CTEQ6.1 PDF analysis

J. Pumplin, D. Stump, W.K. Tung, J. Huston (MSU)
S. Kuhlmann, J. Owens, H. Lai, P. Nadolsky

Work in Progress

1. Technical improvements to the CTEQ6 analysis
   - recalculate K-factors along each eigenvector direction
   - parametrization dependence study (20 vs. 23 (vs. 15?) fit parameters)
2. Inclusive Jet production: Tevatron RunII/RunI, LHC
3. Systematic error studies: Jet Scale dependence
4. Improved treatment of heavy quarks

A coherent picture of PDF uncertainties has emerged: there is reasonable agreement between various methods for estimating the uncertainty –

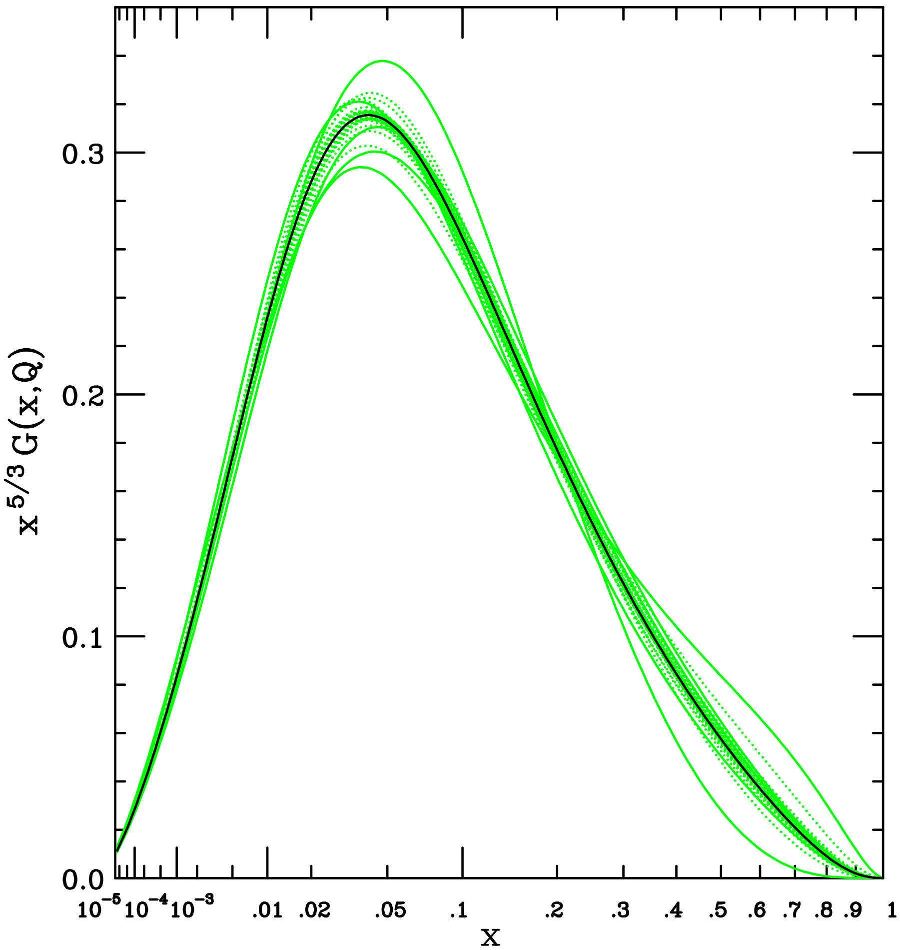
- "Hessian Method" – eigenvectors of the error matrix
- "Lagrange Multiplier Method" – variation of χ^2
- reweighting of experiments
- statistical bootstrap method

Illustrate convergence of uncertainty estimates by two examples:

1. gluon distribution at $Q = 3.16\,\text{GeV}$
2. cdf inclusive jet ratio
   $\frac{d\sigma}{dp_T}(1.96\ \text{TeV})/\frac{d\sigma}{dp_T}(1.80\ \text{TeV})$

These two examples are closely connected, because the existing Tevatron RunI jet measurements provide the main constraint on gluon distributions.

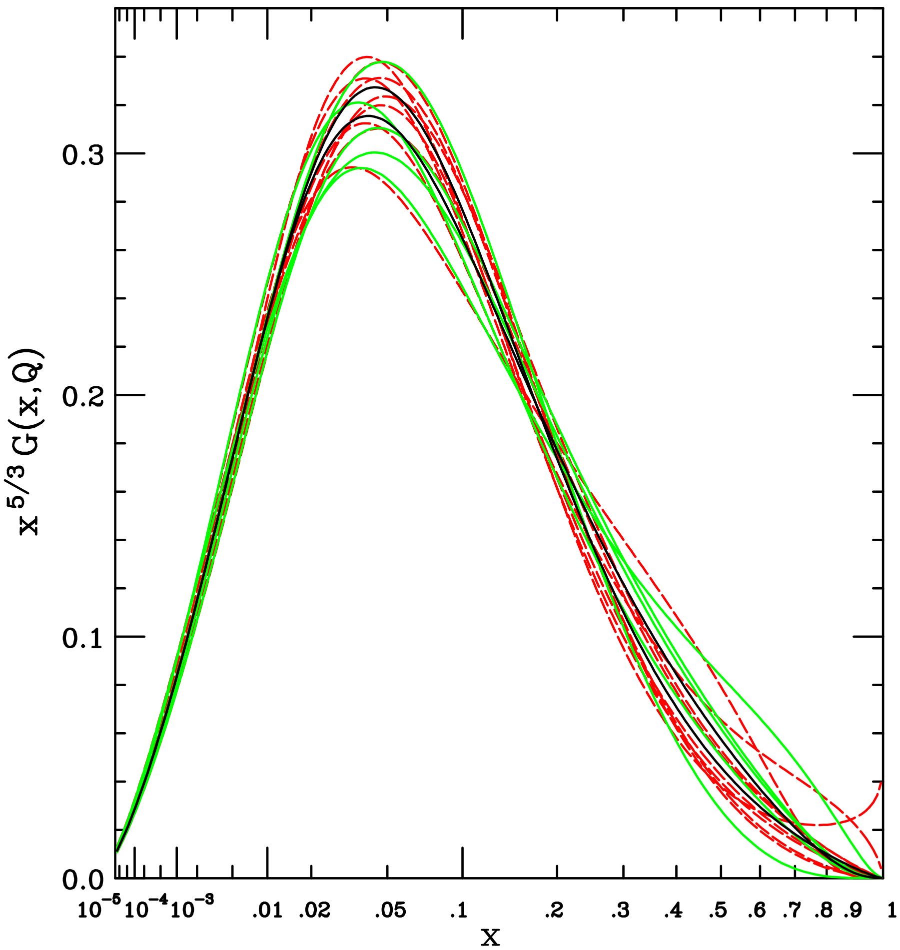
Gluon distribution at $Q = 3.16\,\text{GeV}$



Best Fit (solid) and 40 Eigenvector sets (dotted) of CTEQ6.1 with $\Delta\chi^2 = 100$.

Area under curve is proportional to momentum fraction carried by gluon, which is strongly constrained by DIS data. Hence if $G(x)$ is larger than the central value at $x \approx 0.5$ it must be smaller than the central value at $x \approx 0.05$.

Eigenvector sets with 23 free parameters



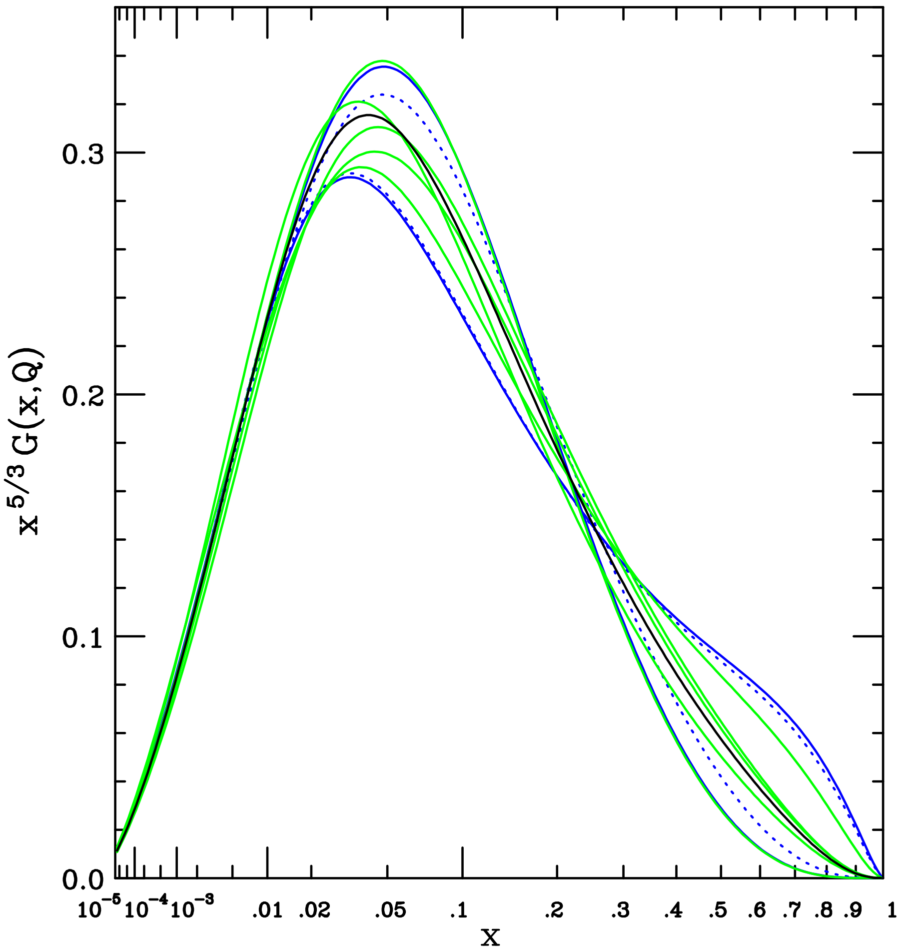
Red curves: Extreme eigenvector sets when 23 free parameters are used.

Green curves: CTEQ6.1 (20 free parameters)

In this and following figures, both methods give similar uncertainty.

A qualitative difference appears only in the extreme $x \to 1$ region, where the 23 parameter fit allows a mildly singular behavior that is not ruled out by the global fit.

Lagrange Multiplier sets with $\Delta\chi^2 = 100$

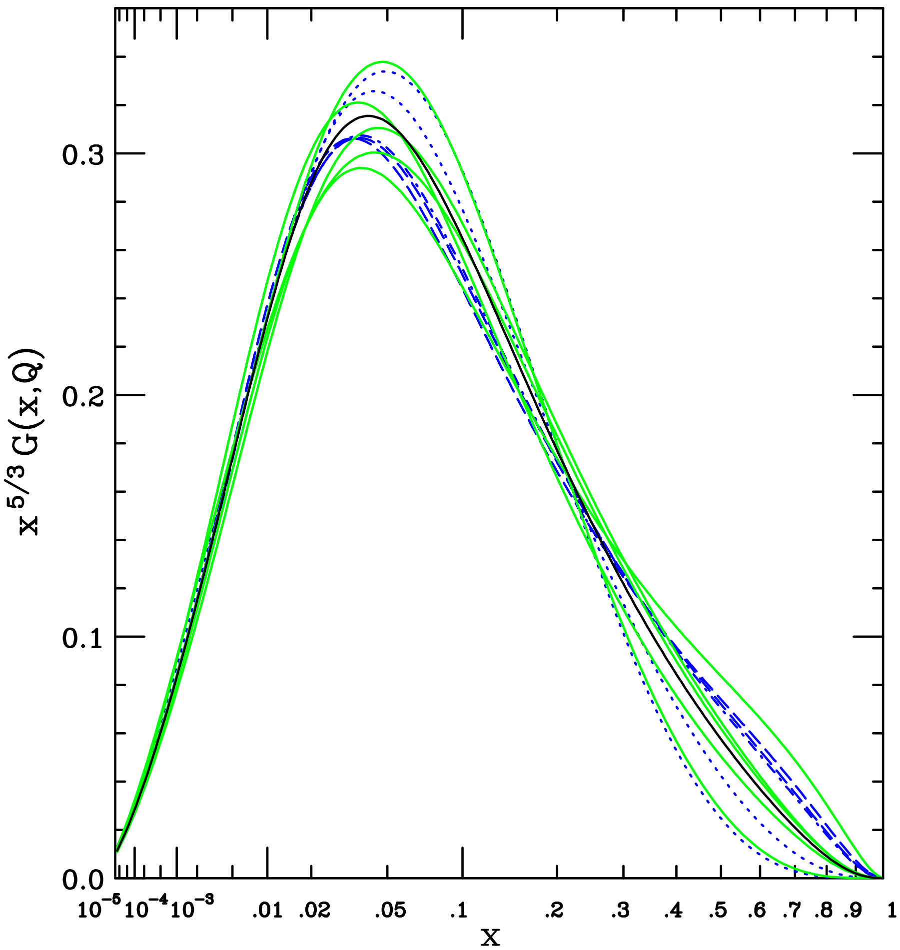


Lagrange Multiplier method used to extremize

1. solid: $(\sqrt{s} = 1.96\,\mathrm{TeV})/(\sqrt{s} = 1.80\,\mathrm{TeV})$ for jets at $p_T = 413\,\mathrm{GeV}$ (last cdf RunI point)
2. dotted: $(p_T = 413\,\mathrm{GeV})/(p_T = 362\,\mathrm{GeV})$ for jets at $\sqrt{s} = 1.8\,\mathrm{TeV}$ (last two cdf RunI points)

Lagrange Multiplier method shows slightly larger uncertainty range, but does so at the expense of marginal fits to CDF and DØ jet experiments, because jet predictions were pushed by LM.

Extra weight for jet data



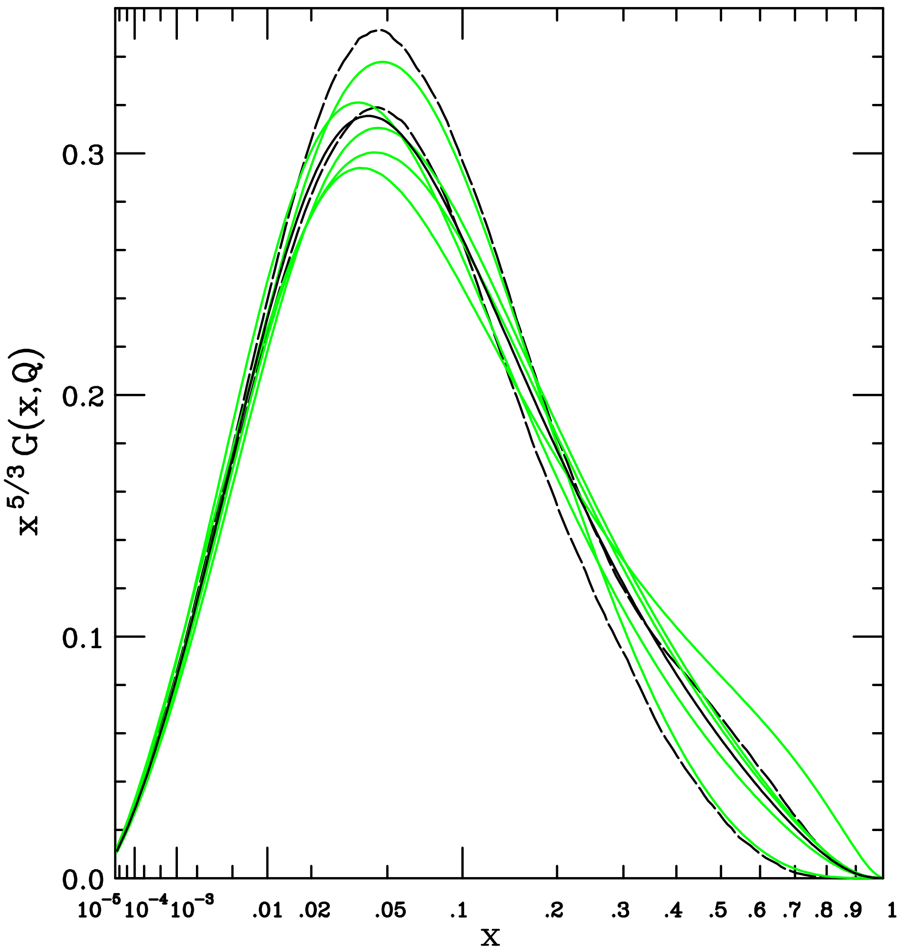
Assign extra weight to Tevatron run I inclusive jet measurements:

- dotted: CDF factor 10, 50
- dashed: DØ factor 10, 50
- dotdash: CDF and DØ factor 10

These are the experiments that most constrain the gluon distribution. The variation is comparable to the 40 eigenvector sets (green curves = CTEQ6.1)

DØ prefers stronger high-x gluon.

Statistical Bootstrap method



Continuous bootstrap method: generate random weights for each of the 16 experiments in global fit by $\frac{dP}{dW_i} = e^{-W_i}$. Find best fit for each set of weights. Repeat 200 times and take the central 90 % at each x as the measure of uncertainty.

Shows a sizable uncertainty with no *ad hoc* assumption such as $\Delta\chi^2 = 100$.

Traditional statistical bootstrap uses integer weights $0 - 16$ defined by random selection. Continuum method is similar but avoids zero weights.

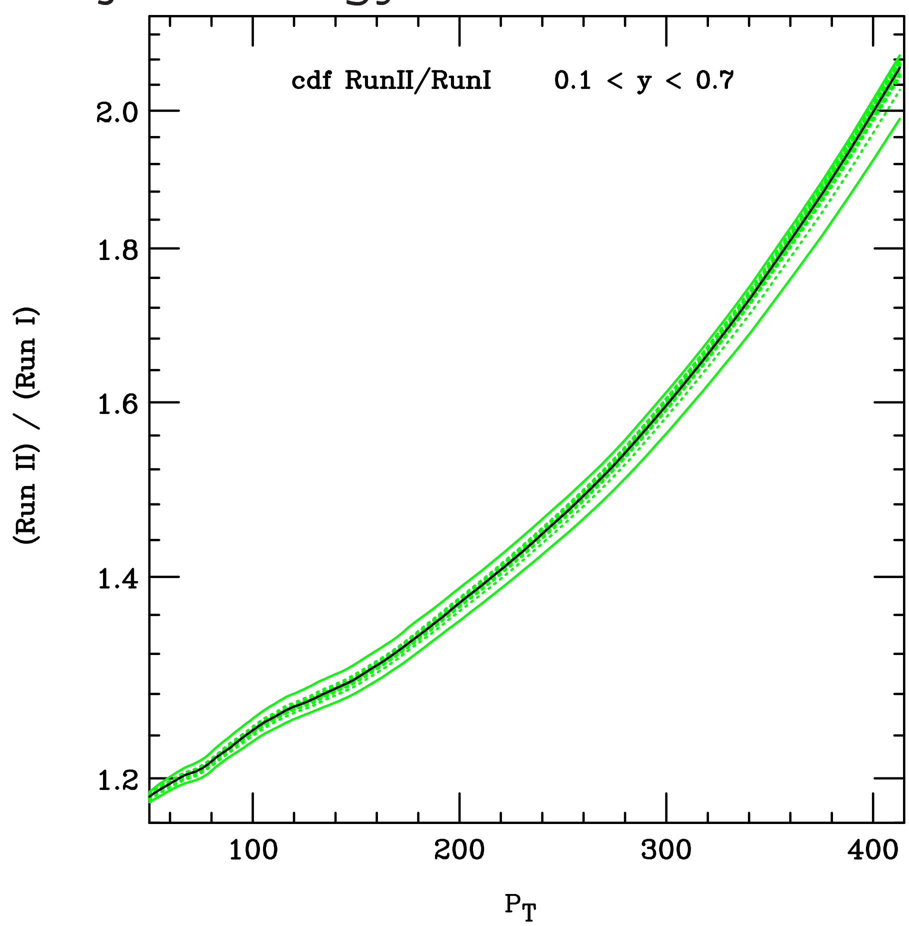
(Central value is slightly different because old K-factor method was used.)

Tevatron inclusive jet ratio

The inclusive jet energy dependence

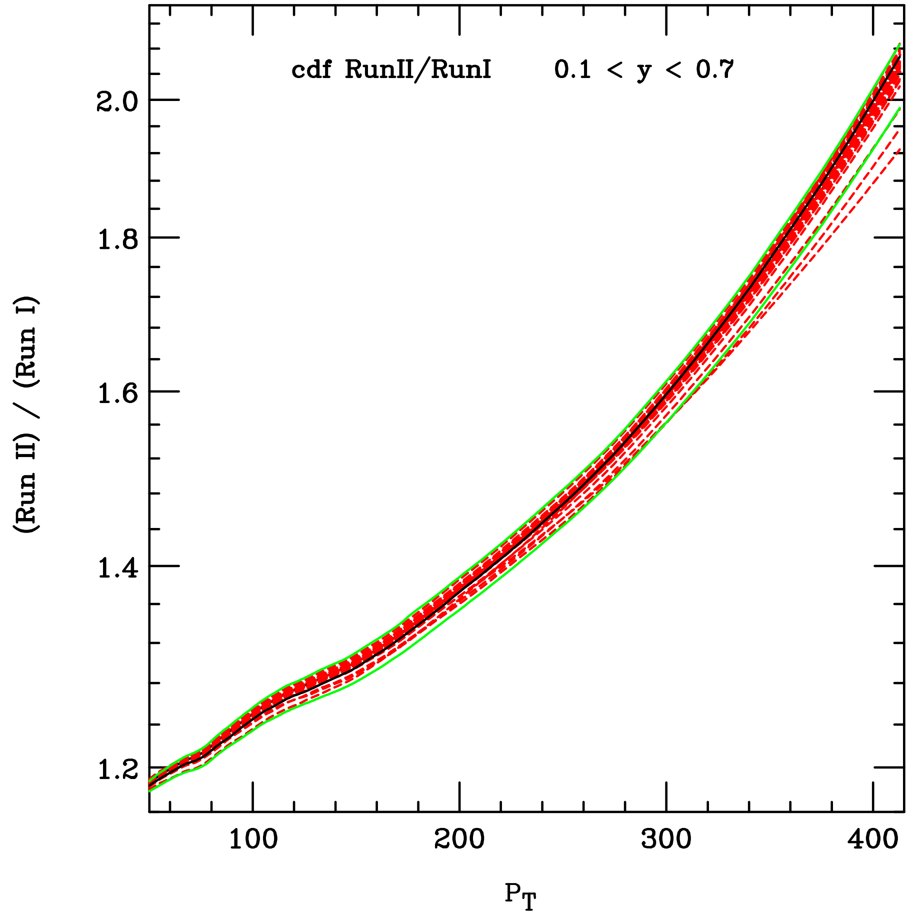
$$\frac{\frac{d\sigma}{dP_T}(1.96\,\text{TeV})}{\frac{d\sigma}{dP_T}(1.80\,\text{TeV})}$$

between Tevatron Run I and run II will eventually give a sensitive test of QCD and a probe for quark substructure, because many systematic errors cancel. At present, it is an important check on the experimental jet "energy scale" calibration.



Best Fit (black) and 40 Eigenvector sets (green) with $\Delta\chi^2 = 100$ of CTEQ6.1. Extremes are given by eigenvector sets +15 and −15 (solid).

Inclusive jet ratio with 23 free parameters

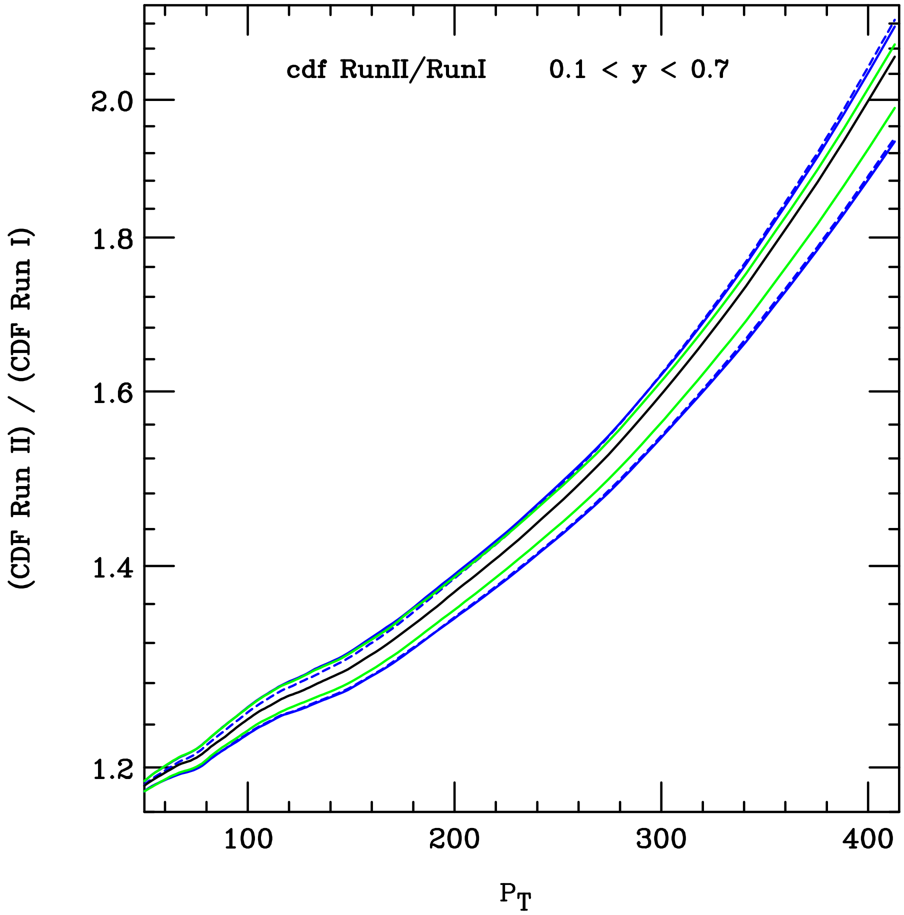


Red curves: Extreme eigenvector sets when 23 free parameters are used.
Green curves: Extremes from CTEQ6.1 (20 free parameters)

Going from 20 fitting parameters to 23 leads to a small expansion in the uncertainty range at large P_T.

(This suggests that the MRST form with only 15 parameters may underestimate the uncertainty; but I have not yet checked that.)

Inclusive jet ratio by Lagrange Multiplier

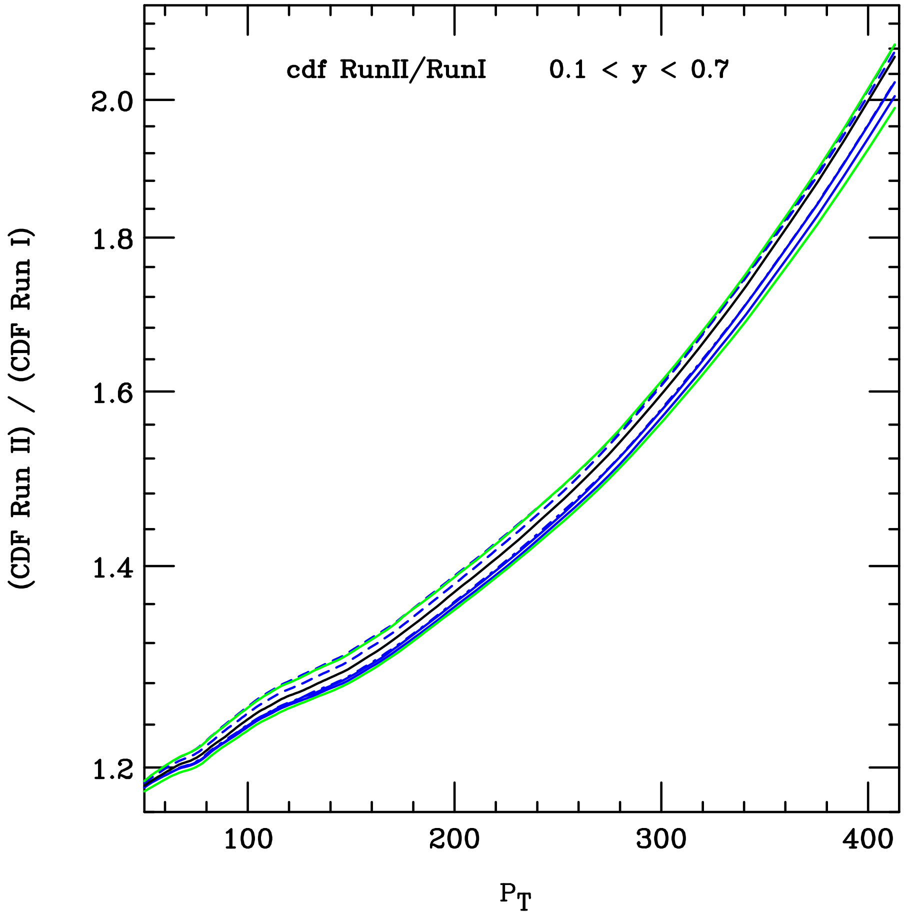


Lagrange Multiplier method used to extremize

1. solid: $(\sqrt{s} = 1.96\,\text{TeV})/(\sqrt{s} = 1.80\,\text{TeV})$ for jets at $p_T = 413\,\text{GeV}$ (last cdf RunI point)
2. dotted: $(p_T = 413\,\text{GeV})/(p_T = 362\,\text{GeV})$ for jets at $\sqrt{s} = 1.8\,\text{TeV}$ (last two cdf RunI points)

Lagrange Multiplier method suggests slightly larger uncertainty range than the 40 eigenvector sets. But does so at the expense of marginal fits to the existing RunI CDF and DØ jet experiments, which bear a sizable fraction of the $\Delta\chi^2 = 100$.

Inclusive jet ratio with extra weight for existing jet data



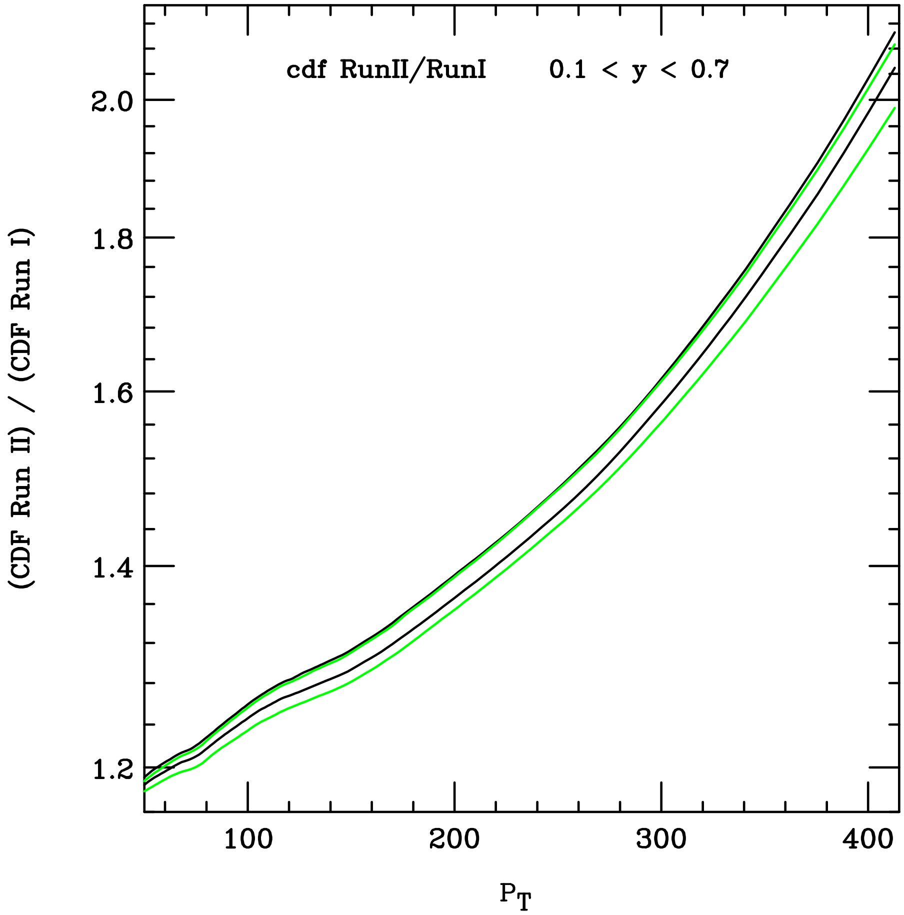
Blue curves: Extra weight assigned to Tevatron run I jet measurements:

- dotted: CDF factor 10, 50
- dashed: DØ factor 10, 50
- dotdash: CDF and DØ factor 10

The variation is comparable to that of the 40 eigenvector sets of CTEQ6.1 (Green curves)

DØ and CDF pull in different directions – possibly because DØ covers a range of rapidities.

Inclusive jet ratio by statistical bootstrap

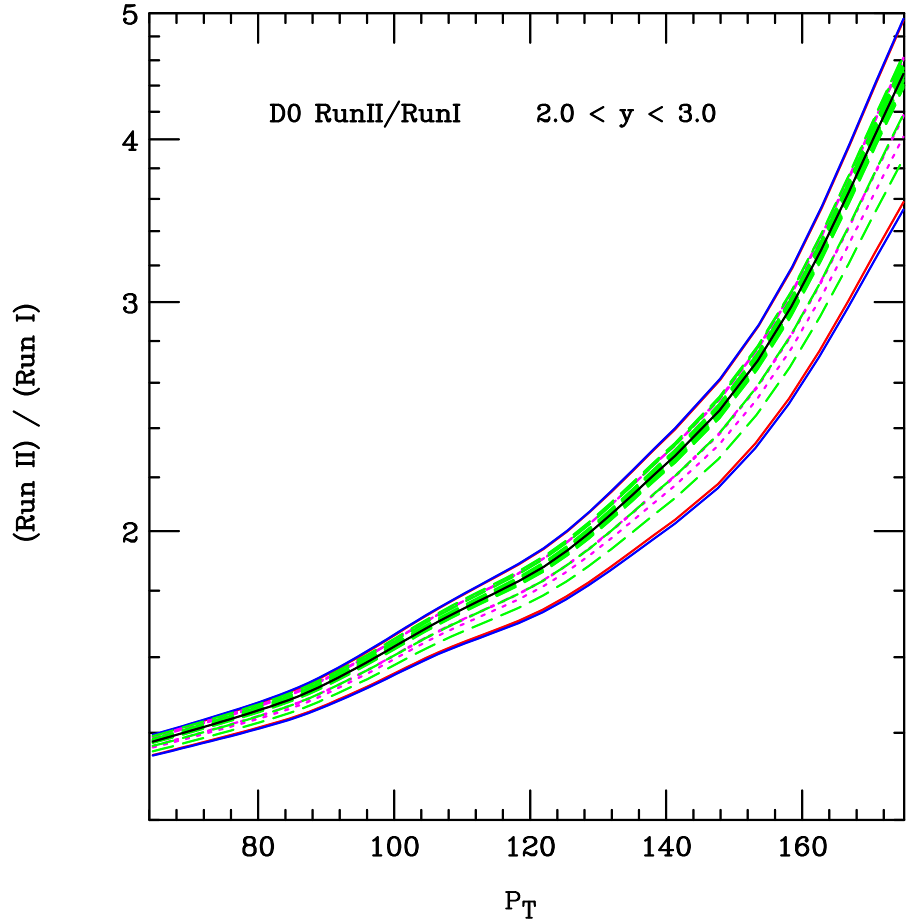


Black curves: Continuous bootstrap method

Generate random weights for the 16 experiments of global fit by $\frac{dP}{dW_i} = e^{-W_i}$. Find best fit with each set of weights. Repeat 200 times and take the central 90 % at each x as the measure of uncertainty.

The uncertainty range is comparable to that of the 40 eigenvector sets of CTEQ6.1 (Green curves). (The central value is somewhat different because the bootstrap sets were generated using older K-factor treatment.)

Inclusive jet ratio at forward rapidity



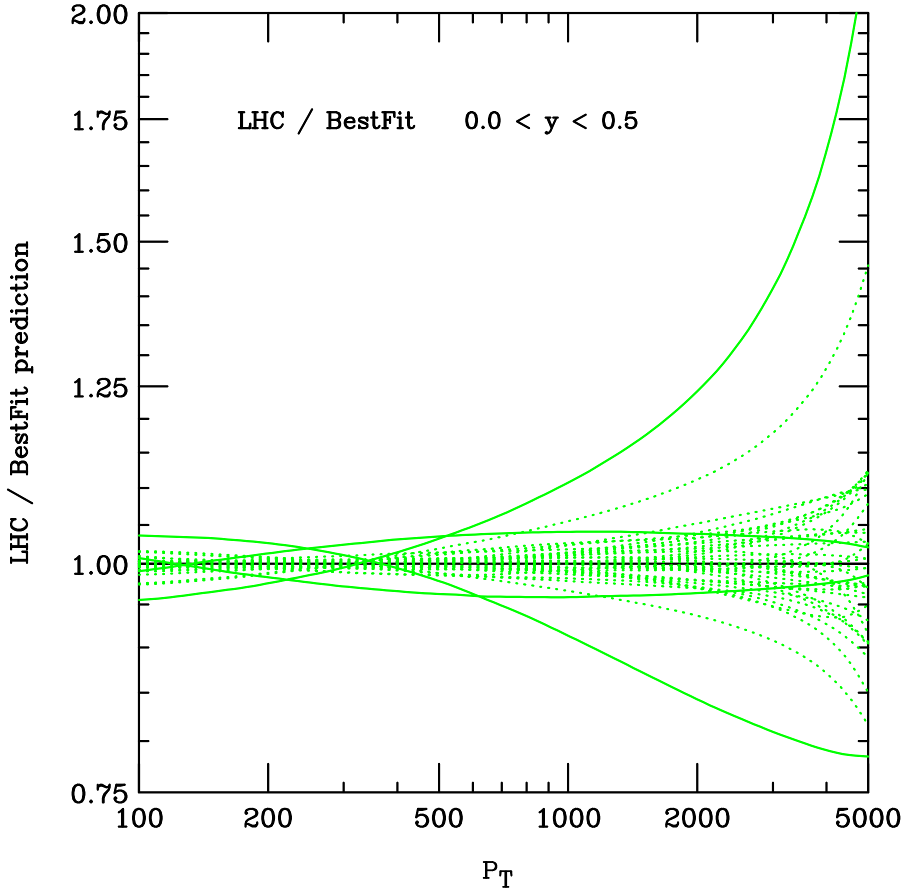
$$\frac{\frac{d\sigma}{dP_T}(1.96\,\text{TeV})}{\frac{d\sigma}{dP_T}(1.80\,\text{TeV})}$$

for $2 < \eta < 3$ – all uncertainty estimates.

References for Statistical Bootstrap method

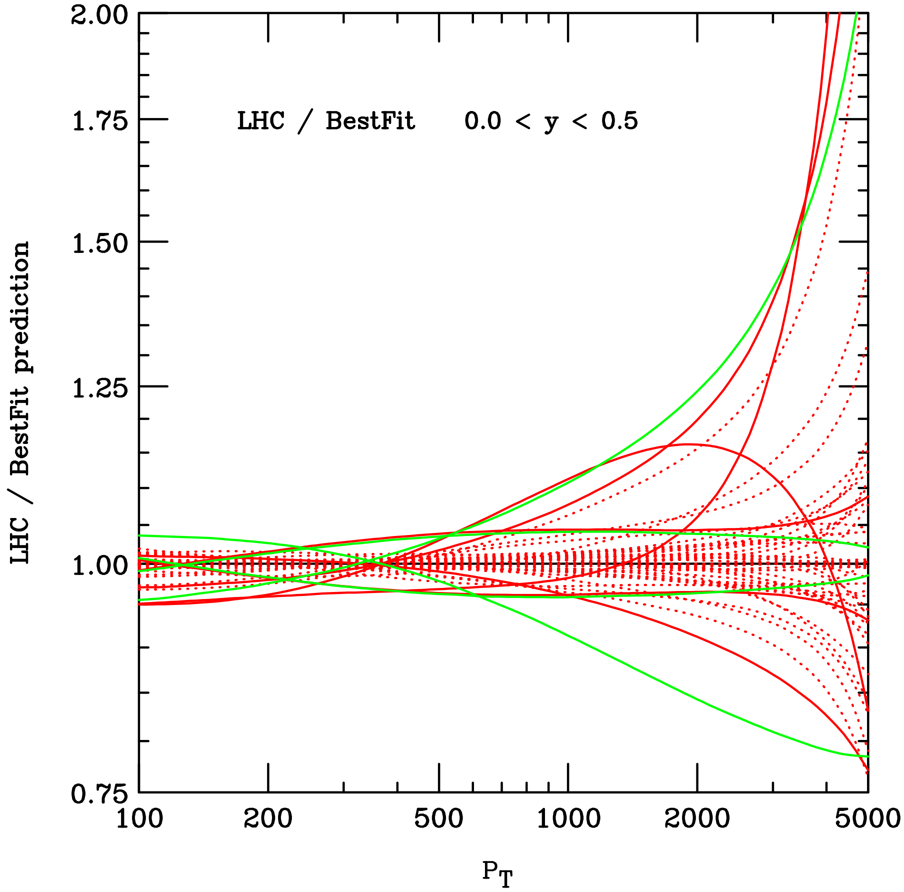
1. Efron and Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall 1993.
2. M. Chernick, *Bootstrap Methods: A Practitioner's Guide*, John Wiley & Sons 1999.

LHC inclusive jet uncertainty



Fractional uncertainty of prediction for LHC inclusive jet cross section. green: 40 Eigenvector sets with $\Delta\chi^2 = 100$ of CTEQ6.1. Extremes are given by eigenvector sets +5, −5, +15, −15 (solid).

LHC jet uncertainty with 23 free parameters

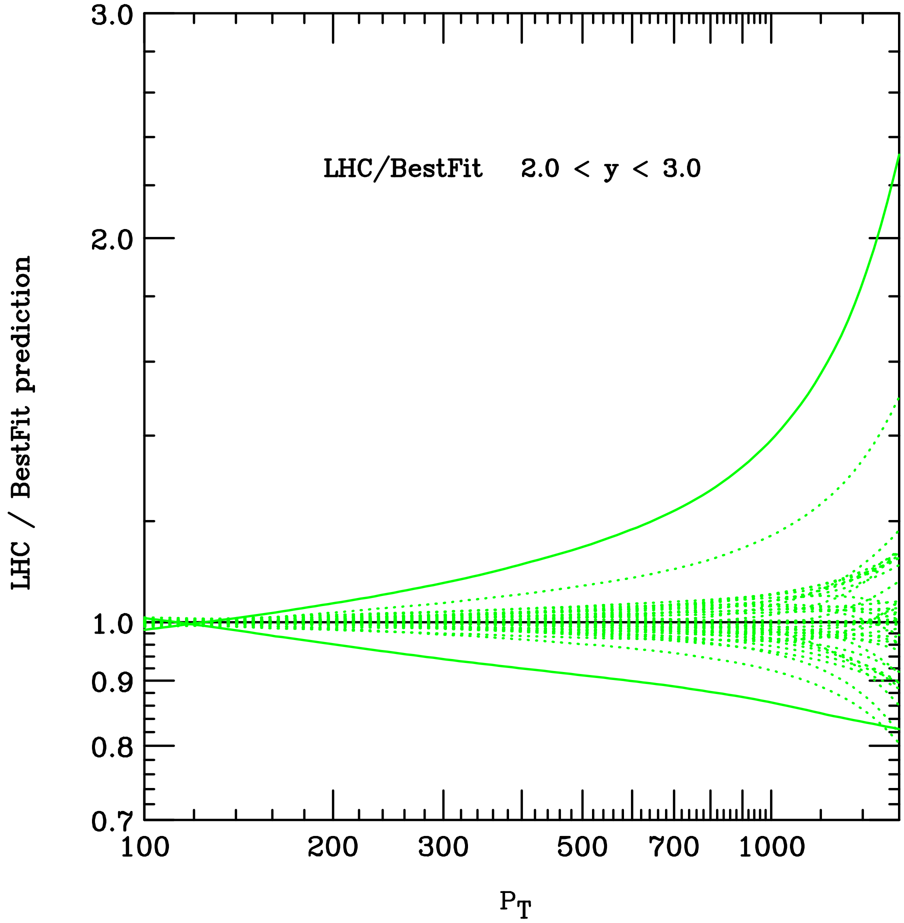


Red curves: Extreme eigenvector sets when 23 free parameters are used.
Green curves: Extremes from CTEQ6.1 (20 free parameters)

Predictions become very uncertain at $P_T > 4000$ GeV.

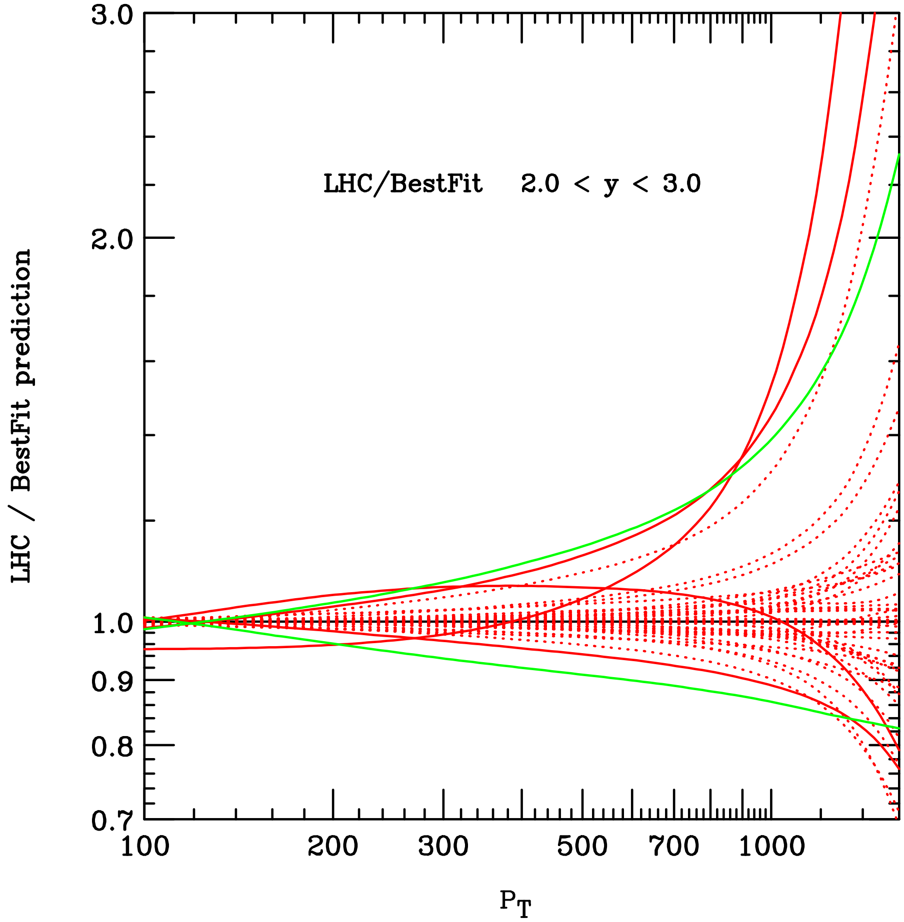
(Electroweak contribution not included.)

LHC jet uncertainty at $2 < y < 3$



Fractional uncertainty of prediction for LHC inclusive jet cross section. green: 40 Eigenvector sets with $\Delta\chi^2 = 100$ of CTEQ6.1. Extremes are given by eigenvector sets +15, −15 (solid).

LHC jet uncertainty at $2 < y < 3$ with 23 free parameters



Red curves: Extreme eigenvector sets when 23 free parameters are used.
Green curves: Extremes from CTEQ6.1 (20 free parameters)

Looks like a repeat of the HJ saga: PDF uncertainty allows a much higher jet cross section at large p_T than the Best Fit would have predicted.