Techniques in Global Fitting of Parton Distribution Functions

Jon Pumplin – 8 October 2007
BNL Workshop on Polarized PDFs from RHIC

The goal of this talk is to give an overview of techniques used to measure Parton Distribution Functions by the CTEQ group, with an eye toward extending those techniques to measure polarized Parton Distributions.

Outline

1. Introduction to PDFs
2. Results from the unpolarized analysis
3. Details of the implementation
   - Handling correlated experimental errors
   - Estimating uncertainties
   - Eigenvector PDF sets
4. The Lagrange multiplier method
5. Outlook for polarized PDFs

Introduction to PDFs

High energy hadrons interact through their quark and gluon constituents. At short distance scales — e.g. large momentum transfer — QCD interactions become weak due to asymptotic freedom, which allows us to apply perturbation theory to a rich variety of experiments.

The complicated nonperturbative long-distance nature of the proton then shows itself only through the Parton Distribution Functions $f(Q,x)$ of momentum scale Q and light-cone momentum fraction x for each flavor. It is convenient to think of these loosely as "one-particle" probability distributions; although that is not strictly correct beyond leading order, where their definition requires singularity management choices such as $\overline{MS}$.

Goals:

1. Test of QCD
2. Needed for Signal and Background calculations at colliders
3. Measures a fundamental aspect of proton structure — perhaps testable against moments calculated on lattice. This is a particularly strong motivation to exploring the spin dependence.

Global QCD analysis

The dependence of the PDFs on Q is determined perturbatively by QCD renormalization group equations (DGLAP). Hence $f(Q,x)$ can be characterized by functions $f_a(Q_0,x)$ at a fixed small Q_0, where $a = g$, u, $\bar{u}$, $d, \ldots$. Those functions are measured by a QCD global analysis: simultaneously fitting a wide range of data from different experiments at $Q \geq Q_0$.

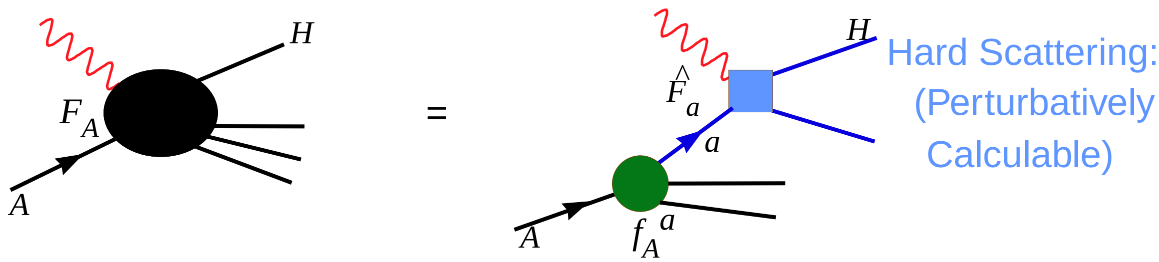
Key points:

- Factorization Theorem – Short distance and long distance are separable, PDFs are process independent
- Asymptotic Freedom – Hard scattering is perturbatively calculable
- DGLAP Evolution – Evolution in Q is perturbatively calculable, so the functions to be measured depend only on x.

If the hard-scattering data include polarization information, it should be possible to extend the PDF formalism to allow separate PDFs for polarization states of each flavor.

Factorization Theorem

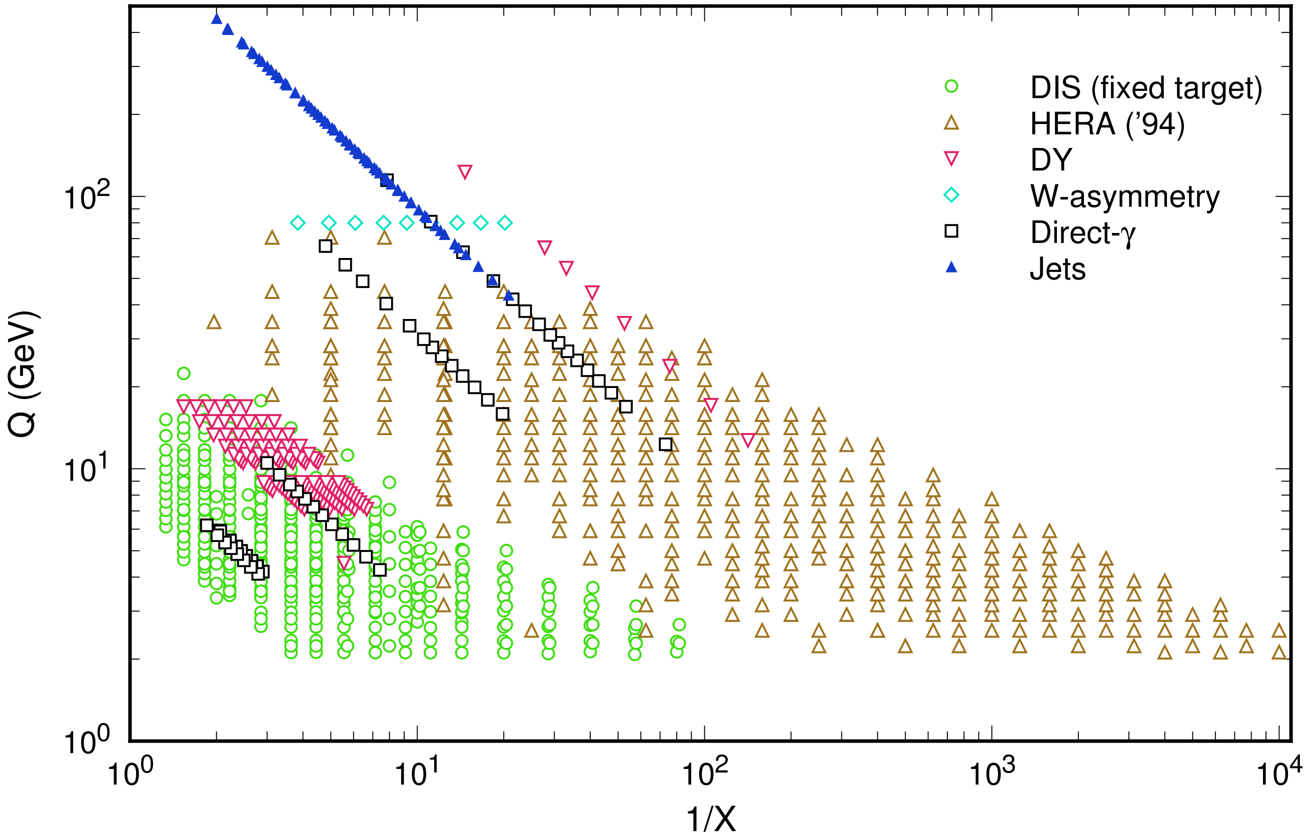
$$F_A^\lambda(x, \frac{m}{Q}, \frac{M}{Q}) = \sum_a f_A^a(x, \frac{m}{\mu}) \otimes \widehat{F}_a^\lambda(x, \frac{Q}{\mu}, \frac{M}{Q}) + \mathcal{O}((\frac{\Lambda}{Q})^2)$$



Experimental
Input

Parton Distributions:
Nonperturbative parametrization at Q_0
DGLAP Evolution to Q

Kinematic region covered by data



Data with a wide range of scales are tied together by the DGLAP renormalization group evolution equation.

Consistency or inconsistency between the different processes can be observed only by applying QCD to tie them together in a global fit.

HERA II, Tevatron run II (W, Z production), and LHC will dramatically extend the range and accuracy.

CTEQ6.6 Global analysis

Input from Experiment:

- ~ 2700 data points with $Q > 2\,\text{GeV}$ from e, μ, ν DIS; lepton pair production (DY); lepton asymmetry in W production; high p_T inclusive jets; dimuon production in neutrino scattering; HERA c and b production; HERA charged current. $\alpha_s(M_Z)$ obtained from LEP.

Data cuts:

- $Q^2 > 4\text{GeV}^2$, $W^2 > 12\text{GeV}^2$

Input from Theory:

- NLO QCD evolution and hard scattering

Assumptions:

- Parametrize at Q_0:
  $A_0\, x^{A_1} (1-x)^{A_2}\, e^{A_3 x} (1 + A_4 x + A_5 x^2)$
- $s = \bar{s}$ with $A_4 = A_5 = 0$;
- No intrinsic b or c

Method

- Construct effective $\chi^2_{\text{global}} = \sum_{\text{i=expts}} W_i\, \chi^2_i$
- Minimize χ^2_{global} to find "Best Fit" PDFs.
- Use χ^2_{global} in neighborhood of the minimum to define uncertainty limits.

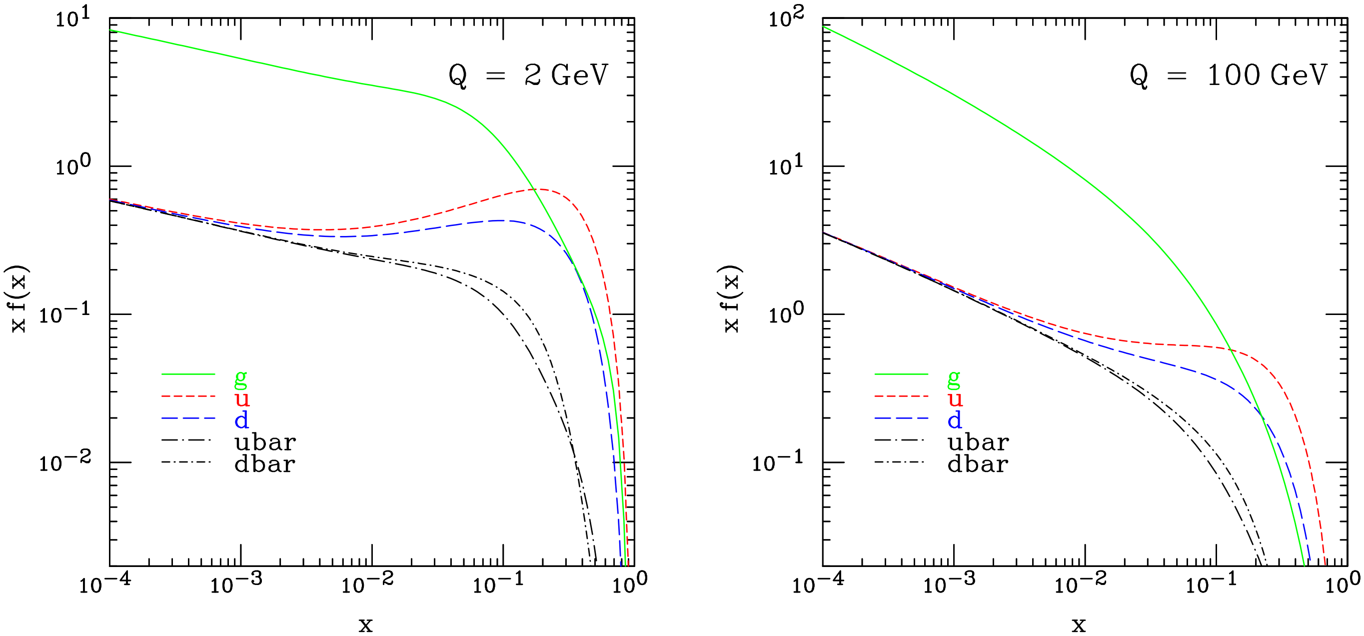
The PDF Paradigm

1. Parameterize x-dependence of each flavor at fixed small Q_0 (1.3 GeV) (parameters $A_1, \ldots, A_N$) $N \sim 20$
2. Compute PDFs $f_a(x, Q)$ at $Q > Q_0$ by DGLAP
3. Compute cross sections for DIS(e,μ,ν), Drell-Yan, Inclusive Jets,. . . by perturbation theory (NLO or NNLO — LO is not good enough, gives much worse χ^2)
4. Compute "χ^2" measure of agreement between predictions and measurements:

$$\chi^2 = \sum_i W_i \left(\frac{\text{data}_i - \text{theory}_i}{\text{error}_i} \right)^2$$

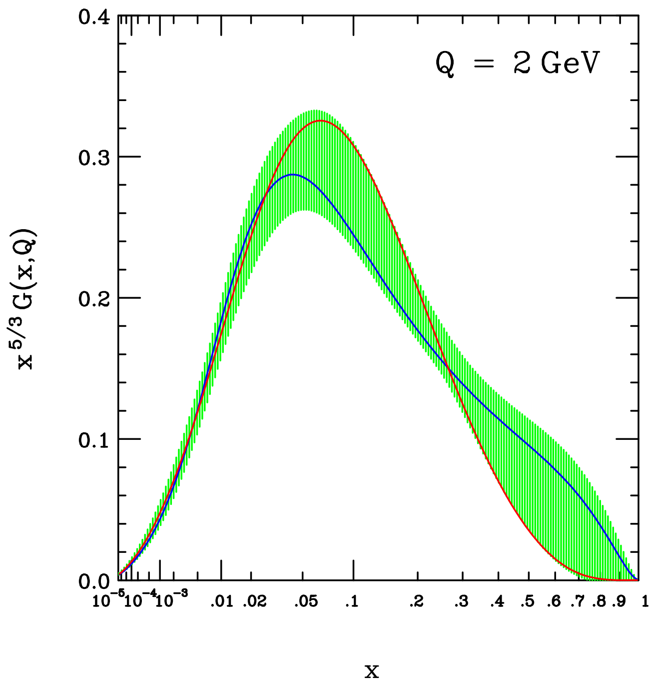
   or generalizations to include correlated systematic errors.
5. Minimize χ^2 with respect to the shape parameters $\{A_i\}$ to find Best Fit PDFs
6. PDF Uncertainty Range is the region in $\{A_i\}$ space where χ^2 is sufficiently close to minimum that all experiments are fit tolerably well.
7. Weight factors W_i are needed to force the fit to pay attention to small data sets.

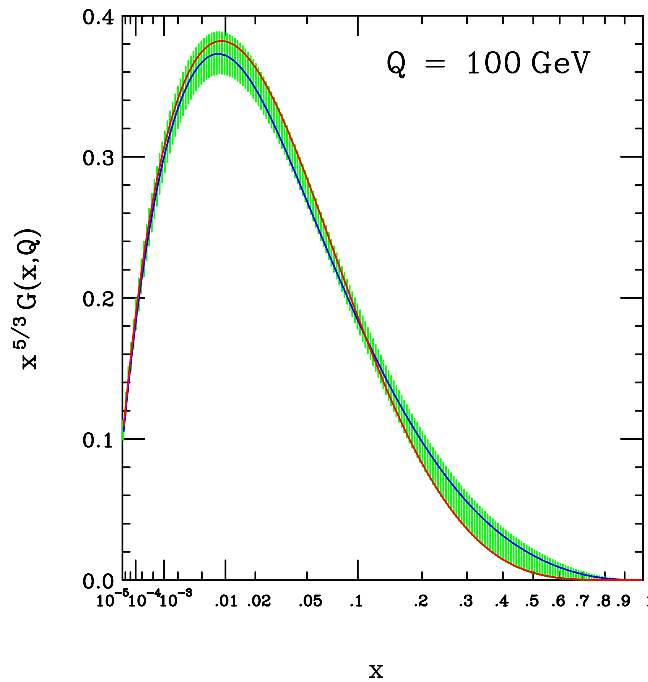
Typical results at $Q = 2$ and 100 GeV



- Valence quarks dominate for $x \to 1$
- Gluon dominates for $x \to 0$, especially at large Q

Uncertainty Results (Gluon)





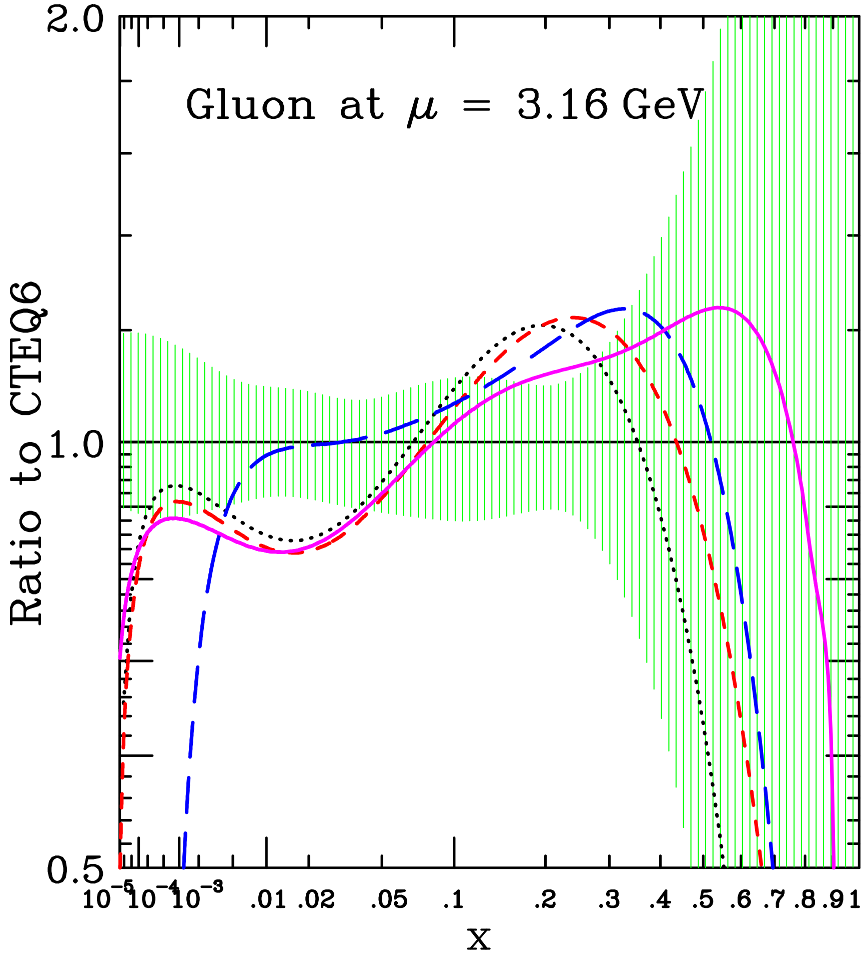
$\Delta\chi^2 = 100$ uncertainty band
Weight 50 for CDF Jets
Weight 50 for DØ Jets

Consistency check: Uncertainty estimated by $\Delta\chi^2 = 100$ (for 1811 data points) is comparable to the difference between "pulls" of nominally similar experiments.

(Area under curve is proportional to momentum fraction carried by gluon — strongly constrained by DIS data. Hence the envelope itself is not an allowed solution.)

Convergent Evolution: Uncertainty is smaller at large Q.

Fractional uncertainty of gluon



mrst2001, mrst2002, mrst2003, mrst2004

- large uncertainty at large x
- Differences between MRST and CTEQ are comparable to the estimated uncertainty—ironic because original motive to study uncertainty systematically was the danger that comparing groups using same basic method would underestimate the uncertainty!

Part II
Some details of the implementation

Including systematic errors in χ^2

The simplest definition

$$\chi_0^2 = \sum_{i=1}^{N} \frac{(D_i - T_i)^2}{\sigma_i^2} \qquad \begin{cases} D_i = & \text{data} \\ T_i = & \text{theory} \\ \sigma_i = & \text{“expt. error”} \end{cases}$$

is optimal for random Gaussian errors,

$$D_i = T_i + \sigma_i r_i \qquad \text{with} \qquad P(r) = \frac{e^{-r^2/2}}{\sqrt{2\pi}}.$$

With systematic errors,

$$D_i = T_i(a) + \alpha_i r_{\text{stat},i} + \sum_{k=1}^{K} r_k \beta_{ki} \,.$$

The fitting parameters are $\{a_\lambda\}$ (theoretical model) and $\{r_k\}$ (corrections for systematic errors).

Published experimental errors:

- α_i is the ‘standard deviation’ of the random uncorrelated error.
- β_{ki} is the ‘standard deviation’ of the k th (completely correlated!) systematic error on D_i.

To take into account the systematic errors, we define

$$\chi'^2(a_\lambda, r_k) = \sum_{i=1}^{N} \frac{(D_i - \sum_k r_k \beta_{ki} - T_i)^2}{\alpha_i^2} + \sum_k r_k^2,$$

and minimize with respect to $\{r_k\}$. The result is

$$\hat{r}_k = \sum_{k'} \left(A^{-1}\right)_{kk'} B_{k'}, \qquad \text{(systematic shift)}$$

where

$$A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N} \frac{\beta_{ki}\beta_{k'i}}{\alpha_i^2}$$

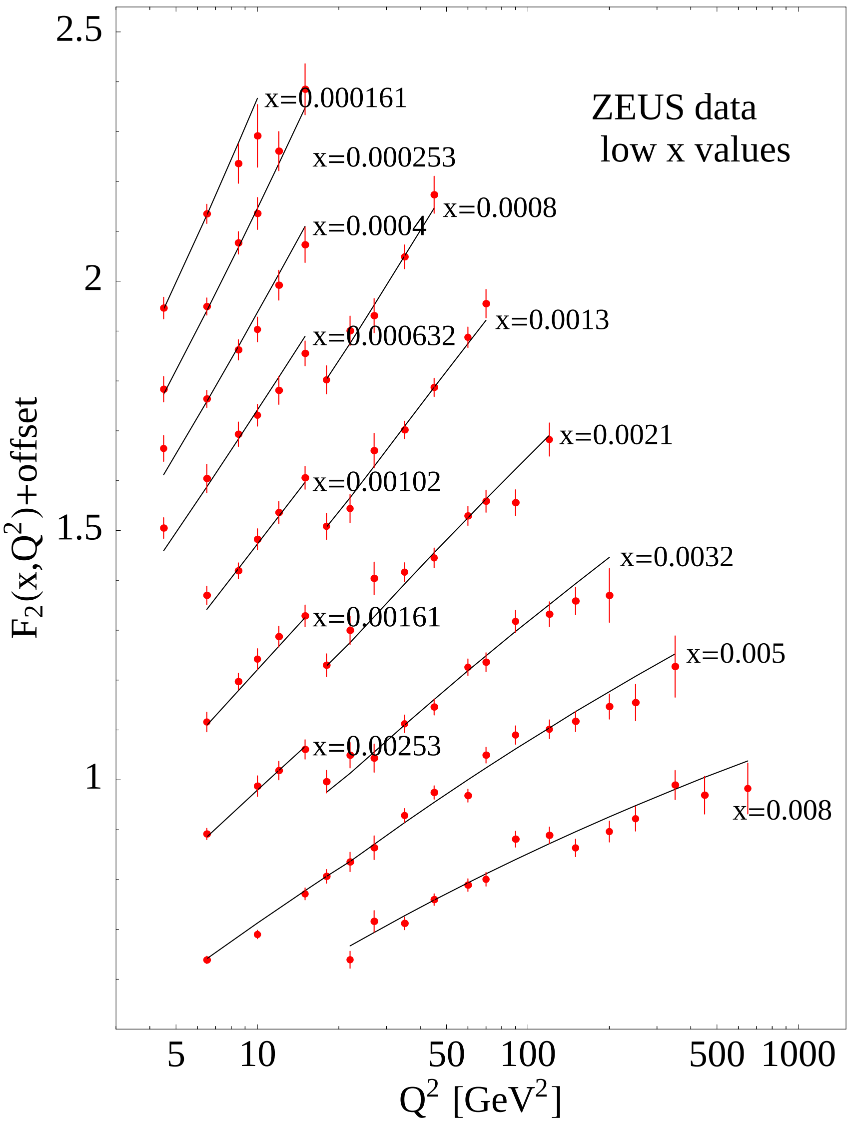
$$B_k = \sum_{i=1}^{N} \frac{\beta_{ki}(D_i - T_i)}{\alpha_i^2}.$$

The $\hat{r}_k$'s depend on the PDF model parameters $\{a_\lambda\}$. We can solve for them explicitly since the dependence is quadratic, so they don't add to the number of parameters fed to Minuit.

We then minimize the remaining $\chi^2(a)$ with respect to the model parameters $\{a_\lambda\}$.

- $\{a_\lambda\}$ determine $f_i(x, Q_0^2)$.
- $\{\hat{r}_k\}$ are are the optimal "corrections" for systematic errors; i.e., systematic shifts to be applied to the data points to bring the data from different experiments into compatibility, within the framework of the theoretical model.

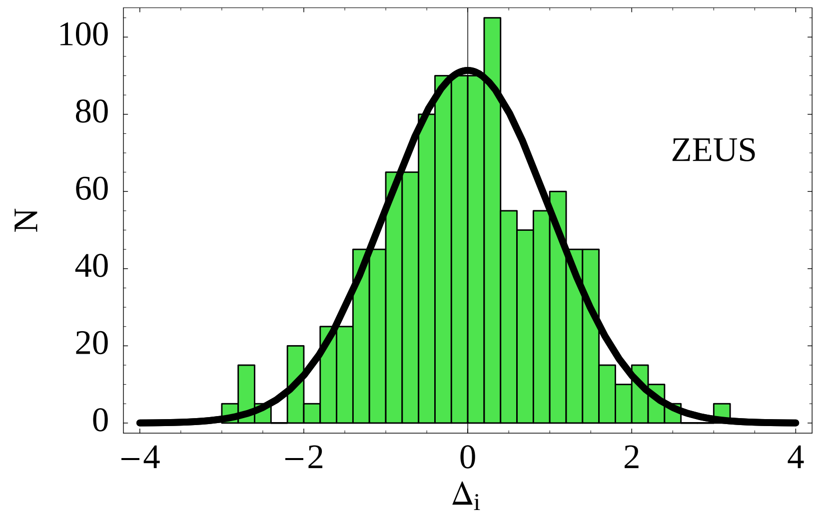
CTEQ6M fit to ZEUS data at low x



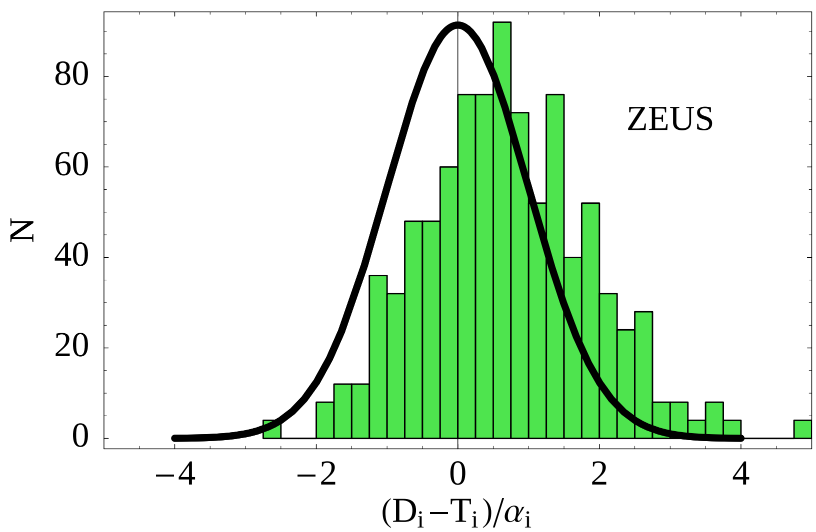
The data points include the estimated corrections for systematic errors. That is to say, the central values plotted have been shifted by an amount that is consistent with the estimated systematic errors, where the systematic error parameters are determined using other experiments via the global fit.

The error bars are statistical errors only.

Systematic Error treatment works



(a) Histogram of residuals for the ZEUS data. The curve is a Gaussian of width 1.



(b) Similar comparison without corrections for systematic errors on the data points.

Minimizing χ^2

The PDFs are determined by finding the fitting parameters $\{A_i\}$ that minimize χ^2. This job is nontrivial because

1. There are lots of parameters — currently 22.
2. To reduce the dependence on parametrization assumptions, new parameters are included in the fit until it is barely stable. Hence there are some "nearly flat" directions, which leaves the χ^2 surface quadratic only very close to the minimum.
3. The parameters are highly correlated.
4. Evaluation of χ^2 for a single choice of $\{A_i\}$ takes a few seconds, so efficiency is needed.

I have extended the classic Minuit to include an iterative method that converges to obtain the eigenvectors of the Hessian matrix.

- Other groups restrict the number of parameters (10-15) to avoid convergence problems at the cost of more dependence on the assumed parametrization.
- When implementing the calculation, it is crucial that χ^2 is a smooth function of the parameters — e.g., don't use adaptive integration in the theory calculations!

Minimizing χ^2 - ctd

In the neighborhood of the minimum, χ^2 can be approximated by a quadratic form

$$\chi^2 = \chi_0^2 + \sum_{ij} H_{ij}\,(A_i - A_i^{(0)})\,(A_j - A_j^{(0)})$$

where the Hessian matrix H is the inverse of error matrix.

After the iteration has converged, this is expressed in a diagonal form by making use of the eigenvectors of H:

$$\chi^2 = \chi_0^2 + \sum_i z_i^2$$

$$A_i = A_i^{(0)} + \sum_j w_{ij}\,z_j$$

In this way, χ^2 is probed in each eigenvector

direction at the appropriate scale of $\Delta\chi^2$. The distances moved in parameter space range from very small ("steep directions", well-determined features) to very large ("flat" directions).

The ratio of these distances corresponds to the spectrum of eigenvalues of the Hessian, which span a range of many orders of magnitude.

Uncertainties from eigenvector sets

The uncertainty of PDFs can be characterized by a collection of fits that are created by stepping away from the minimum of χ^2 along each eigenvector direction of the local quadratic form (Hessian matrix). The distance to go along each direction must be such that the fit remains "acceptable". In CTEQ6.1, this was estimated to be $\Delta\chi^2 = 100$ for 90% confidence. (1811 data points)

The current CTEQ6.6 fits have 22 free parameters and hence 44 "Eigenvector uncertainty sets".

The PDF uncertainty for any predicted quantity is obtained by evaluating that quantity with each of the eigenvector sets and then applying a simple asymmetric formula: the square root of the sum of the squares of the upward (downward) deviations from the value given by the central fit gives the estimated upper (lower) limit.

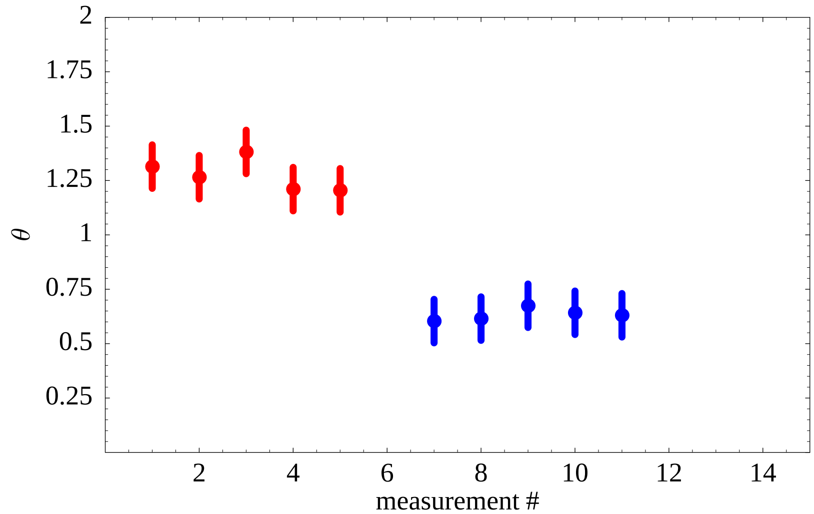
Sources of uncertainty

- Experimental errors included in χ^2
- Unknown experimental errors
- Higher-order QCD corrections + Large Logs
- Power Law QCD corrections ("higher twist")
- Parametrization dependence

Essential Difficulties

- Experiments run until systematic errors dominate $\Rightarrow$ remaining systematic errors involve guesswork
- Systematic errors of the theory and their correlations are even harder to guess
- Some combinations are unconstrained, e.g. $s-\bar{s}$ before NuTeV dimuon data
- No data at very small x (after $Q > 2\,\text{GeV}$, $W^2 > 12\,\text{GeV}$ cuts)
- Major problem: experiments are not consistent with each other. (Note, we would have been easily mislead if we had only one of H1/Zeus or CDF/DØ!)

The basic uncertainty issue



Suppose θ is measured in two different experiments. What do you quote as Best Fit and Uncertainty?

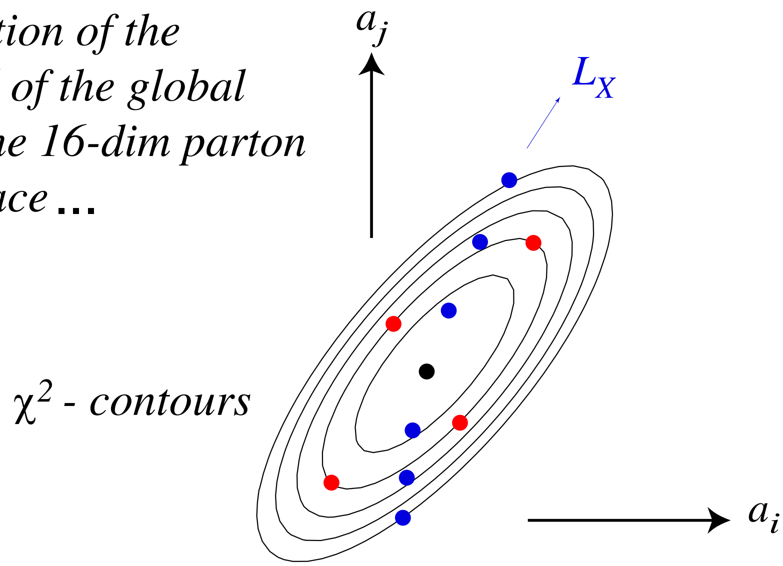
(Maybe you expand the errors so the uncertainty range covers both data sets. Or maybe you expand it even more, using the difference between experiments as a measure of the uncertainty.)

What happens to the Best Fit value when the relative weight of the two experiments is varied?

That is the method we use to assess uncertainties of the PDF Global Fit: we vary weights of the experiments to estimate a range of acceptable $\Delta\chi^2$ above the minimum value, in place of the classical $\Delta\chi^2 = 1$.

Complementary uncertainty methods

2-dim illustration of the neighborhood of the global minimum in the 16-dim parton parameter space ...



- Hessian Matrix / Eigenvector Method:
  Eigenvectors of error matrix yield 44 sets $\{S_i^{\pm}\}$ that are tolerably good fits to all of the data. Get uncertainty of any prediction by sum of squares of deviations; or more crudely just from the extremes from the 44 sets.

- Lagrange Multiplier Method:
  Track χ^2 as function of F (e.g. σ_W) by minimizing $\chi^2 + \lambda F$. Yields special-purpose PDFs that give extremes of σ_W, or $\langle y \rangle$ for rapidity distribution of W, or σ for $t\bar{t}$ production; or . . .

Outlook for Spin-dependent PDFs

- The standard PDF analysis already deals with quantities that are very well determined, such as $u(x)$ at moderate x, along with quantities that correspond to "flat" directions, such as $g(x)$ at large x. Hence no problem is anticipated in extracting information on quantities like spin dependence that are not strongly constrained.

- Rather than making a new full global fit that includes all of the current unpolarized data set used in the CTEQ fit, together with the polarized data to constrain the spin-dependent degrees of freedom, it may be possible to do a good job much more easily by freezing the unpolarized PDFs at their current forms in the CTEQ fit, and just fitting the new helicity-dependent functions to the polarized data.

What needs to be done?

The following issues need to be handled to extract spin-dependent PDFs from the polarized data. Perhaps some of them have already been solved by Spin Experts.

1. What are the spin-dependent PDFs that one can hope to extract? — is it just replacing each unpolarized parton distribution (d, u, s, $\bar{d}$, $\bar{u}$, $\bar{s}$, g) by a pair of functions with helicity parallel or antiparallel to the proton helicity?
2. How should these functions be parametrized at Q_0? — models will be needed at first to keep the number of new fitting parameters small enough to be determinable by the data.
3. Is a NLO DGLAP evolution package available for the polarized PDFs?
4. What data sets are or will be available? — can we afford the luxury of cuts like $Q^2 > 4\,\mathrm{GeV}^2$ to suppress non-leading twist effects?
5. Are NLO calculations of the relevant polarization-dependent observables available? (Or may LO calculations be good enough for the polarization?)