration of this space, particularly the a priori unknown "flat directions"

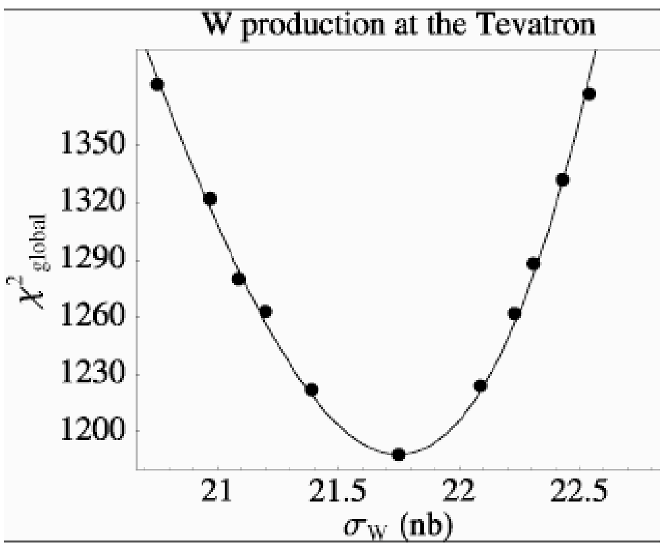
Some flat directions can be important; see inclusive jet cross section.



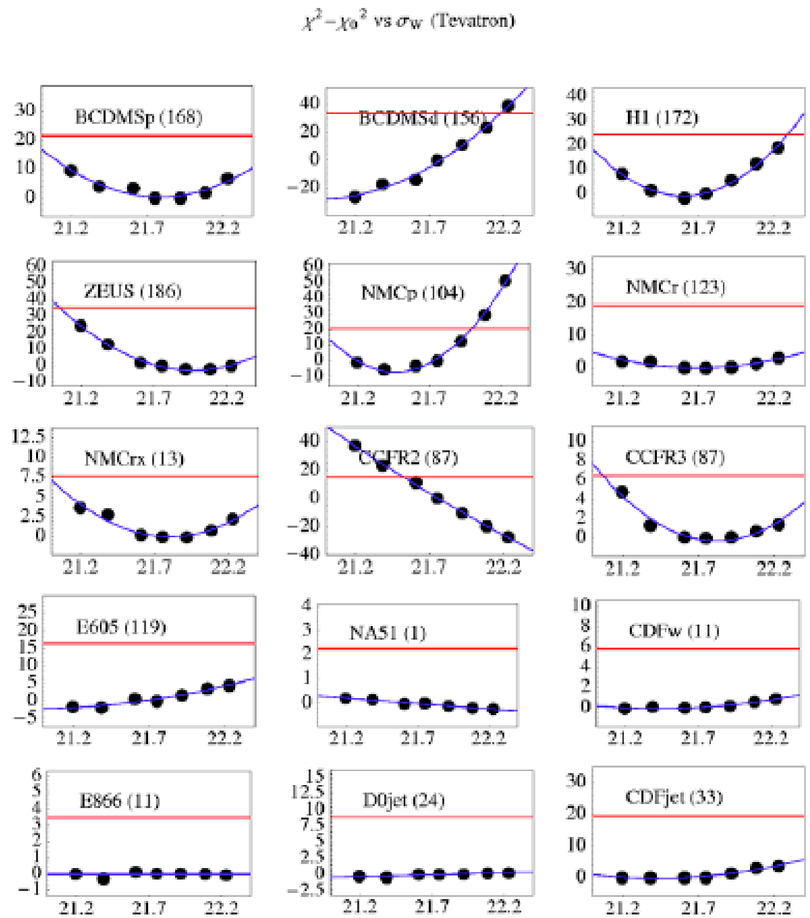
W cross section

Case study: CTEQ global analysis of σ_W (χ^2 method)

Estimate the uncertainty on the predicted cross section for $pp_{bar} \rightarrow W+X$ at the Tevatron collider.



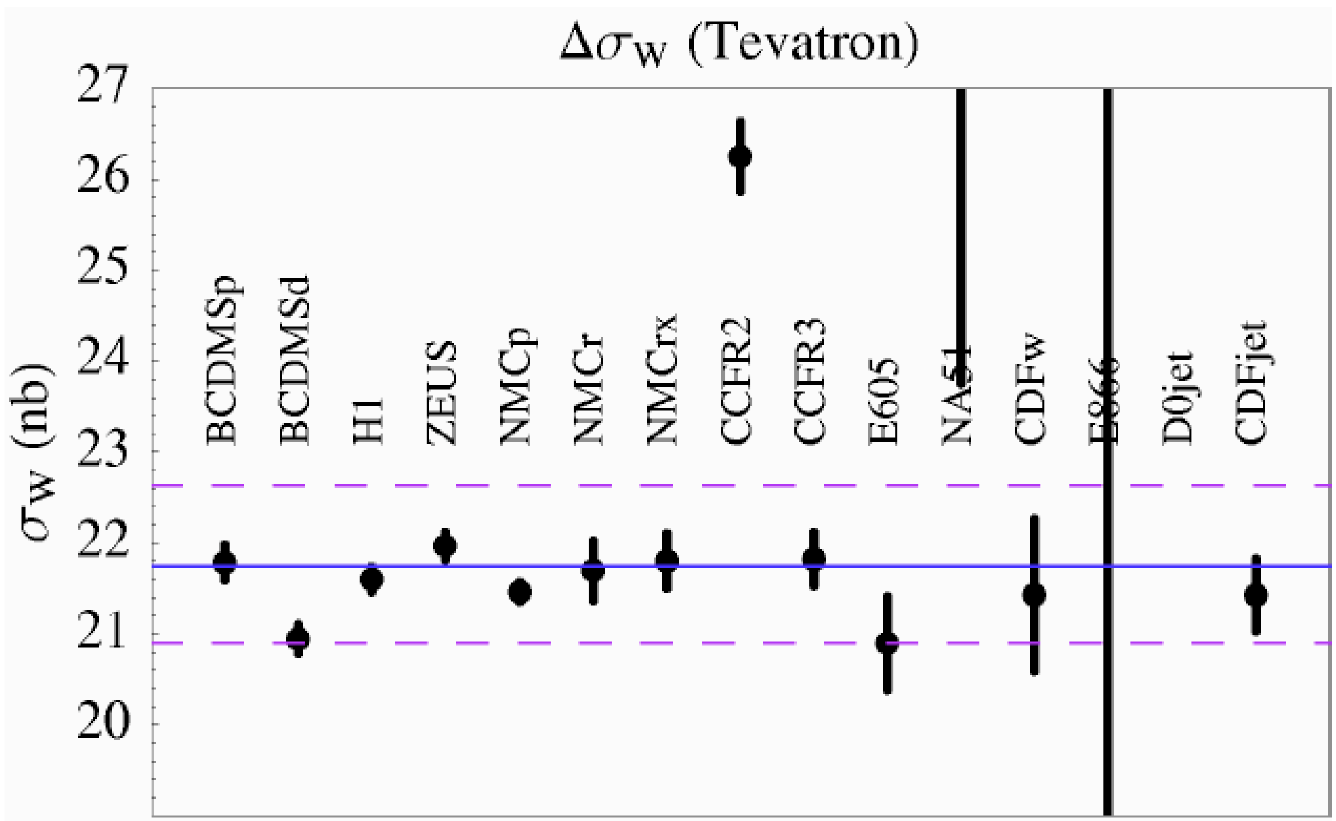
global χ^2



local χ^2's

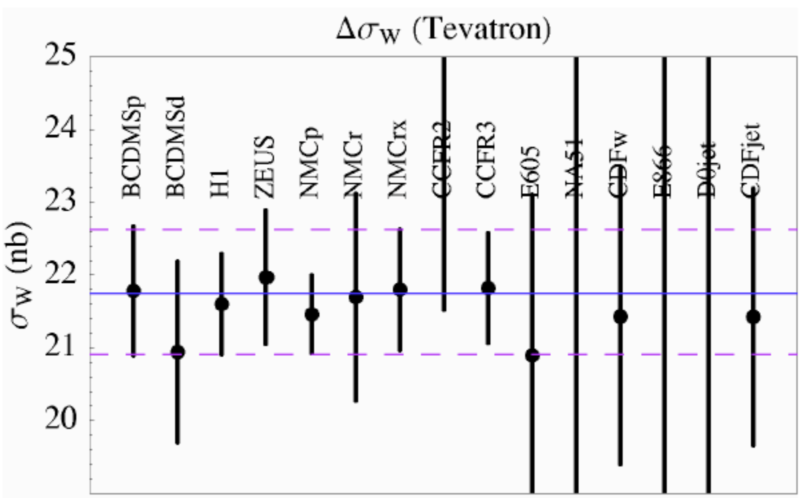
Determining tolerance

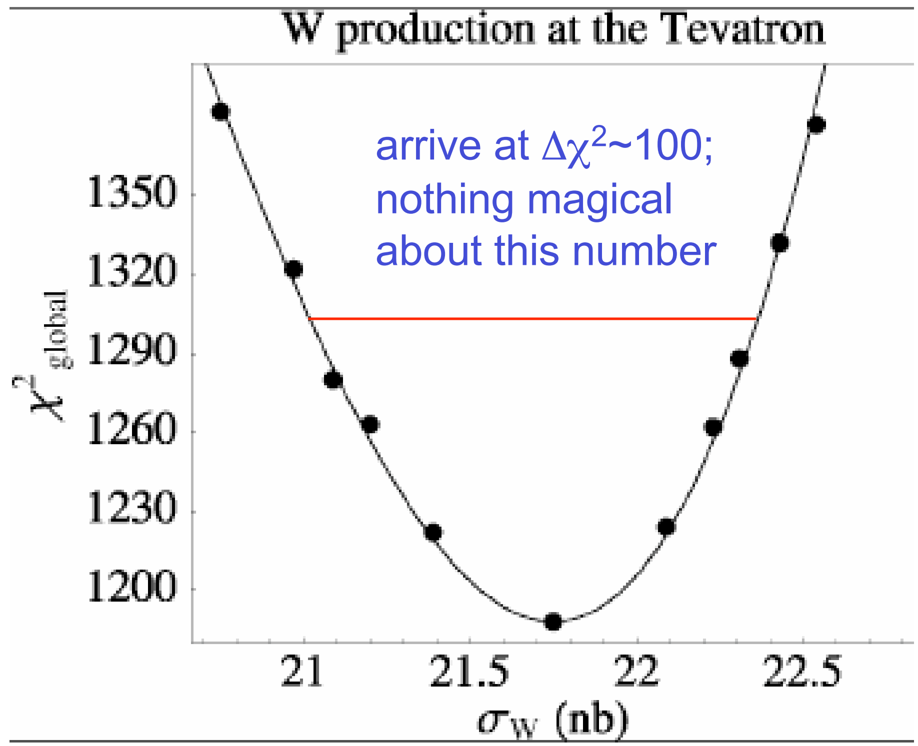
Each experiment defines a "prediction" and a "range".
This figure shows the $\Delta\chi^2 = 1$ ranges.



Determining tolerance

This figure shows broader ranges for each experiment based on the "90% confidence level" (cumulative distribution function of the rescaled χ^2).





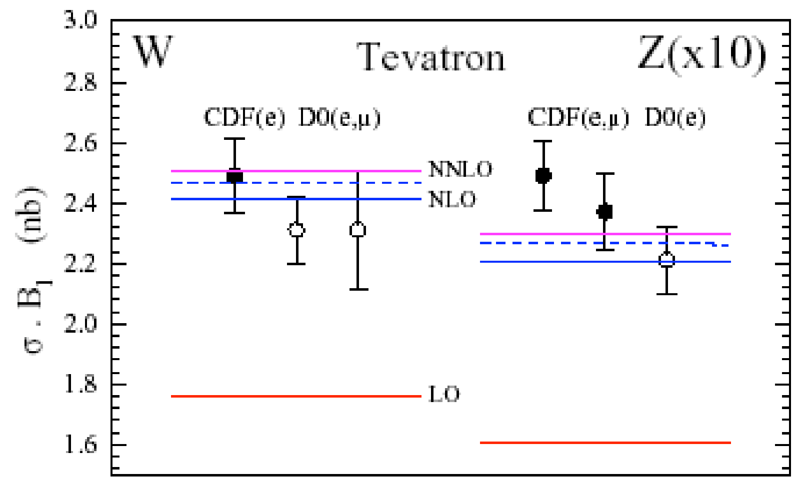
global χ^2

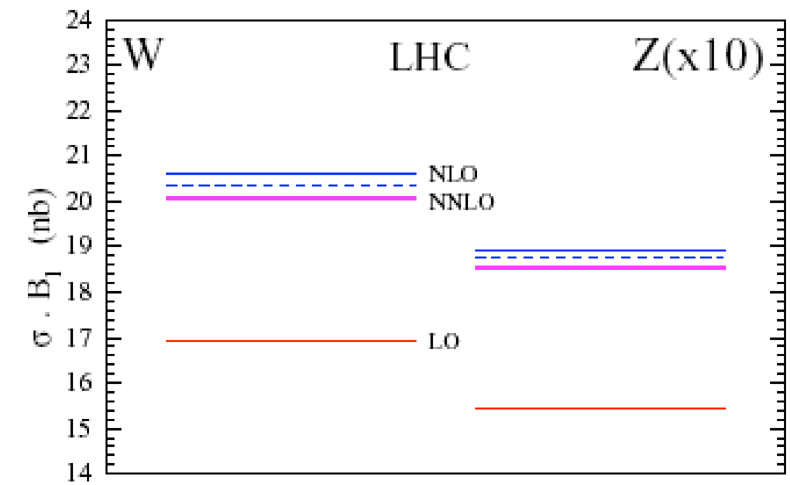
More on W theoretical uncertainties

Reasonable stability order by order for (quark-dominated) W and Z cross-sections.

However, changes of order 4%. Much bigger than uncertainty due to experimental errors.

This fairly good convergence is largely guaranteed because the quarks are fit directly to data.





CTEQ obtain for $\alpha_S = 0.118$

$$\Delta\sigma_W(\mathrm{LHC}) \approx \pm 4\% \quad \Delta\sigma_{\mathrm{W}}(\mathrm{Tev}) \approx \pm 4$$

$$\Delta\sigma_H(\mathrm{LHC}) \approx \pm 5\%.$$

MRST use a wider range of data, and if $\Delta\chi^2 \sim 50$ find for $\alpha_S = 0.119$

$$\Delta\sigma_W(\mathrm{Tev}) \approx \pm 1.2\% \quad \Delta\sigma_{\mathrm{W}}(\mathrm{LHC}) \approx \pm 2\%$$

$$\Delta\sigma_H(\mathrm{Tev}) \approx \pm 4\% \quad \Delta\sigma_{\mathrm{H}}(\mathrm{LHC}) \approx \pm 2\%.$$

CTEQ W cross section at Tevatron is ~2% lower than MRST;
Alekhin is ~4% higher → What does this say about uncertainty range?

PDF uncertainties for Run 1 jet cross section

- 20 free parameters in the fit
- In the Hessian method, a 20X20 matrix is diagonalized and 20 orthogonal eigenvector directions in parameter space are determined

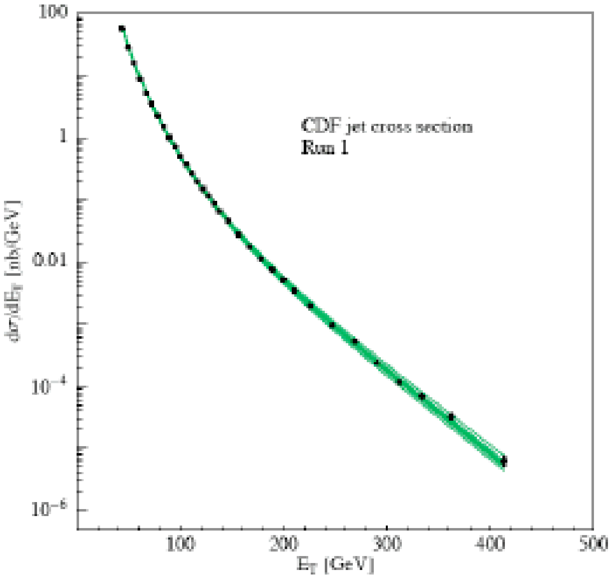


Figure 9: Calculations of $d\sigma/dE_T$ for the 40 eigenvector sets.

•Each eigenvector direction corresponds to some linear combination of pdf parameters
•Large eigenvalues correspond to highly determined directions (e.g. valence quarks)
•Small eigenvalues correspond to poorly determined directions (high x gluon)
•Result is 40 pdf's (go along + and - direction $\Delta\chi^2$ of 100 for each eigenvalue)

Note 1 eigenvector(15+) leads to noticeably larger prediction than the others

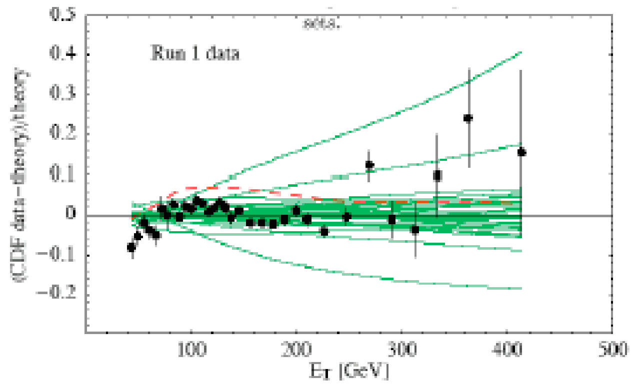


Figure 10: Calculations for the 40 eigenvector basis sets, plotted as fractional differences compared to CTEQ6.1M. The points are the CDF measurements. The error bars are statistical errors only; the systematic shifts are not subtracted from the data. The dashed curve is CTEQ5HJ compared to CTEQ6.1M.

See hep-ph/0303013 **Inclusive Jet Production, Parton Distributions, and the Search for New Physics**

CDF jet cross section uncertainties

- On the right are shown the uncertainties for the CDF jet cross section along each eigenvector ($\Delta\chi^2$ = 100)
  - jet cross section most sensitive to eigenvector 15
    - which mainly contains parameters relating to behavior of high x gluon

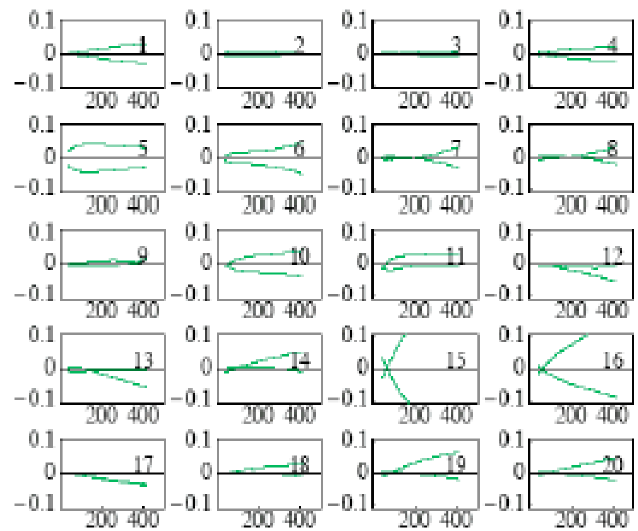
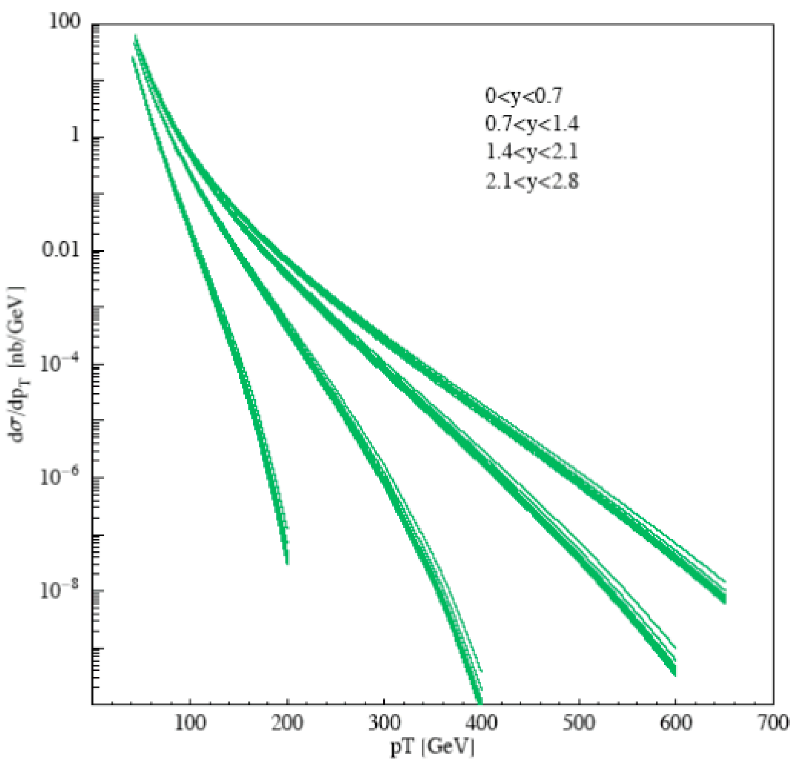


Figure 11: The cross sections for + and − displacements along the 20 eigenvectors, plotted as fractional differences compared to CTEQ6.1M.

→ one of the flat directions

Uncertainties for Run 2 jet cross sections

- CDF will measure the inclusive jet cross section in the forward regions as well



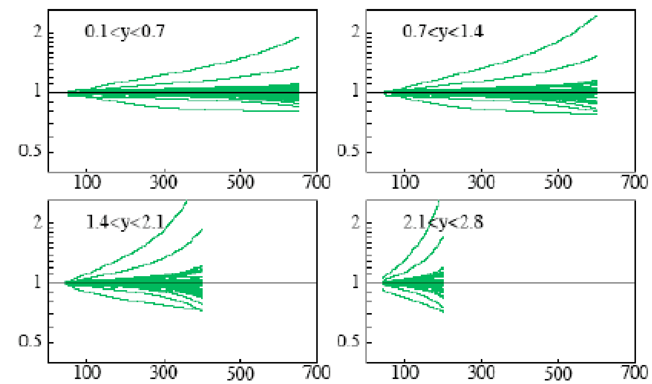
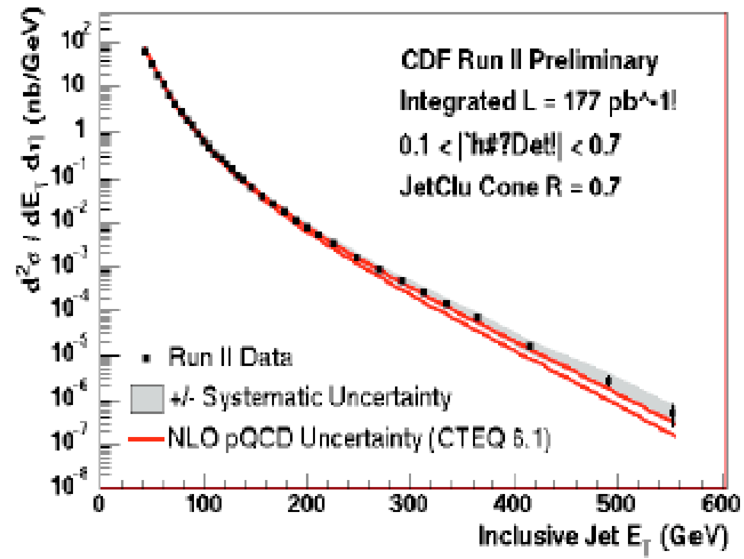
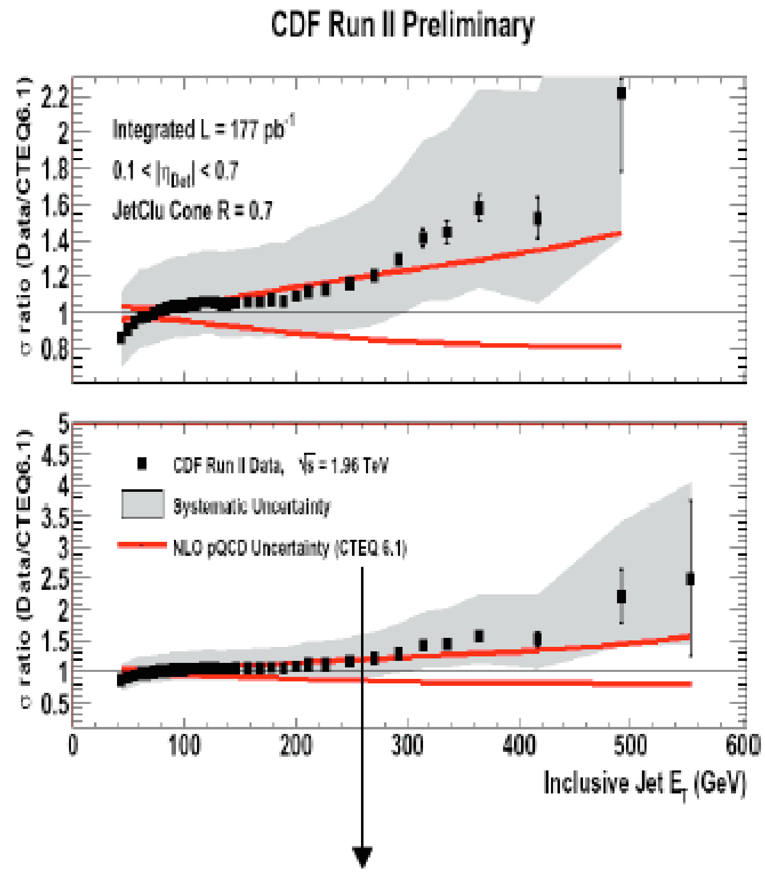


Figure 28: Uncertainty range of the Run 2 cross section for the CDF rapidity bins. The curves show the ratios of the 40 eigenvector basis sets compared to the central (CTEQ6.1M) prediction

NB: new physics is primarily central;
a pdf explanation should work everywhere

CDF Run 2 jet cross sections





Jet cross section in agreement with theoretical prediction within errors

Jet energy scale error will decrease as understanding of Run 2 detector improves

Dominant theoretical error is due to pdf uncertainty

CTEQ6.1 already has an enhanced high x gluon due to influence of Run 1 jet data

Effective use of pdf uncertainties

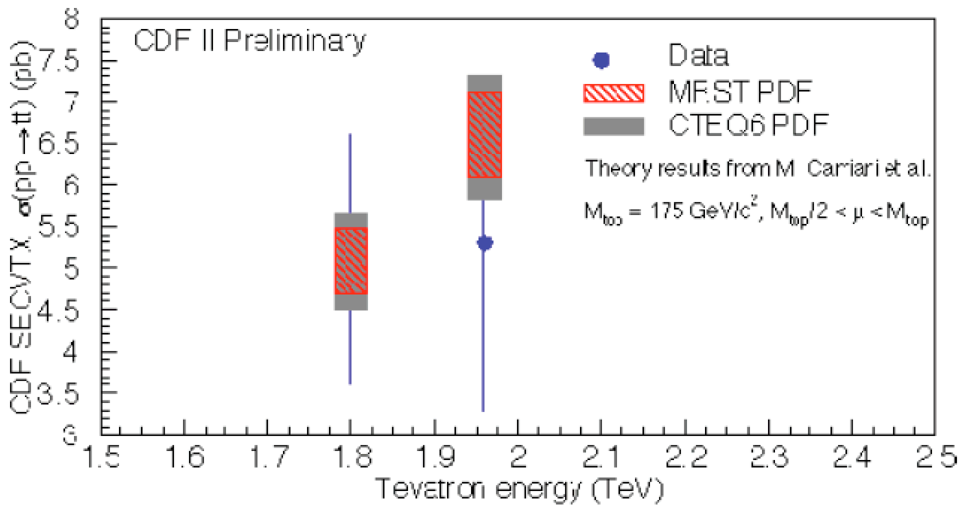
- PDF uncertainties are important both for precision measurements (W/Z cross sections) as well as for studies of potential new physics (a la jet cross sections at high E_T)
- Most Monte Carlo/matrix element programs have "central" pdf's built in, or can easily interface to PDFLIB
- Determining the pdf uncertainty for a particular cross section/distribution might require the use of many pdf's
  - CTEQ Hessian pdf errors require using 40 pdf's
  - GKK on the order of 100
  - MRST2002 requires 30 pdf's
- Too clumsy to attempt to includes grids for calculation of all of these pdf's with the MC programs
- **->Les Houches accord #2 (LHAPDF)**
  - each pdf can be specified by a few lines of information, if MC programs can perform the evolution
  - fast evolution routine used to construct pdf grids as needed

Les Houches accord #2 (LHAPDF)

- Using the interface is as easy as using PDFLIB (and much easier to update)
- Current version has all modern pdf’s, including error sets
- LHAPDF has been handed off to Durham who will provide support in perpetuity
  - http://durpdg.dur.ac.uk/lhapdf/downloads.htm
- call InitiPDFset(*name*)
  - called once at the beginning of the code; *name* is the file name of external PDF file that defines PDF set
- call InitPDF(*mem*)
  - *mem* specifies individual member of pdf set
- call evolvePDF(*x*,*Q*,*f*)
  - returns pdf momentum densities for flavor *f* at momentum fraction *x* and scale *Q*
- In new version, can store all pdf’s in memory at same time
  - run program with central pdf, and then store pdf*pdf weight for all error pdf’s
  - in MCFM now

Using pdf uncertainties

- CTEQ or MRST pdf's are meant to be used with NLO programs
- It's relatively straightforward, for example, to calculate the pdf uncertainty for something like the t-tbar total cross section using a NLO calculation



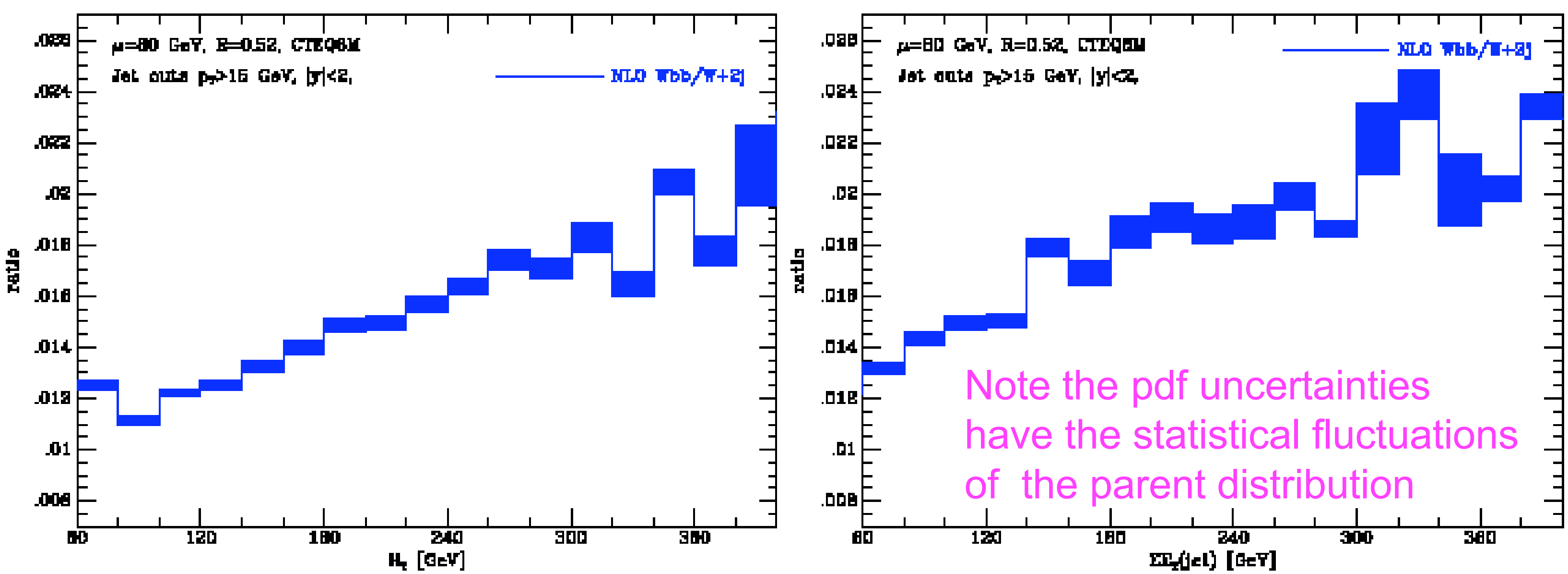
Using pdf uncertainties

NLO programs can be slow, especially if you have to run 41 pdf's
But if new version of LHAPDF is used, can run full cross section with central pdf and store pdf*pdf luminosity for each event and then re-weight

- Total cross-section uncertainty: Using MCFM, see CDF6849

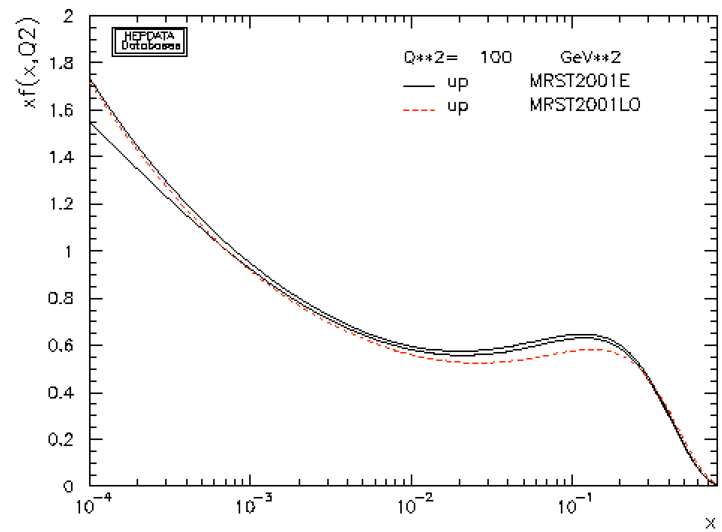
$$W b\bar{b} \to 2.5\% \;, \qquad W + 2j \to 1.5\%.$$

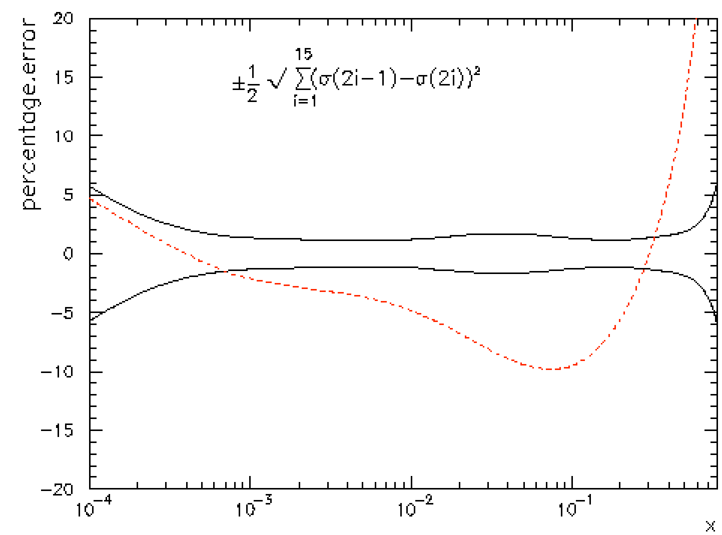
- Uncertainty in the $(Wb\bar{b}/W + 2$ jet) ratio:



Using pdf uncertainties

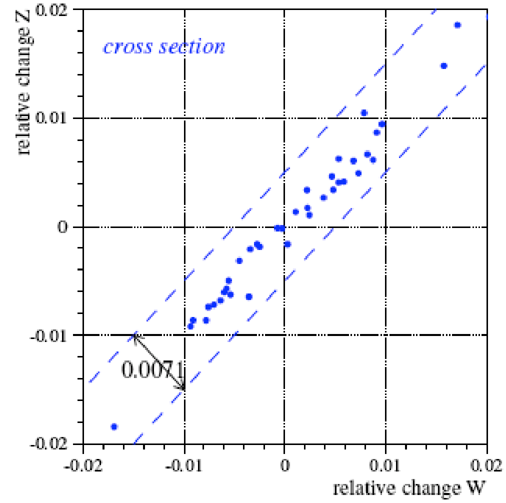
- More often, though, we need to calculate the pdf uncertainty on something like an acceptance
- And we need to use a Monte Carlo so that we can work at the detector level
- Technically, should use LO pdf's with MC's
  - no LO error pdf's
  - but resulting error of using NLO pdf's is NLO, beyond scope of MC
- But, LO central fit can differ from NLO central fit by more than NLO pdf uncertainty
- Also, have to worry about generating enough Monte Carlo events to reduce statistical errors



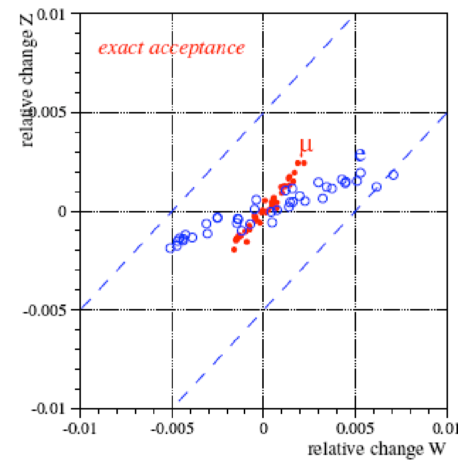


Faster ways

- It's probably not unreasonable to calculate the pdf uncertainty on quantities like acceptances by using the pdf error sets with LO calculations
- This was done for example by the W/Z cross section working group using an analytic parametrisation of the acceptance
  - ok since acceptance for W/Z is a very simple problem



correlation between $\sigma(W)$ and $\sigma(Z)$



correlation between A_W and A_Z

Use NLO and LO ME for t-tbar production

- Run t-tbar in MCFM
  - Total cross section at NLO
  - Lepton + jets and dilepton at LO using CDF cuts
  - Lepton+jets cuts
  - $p_T^{electron}$>20 GeV
  - $|\eta^{electron}|$<1.1
  - $p_T^{neutrino}$>20 GeV
  - R_{cone}=0.4
  - E_T^{jet}>15 GeV/c; $|\eta^{jet}|$<2.0
  - Dilepton cuts
  - lepton E_T: > 20 GeV
  - electrons: eta < 2
  - muons: eta ~< 1.1
  - jets: > 15 GeV eta < 2.5
  - MET: > 25 GeV
  - HT (leptons + MET + all jets): > 200 GeV

- For full cross section at NLO, find cross section uncertainty of 6.0 +0.42/-0.28 pb
  - +7%, -4.7%
  - a bit lower than Mangano but no resum
- For full cross section at LO, find cross section uncertainty of 5.29 + 0.31/-0.18
  - +5.9%, -3.4%
  - cross section is a bit smaller since using NLO pdf's rather than LO

Faster ways for t-tbar

- Lepton + jets
  - 94 +4.5/-2.8 fb
  - +4.8%, -3%
- Dilepton
  - 46.7 +2.8/-1.4 fb
  - +6%, -3%

- For full cross section at NLO, find cross section uncertainty of 6.0 +0.42/-0.28 pb
  - +7%, -4.7%
  - a bit lower than Mangano but no resum
- For full cross section at LO, find cross section uncertainty of 5.29 + 0.31/-0.18
  - +5.9%, -3.4%
  - cross section is a bit smaller since using NLO pdf's rather than LO

Complications/subtleties

- How much do detector effects/initial/final state radiation/ hadronization affect cross section/acceptance uncertainties?
  - correlation between ISR and pdf uncertainties, e.g. eigenvector 15
- Compare LO ME results with very high statistics MC for central pdf, a few eigenvectors and LO pdf
- Compare MC@NLO with Herwig (with spin correlations turned on/off)

Summary/discussion

- Week 2 of Santa Barbara was devoted mostly to pdf discussions with Stirling, Tung, Alekhin, Giele, etc
  - "shootout" between $\Delta\chi^2=1$ people and $\Delta\chi^2>1$ people
  - cross-studies by CTEQ and MRST currently underway to understand any differences
  - Santa Barbara accord?: MRST/CTEQ use same $\Delta\chi^2$ criteria
- See http://www.pa.msu.edu/~huston/santa_barbara/collider04.html for all talks/notes/discussion/Big Questions

Two more things

- Pedagogical guide for grad students coming out soon

Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics

Editors: *M.A. Dobbs[1], S. Frixione[2], E. Laenen[3], K. Tollefson[4]*
Contributing Authors: *H. Baer[5], E. Boos[6], B. Cox[7], M.A. Dobbs[1], R. Engel[8], S. Frixione[2], W. Giele[9], J. Huston[4], B. Kersevan[10], F. Krauss[11], Y. Kurihara[12], E. Laenen[3], L. Lönnblad[13], F. Maltoni[14], M. Mangano[11], S. Odaka[12], P. Richardson[11], A. Ryd[15], T. Sjöstrand[13], P. Skands[13], Z. Was[16], B.R. Webber[17], D. Zeppenfeld[18]*

[1]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

This work was supported in part by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

[2]INFN, Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy
[3]NIKHEF Theory Group, Kruislaan 409, 1098 SJ Amsterdam, The Netherlands
[4]Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824-1116, USA
[5]Department of Physics, Florida State University, 511 Keen Building, Tallahassee, FL 32306-4350, USA
[6]Moscow State University, Moscow, Russia
[7]Dept of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, England M13 9PL
[8]Institut fr Kernphysik, Forschungszentrum Karlsruhe, Postfach 3640, D - 76021 Karlsruhe, GERMANY
[9]Fermi National Accelerator Laboratory, Batavia, IL 60510-500, USA
[10]Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana,Slovenia; Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19,SI-1000 Ljubljana, Slovenia
[11]CERN, CH–1211 Geneva 23, Switzerland
[12]KEK, Oho 1-1, Tsukuba, Ibaraki 305-0801, Japan
[13]Department of Theoretical Physics, Lund University, S-223 62 Lund, Sweden
[14]Centro Studi e Ricerche "Enrico Fermi", via Panisperna, 89/A - 00184 Rome, Italy
[15]Caltech, 1200 E. California Bl., Pasadena CA 91125, USA
[16]Institute of Nuclear Physics PAS, 31-342 Krakow, ul. Radzikowskiego 152, Poland
[17]Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, U.K.
[18]Department of Physics, University of Wisconsin, Madison, WI 53706, USA

Abstract

Recently the collider physics community has seen signifi cant advances in the formalisms and implementations of event generators. This review is a primer of the methods commonly used for the simulation of high energy physics events at particle colliders. We provide brief descriptions, references, and links to the specifi c computer codes which implement the methods. The aim is to provide an overview of the available tools, allowing the reader to ascertain which tool is best for a particular application, but also making clear the limitations of each tool.

Compiled by the Working Group on Quantum ChromoDynamics and the Standard Model for the Workshop "Physics at TeV Colliders", Les Houches, France, May 2003.

February 25, 2004

W+jets

- I’m working on a comparison of ME+MC/CKKW-enhanced MC/NLO predictions for W + jet cross sections to see what this may tell us about improving predictions for W + jets final states