Parton distributions for the LHC era

Pavel Nadolsky (Southern Methodist University)

presented by

Joey Huston (Michigan State University)

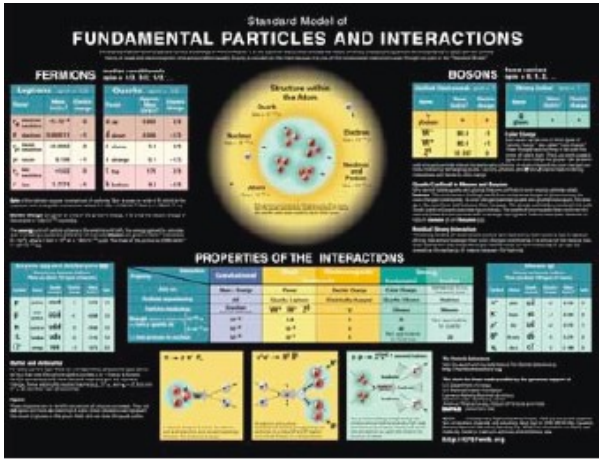
in collaboration with

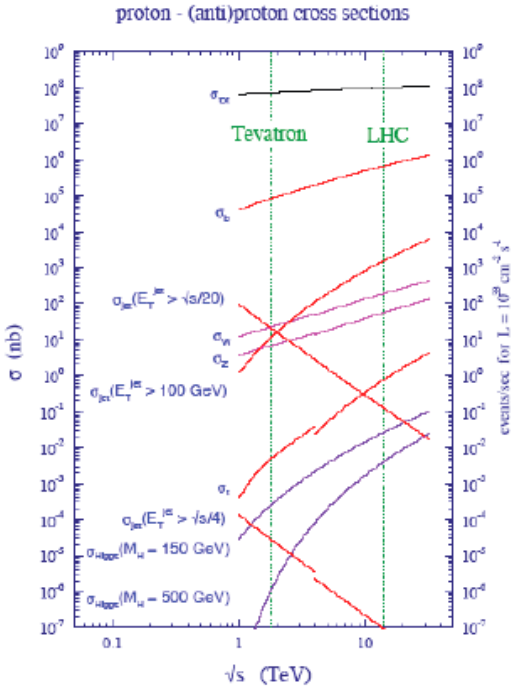
Q.-H. Cao, H.-L. Lai, J. Pumplin,
D. Stump, W.-K. Tung, and C.-P. Yuan

November 12, 2008

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
  - and in particular (for these lectures at least) that pdf's and pdf uncertainties are understood properly

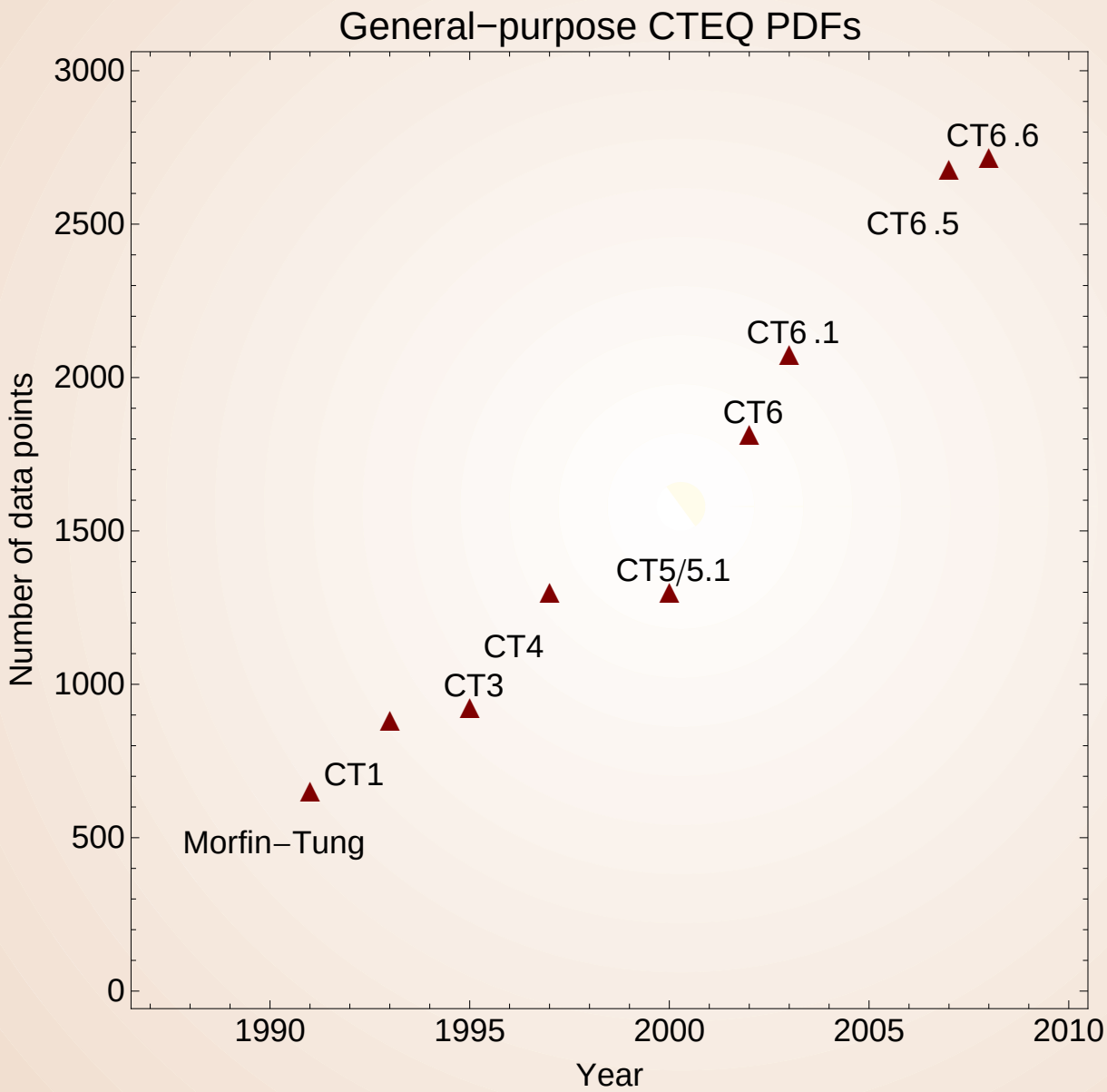




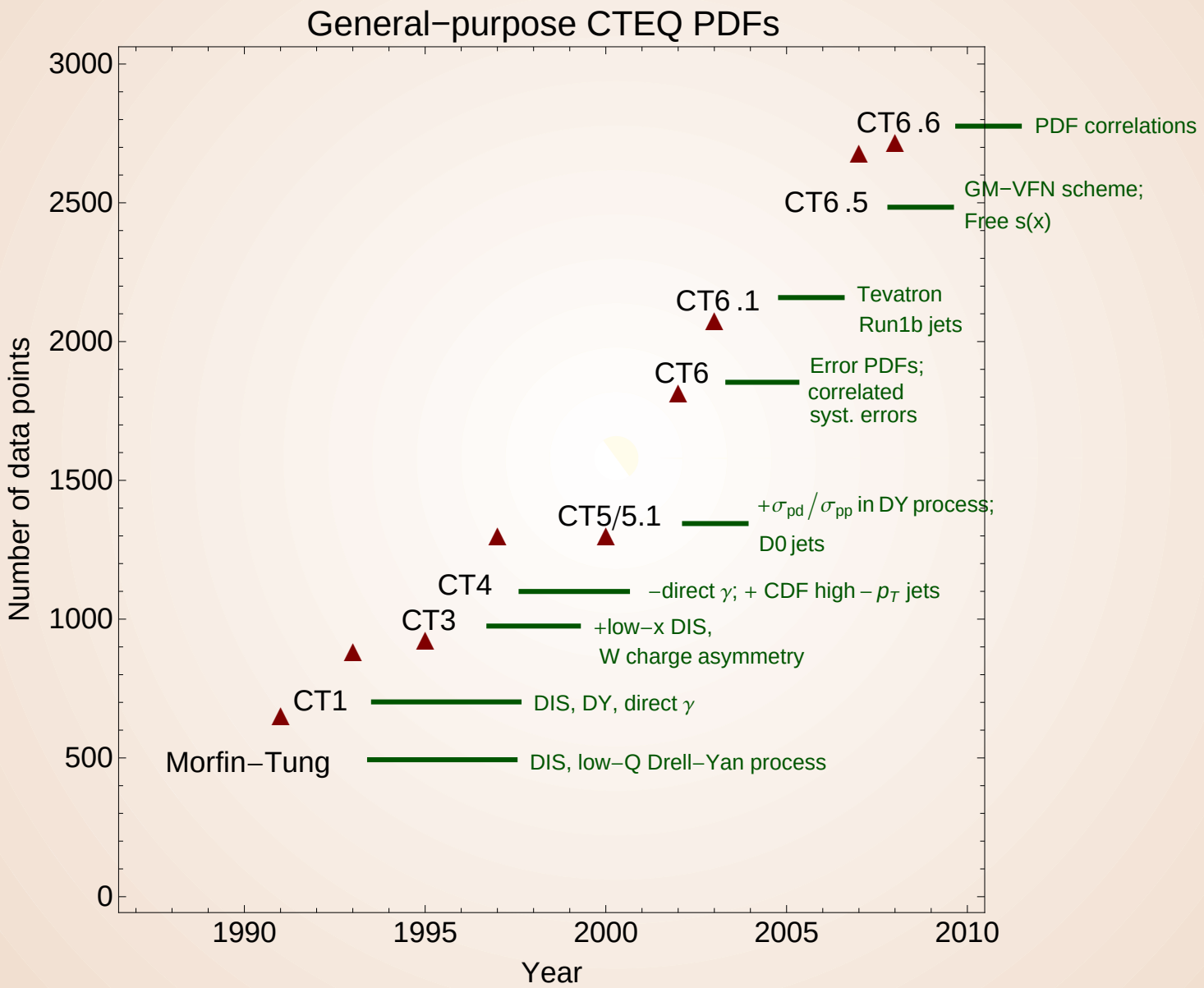
Global analysis at MSU-SMU-Taiwan-Washington

- a part of the Coordinated Theoretical Experimental study of QCD (CTEQ) in U.S.A.
- development of general-purpose parton distribution functions (PDFs) *(Wu-Ki Tung and collaborators)*
- analysis of implications for high-energy processes
- new CTEQ6.6M standard set and 44 extreme eigenvector sets *(Phys Rev, D78, 013004 (2008))*
  - available in the LHAPDF-5.4 library and at www.cteq.org

Timeline of CTEQ PDF's



Timeline of CTEQ PDF's



Data sets used in global fits (CTEQ6.6)

- BCDMS F_2^{proton} (339 data points)
- BCDMS $F_2^{deuteron}$ (251 data points)
- NMC F_2 (201 data points)
- NMC F_2^d/F_2^p (123 data points)
- F_2(CDHSW) (85 data points)
- F_3(CDHSW) (96 data points)
- CCFR F_2 (69 data points)
- CCFR F_3 (86 data points)
- H1 NC e^-p (126 data points; 1998-98 reduced cross section)
- H1 NC e^-p (13 data points; high y analysis)
- H1 NC e^+p (115 data points; reduced cross section 1996-97)
- H1 NC e^+p (147 data points; reduced cross section; 1999-00)
- ZEUS NC e^-p (92 data points; 1998-99)
- ZEUS NC e^+p (227 data points; 1996-97)
- ZEUS NC e^+p (90 data points; 1999-00)
- H1 F_2^c e^+p (8 data points;1996-97)
- H1 $R\sigma^c$ for ccbar e^+p (10 data points;1996-97)
- H1 R_σ^b for bbbar e^+p (10 data points; 1999-00)
- ZEUS F_2^c e^+p (18 data points; 1996/97)
- ZEUS F_2^C e+p (27 data points; 1998/00)
- H1 CC e^-p (28 data points; 1998-99)
- H1 CC e^+p (25 data points; 1994-97)
- H1 CC e^+p (28 data points; 1999-00)
- ZEUS CC e^-p (26 data points; 1998-99)
- ZEUS CC e^+p (29 data points; 1994-97)
- ZEUS CC e^+p (30 data points; 1999-00)
- NuTev neutrino dimuon cross section (38 data points)
- NuTev anti-neutrino dimuon cross section (33 data points)
- CCFR neutrino dimuon cross section (40 data points)
- CCFR anti-neutrino cross section (38 data points)
- E605 dimuon (199 data points)
- E866 dimuon (13 data points)
- Lepton asymmetry from CDF (11 data points)
- CDF Run 1B jet cross section (33 data points)

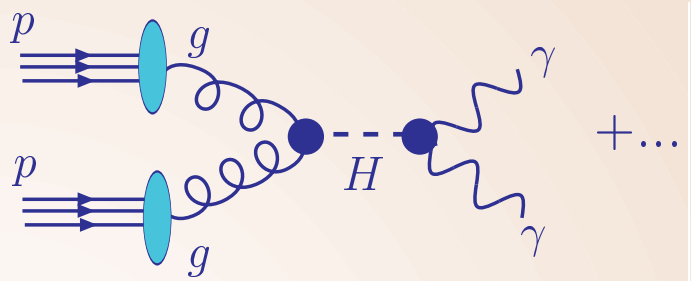
- 2794 data points from DIS, DY, jet production
- All with (correlated) systematic errors that must be treated correctly in the fit
- Note that DIS is the 800 pound gorilla of the global fit with many data points and small statistical and systematic errors
  - and fixed target DIS data still have a significant impact on the global fitting, even with an abundance of HERA data
- To avoid non-perturbative effects, kinematic cuts on placed on the DIS data
  - $Q^2>5$ GeV^2
  - $W^2(=m^2+Q^2(1-x)/x)>12.25$ GeV^2

Introduction to the global PDF analysis

A typical perturbative QCD calculation

Higgs boson production $pp \to (H \to \gamma\gamma)X$ via $gg \to H$

Total cross section $\sigma_{pp\to H\to\gamma\gamma}$:



$$\sigma_{pp\to H\to\gamma\gamma} = \widehat{\sigma}_{gg\to H\to\gamma\gamma} f_{g/p}(x_1, M_H) f_{g/p}(x_2, M_H) + \ldots$$

- $\widehat{\sigma}_{gg\to H\to\gamma\gamma}$ is the **hard-scattering cross section**, given by a **perturbation series in α_s (at least formally)**
- $f_{g/p}(x,\mu)$ is the parton distribution function for finding a gluon g with momentum $x\vec{P}$ in a proton with momentum $\vec{P}$ at a typical momentum μ; $\mu \sim \left|\vec{P}\right| \gg 1$ GeV

$f_{g/p}(x,\mu)$ are **universal** (process-independent) **nonperturbative** functions

Perturbative evolution of $f_{i/p}(x,\mu)$; global fits

At some $\mu_0 \sim 1$ GeV, the PDF's are **parametrized** as

$$f_{i/p}(x,\mu_0) = a_0 x^{a_1}(1-x)^{a_2} F(a_3, a_4, ...)$$

At $\mu > \mu_0$, $f_{i/p}(x,\mu)$ are determined by solving **Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations**,

$$\mu \frac{df_{i/p}(x,\mu)}{d\mu} = \sum_{j=g,u,\bar{u},d,\bar{d},....} \int_x^1 \frac{dy}{y} P_{i/j}\left(\frac{x}{y}, \alpha_s(\mu)\right) f_{j/p}(y,\mu),$$

with $P_{i/j}$ known to order α_s^3 (NNLO):

$$P_{i/j}(x,\alpha_s) = \alpha_s P_{i/j}^{(1)}(x) + \alpha_s^2 P_{i/j}^{(2)}(x) + \alpha_s^3 P_{i/j}^{(3)}(x) + ...$$

Free parameters a_i **and their uncertainties** are determined from a global fit to hadron scattering data

Back to global fits

- With the DGLAP equations, we know how to evolve pdf's from a starting scale Q_0 to any higher scale
- …but we can't calculate what the pdf's are ab initio
  - one of the goals of lattice QCD
- We have to determine them from a global fit to data
  - factorization theorem tells us that pdf's determined for one process are applicable to another

- So what do we need
  - a value of Q_o (1.3 GeV for CTEQ, 1 GeV for MSTW) lower than the data used in the fit (or any prediction)
  - a parametrization for the pdf's
  - a scheme for the pdf's
  - hard-scattering calculations at the order being considered in the fit
  - pdf evolution at the order being considered in the fit
  - a world average value for α_s
  - a lot of data
    - with appropriate kinematic cuts
  - a treatment of the errors for the experimental data
  - MINUIT

Back to global fits

- Parametrization: initial form
  - $f(x) \sim x^{\alpha}(1-x)^{\beta}$
  - estimate β from quark counting rules
    - $\beta = 2n_s - 1$ with n_s being the minimum number of spectator quarks
    - so for valence quarks in a proton (qqq), $n_s=2$, $\beta=3$
    - for gluon in a proton (qqqg), $n_s=3$, $\beta=5$
    - for anti-quarks in a proton (qqqqqbar), $n_s=4$, $\beta=7$
  - estimate α from Regge arguments
    - gluons and anti-quarks have $\alpha \sim -1$ while valence quarks have $\alpha \sim = 1/2$
  - but at what Q value are these arguments valid?

- What do we know?

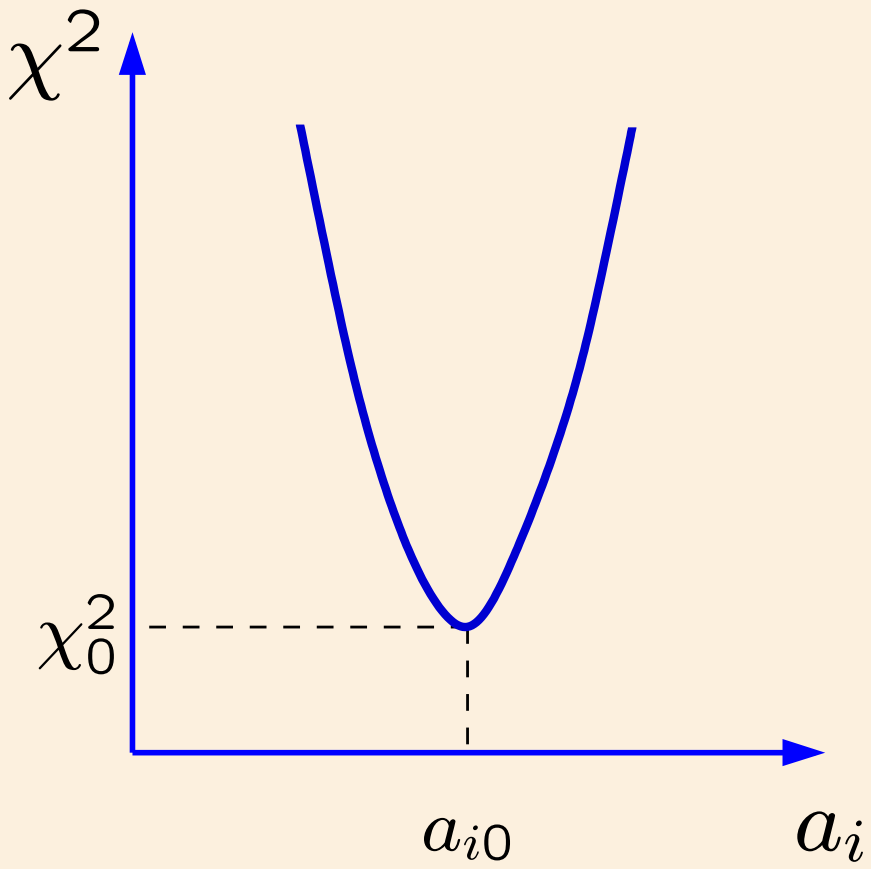
  we know that the sum of the momentum of all partons in the proton is 1 (but see later for modified LO fits)

  we know the sum of valence quarks is 3
  - and 2 of them are up quarks and 1 of them is a down quark
  - we know that the net number of anti-quarks is 0, but what about dbar=ubar

  we know that the net number of strange quarks (charm quarks/bottom quarks) in the proton is 0
  - but we don't know if s=sbar locally

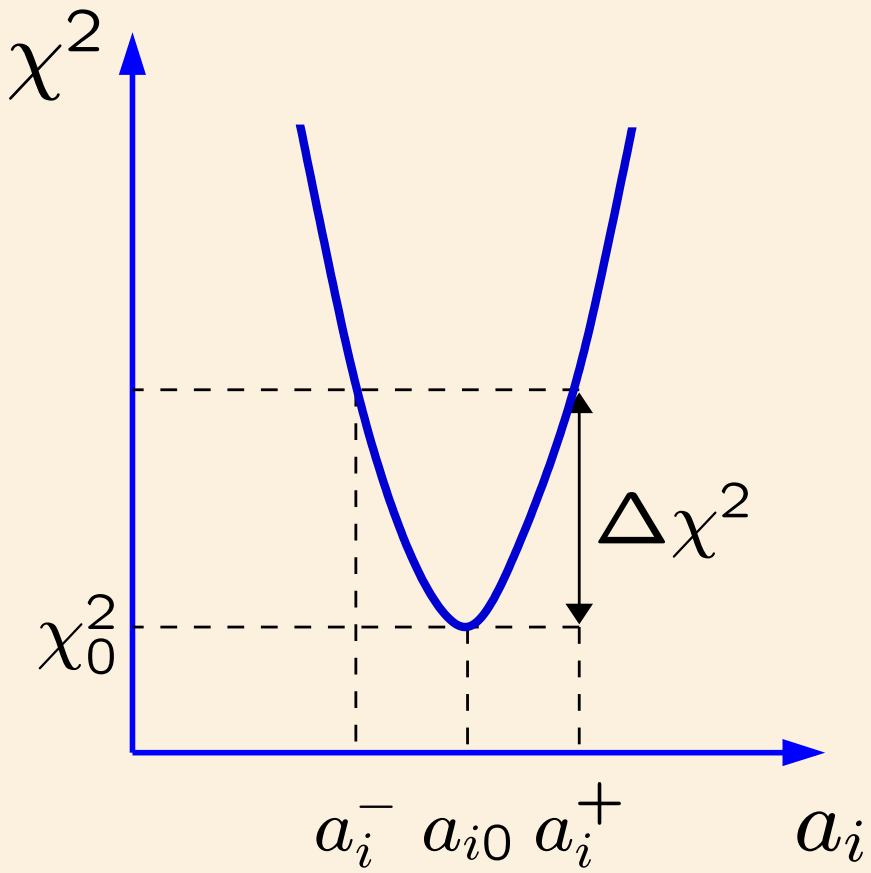
This already puts a lot of restrictions on the pdf's

Multi-dimensional PDF error analysis



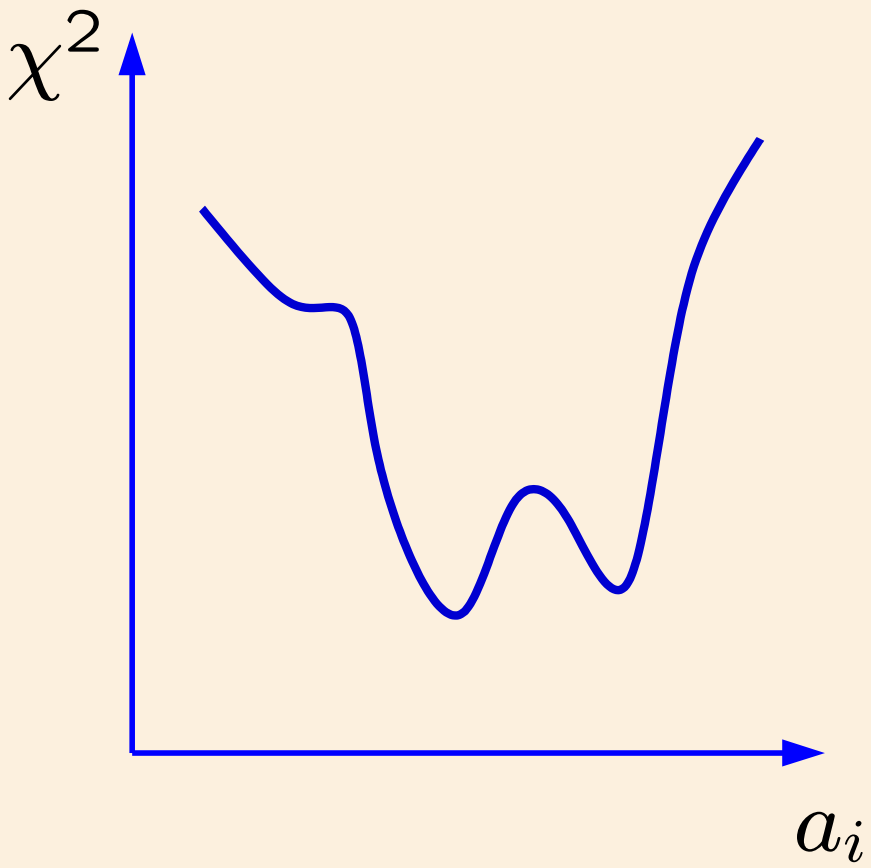
- Minimization of a likelihood function (χ^2) with respect to ~ 30 theoretical (mostly PDF) parameters $\{a_i\}$ and > 100 experimental systematical parameters
  - partly analytical (offset method) and partly numerical

Multi-dimensional PDF error analysis



- Establish a confidence region for $\{a_i\}$ for a given tolerated increase in χ^2

Multi-dimensional PDF error analysis

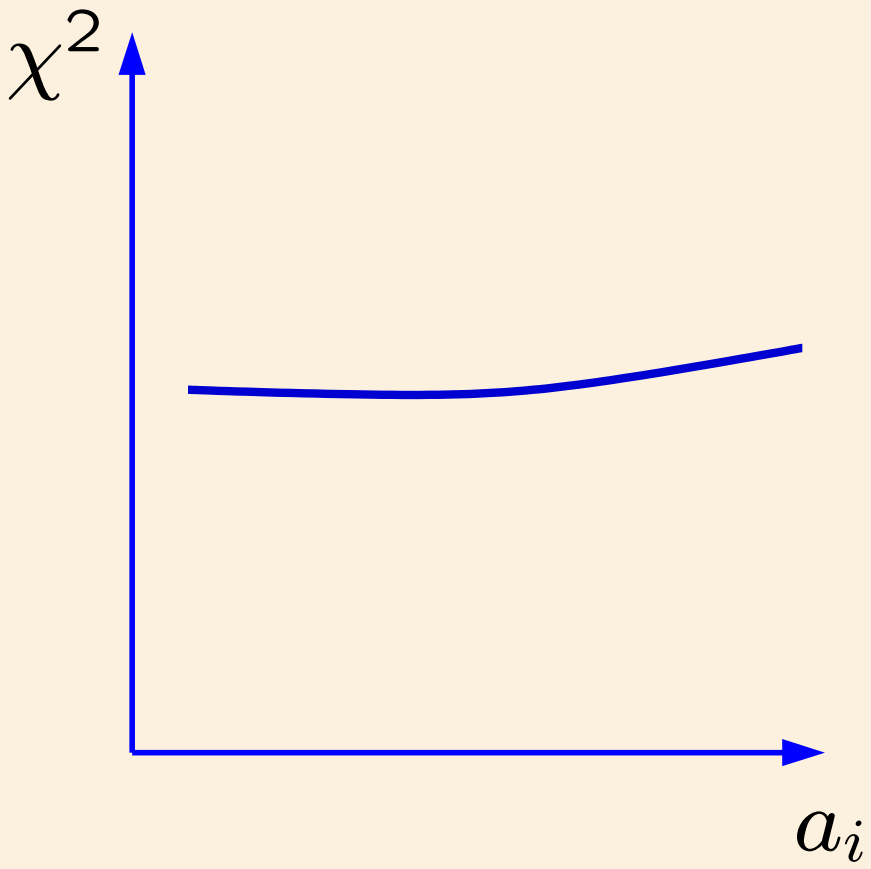


Pitfalls to avoid

- "Landscape"
  - disagreements between the experiments

In the worst situation, significant disagreements between M experimental data sets can produce up to $N \sim M!$ possible solutions for PDF's, with $N \sim 10^{500}$ reached for "only" about 200 data sets

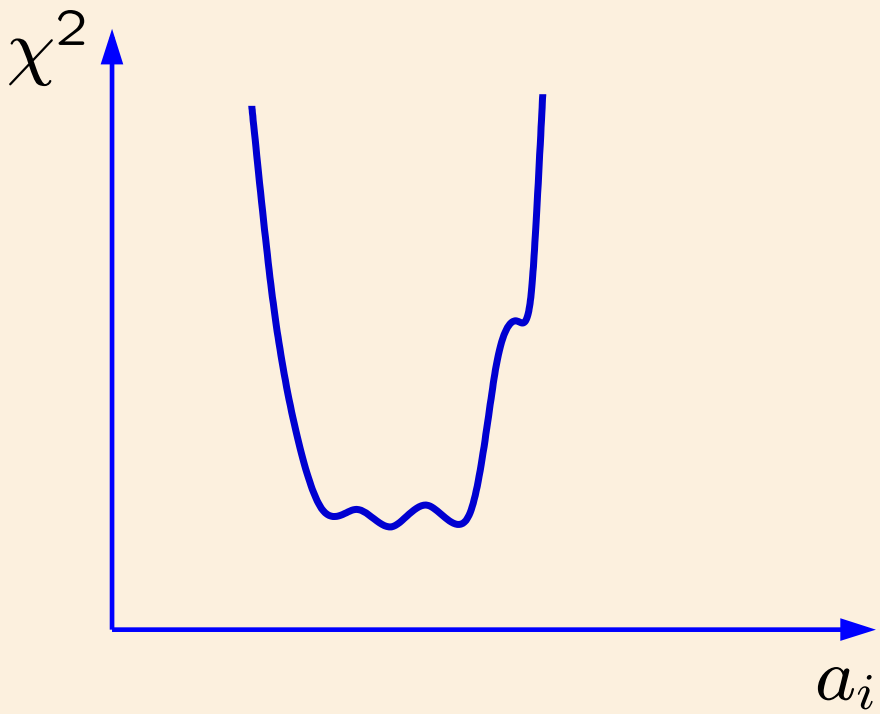
Multi-dimensional PDF error analysis



Pitfalls to avoid

- Flat directions
  - unconstrained combinations of PDF parameters
  - dependence on free theoretical parameters, especially in the PDF parametrization

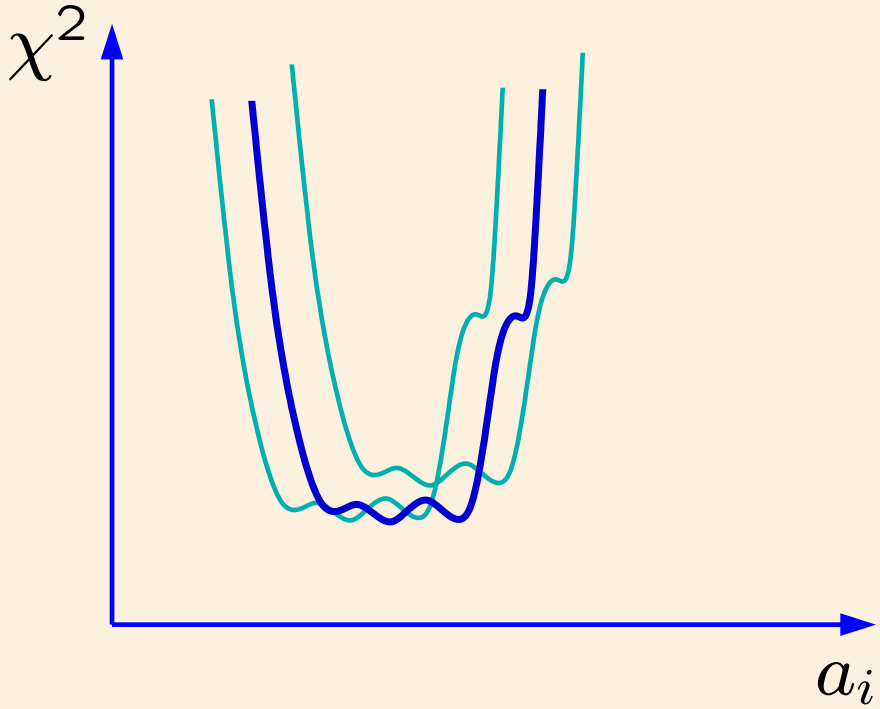
Multi-dimensional PDF error analysis



The actual χ^2 function shows

- a well pronounced global minimum χ^2_0
- weak tensions between data sets in the vicinity of χ^2_0 (mini-landscape)
- some dependence on assumptions about flat directions

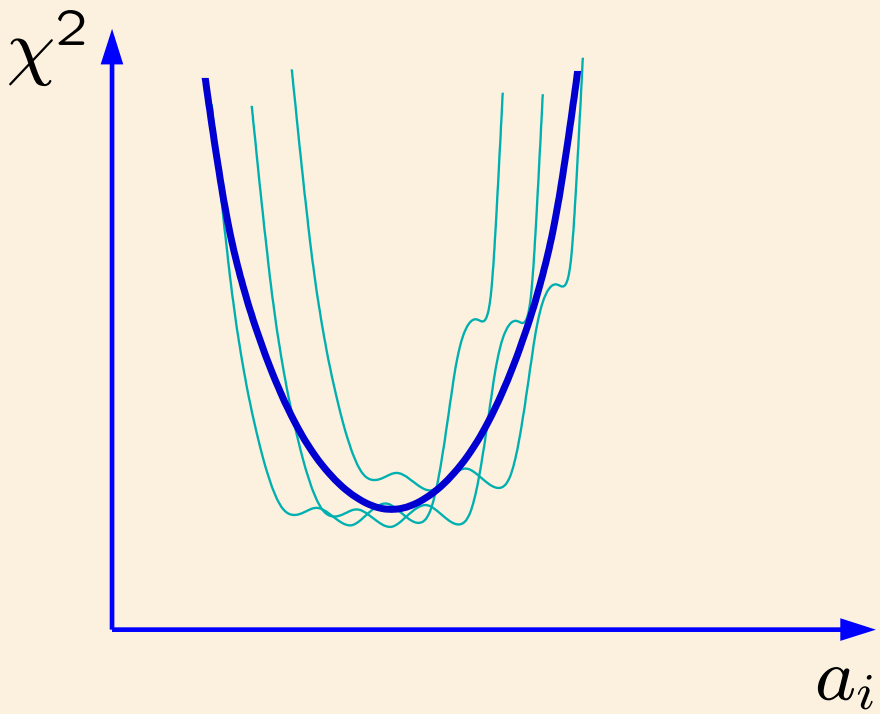
Multi-dimensional PDF error analysis



The actual χ^2 function shows

- a well pronounced global minimum χ^2_0
- weak tensions between data sets in the vicinity of χ^2_0 (mini-landscape)
- some dependence on assumptions about flat directions

Multi-dimensional PDF error analysis

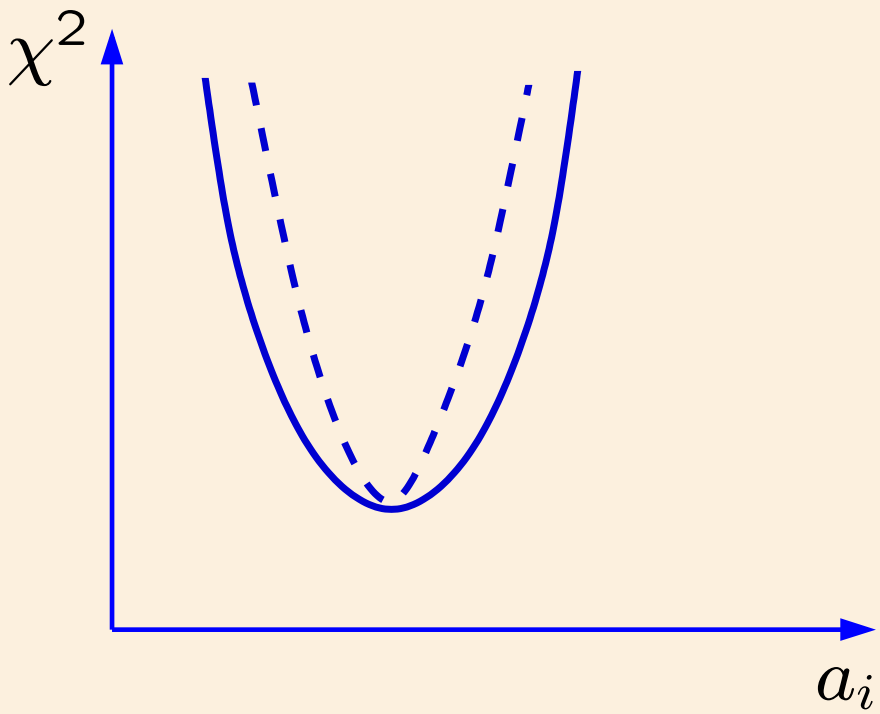


The actual χ^2 function shows

- a well pronounced global minimum χ^2_0
- weak tensions between data sets in the vicinity of χ^2_0 (mini-landscape)
- some dependence on assumptions about flat directions

The likelihood is approximately described by a quadratic χ^2 with a revised tolerance condition $\Delta\chi^2 \leq T^2$

Multi-dimensional PDF error analysis

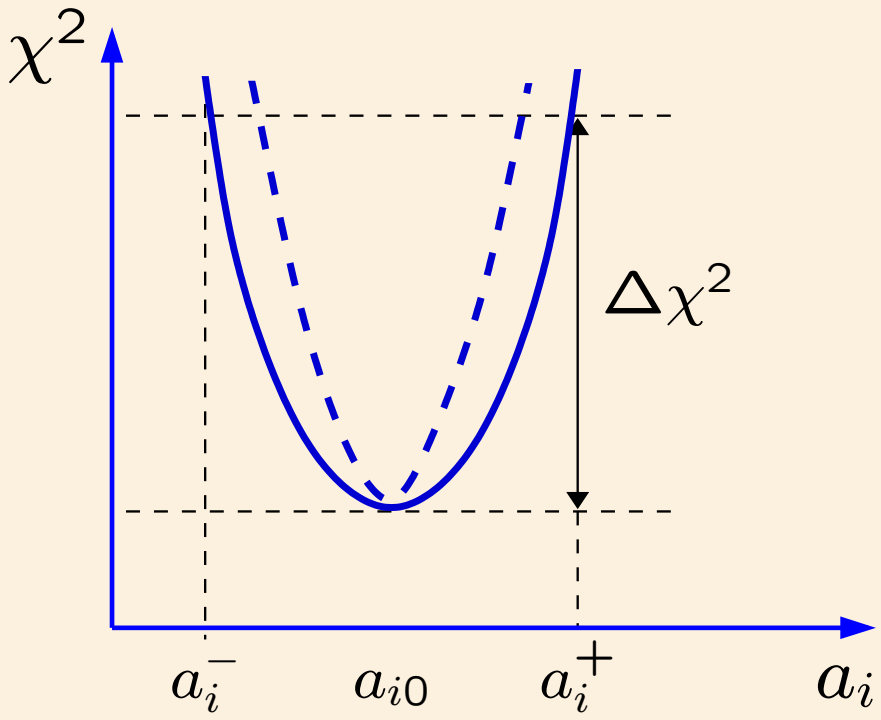


The actual χ^2 function shows

- a well pronounced global minimum χ^2_0
- weak tensions between data sets in the vicinity of χ^2_0 (mini-landscape)
- some dependence on assumptions about flat directions

The likelihood is approximately described by a quadratic χ^2 with a revised tolerance condition $\Delta\chi^2 \leq T^2$

Multi-dimensional PDF error analysis



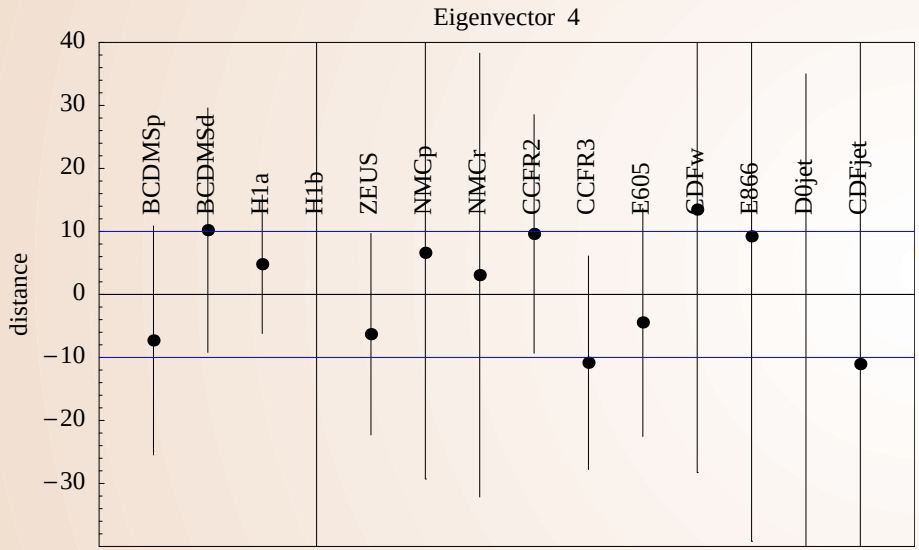
The actual χ^2 function shows

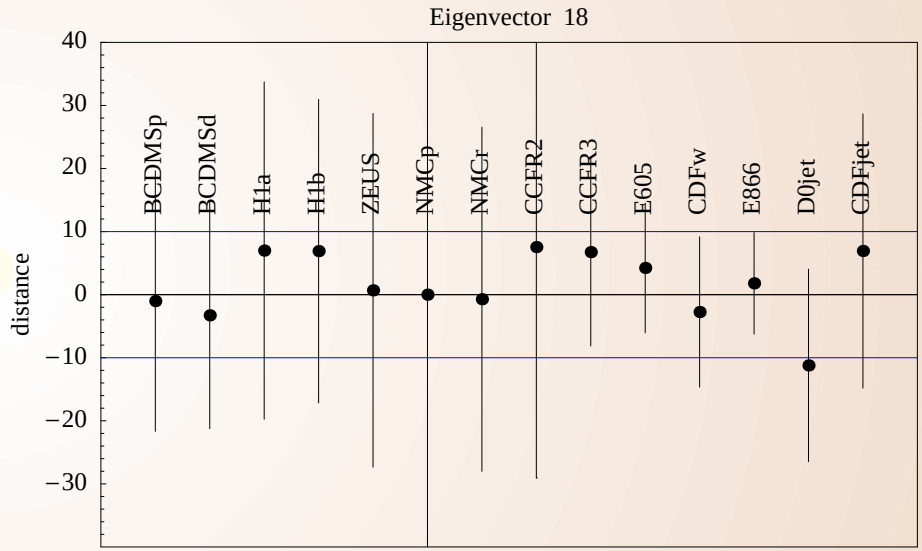
- a well pronounced global minimum χ^2_0
- weak tensions between data sets in the vicinity of χ^2_0 (mini-landscape)
- some dependence on assumptions about flat directions

The likelihood is approximately described by a quadratic χ^2 with a revised tolerance condition $\Delta\chi^2 \leq T^2$

CTEQ6 tolerance criterion (2001)

Acceptable values of PDF parameters must agree at ≈90% c.l. with all experiments included in the fit, *for a plausible range of assumptions about the PDF parametrization, scale dependence, experimental systematics, ...*





Can be crudely approximated (but does not have to) by assuming $T \approx 10$ for all PDF parameters

A somewhat stricter variant of this criterion is applied in the MSTW'08 analysis

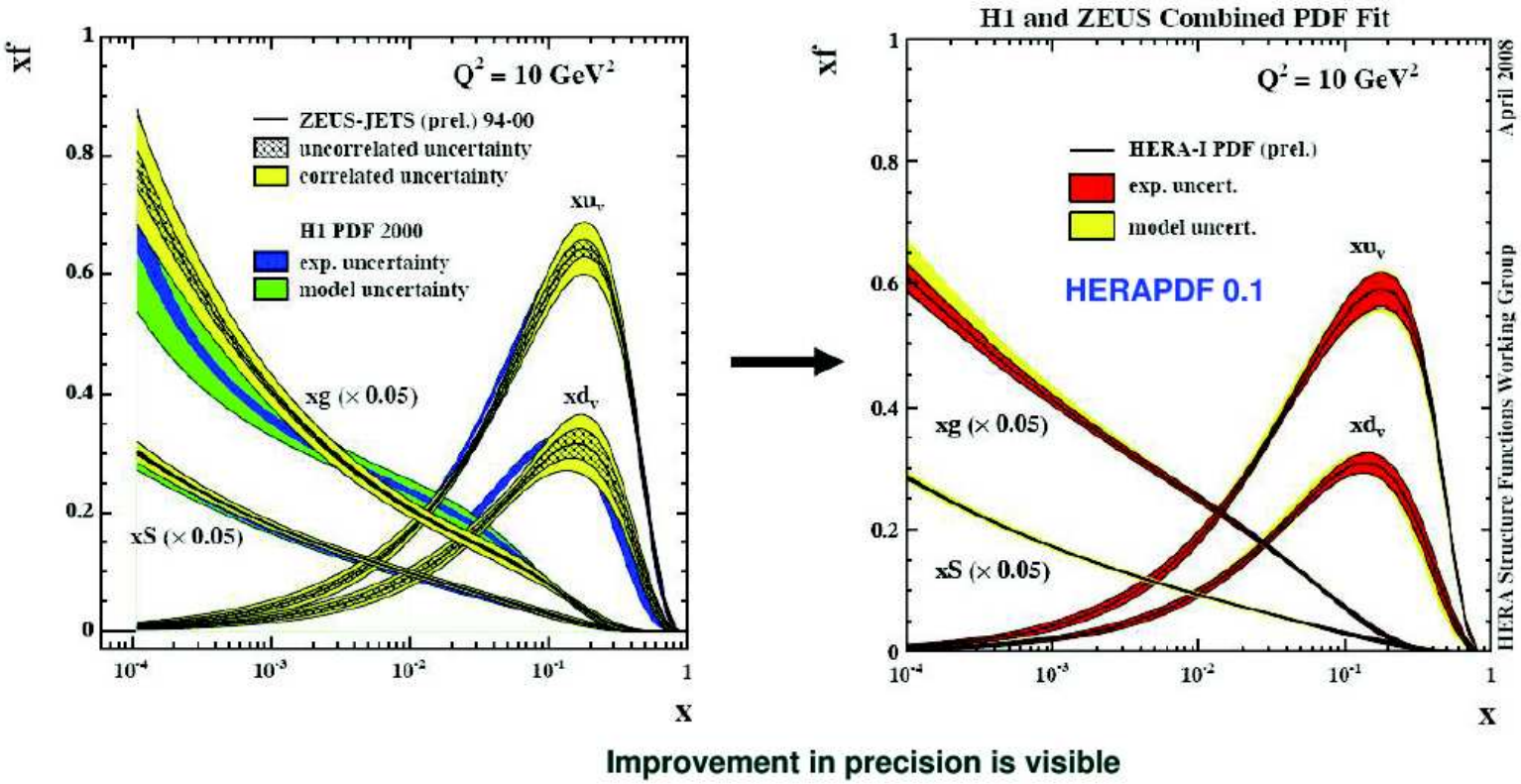
HERAPDF0.1 set based on the combined neutral-current DIS data (2008)

The common fit of the combined data

Partons parametrized at $Q_0^2 = 4$ GeV^2 (Data Q^2>3.5 GeV^2)

Experimental+Model uncertainties taken into account

Errors of the fit estimated using $\Delta\chi^2$=1

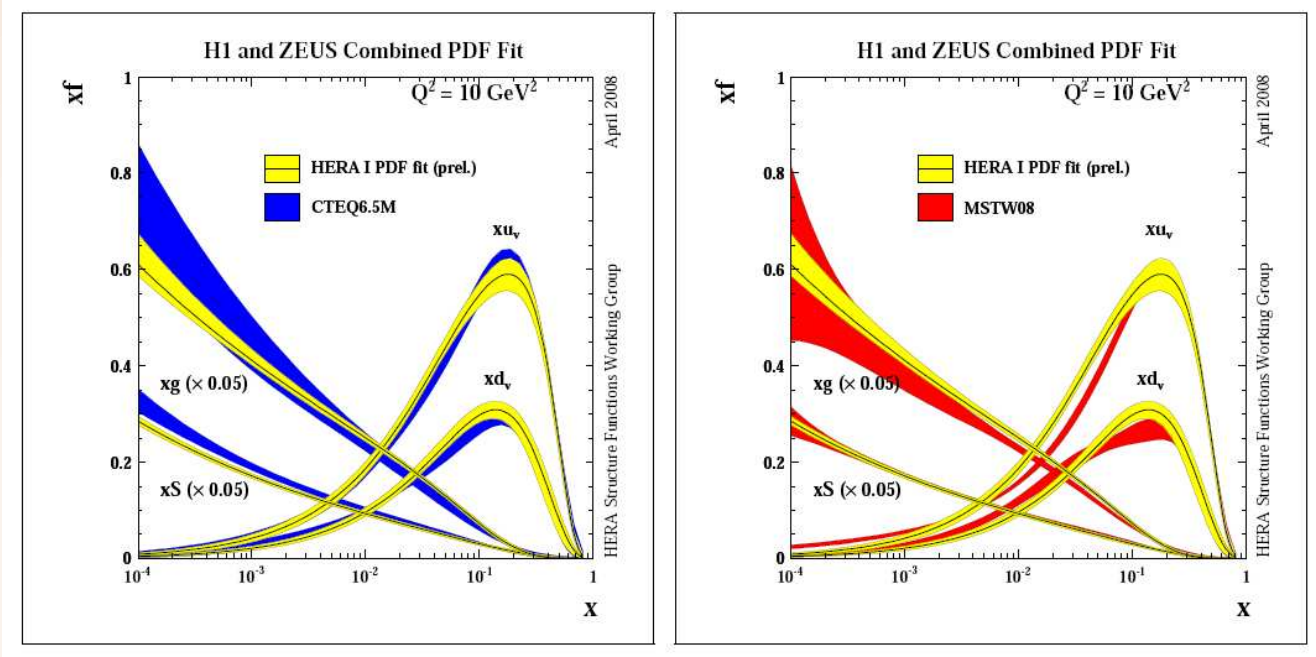


Improvement in precision is visible

Christinel Diaconu. HERA-LHC workshop, 2008

HERAPDF0.1 set based on the combined H1+ZEUS data

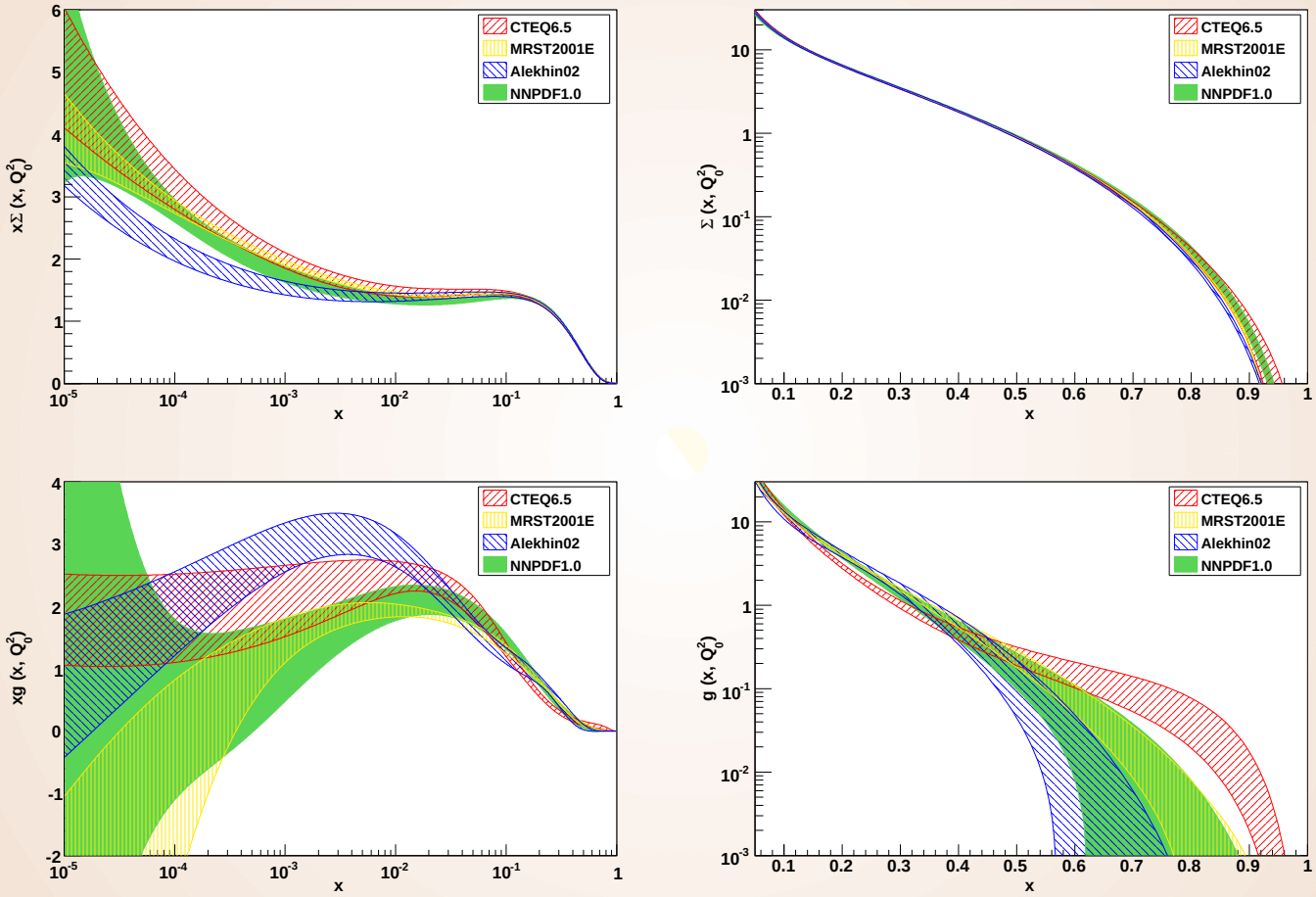
Smaller nominal PDF uncertainty as compared to CTEQ/MSTW; but



- insufficient PDF flavor separation (neutral-current DIS probes only $4/9\,(u+\bar{u}+c+\bar{c})+1/9\,(d+\bar{d}+s+\bar{s})$)
- more rigid PDF parametrizations (e.g., $g(x,Q_0)=Ax^B(1-x)^C$) $\Rightarrow$ less flexibility to probe the PDF behavior, notably at small x
- so far, a simplified (zero-mass scheme) treatment of charm and bottom mass dependence

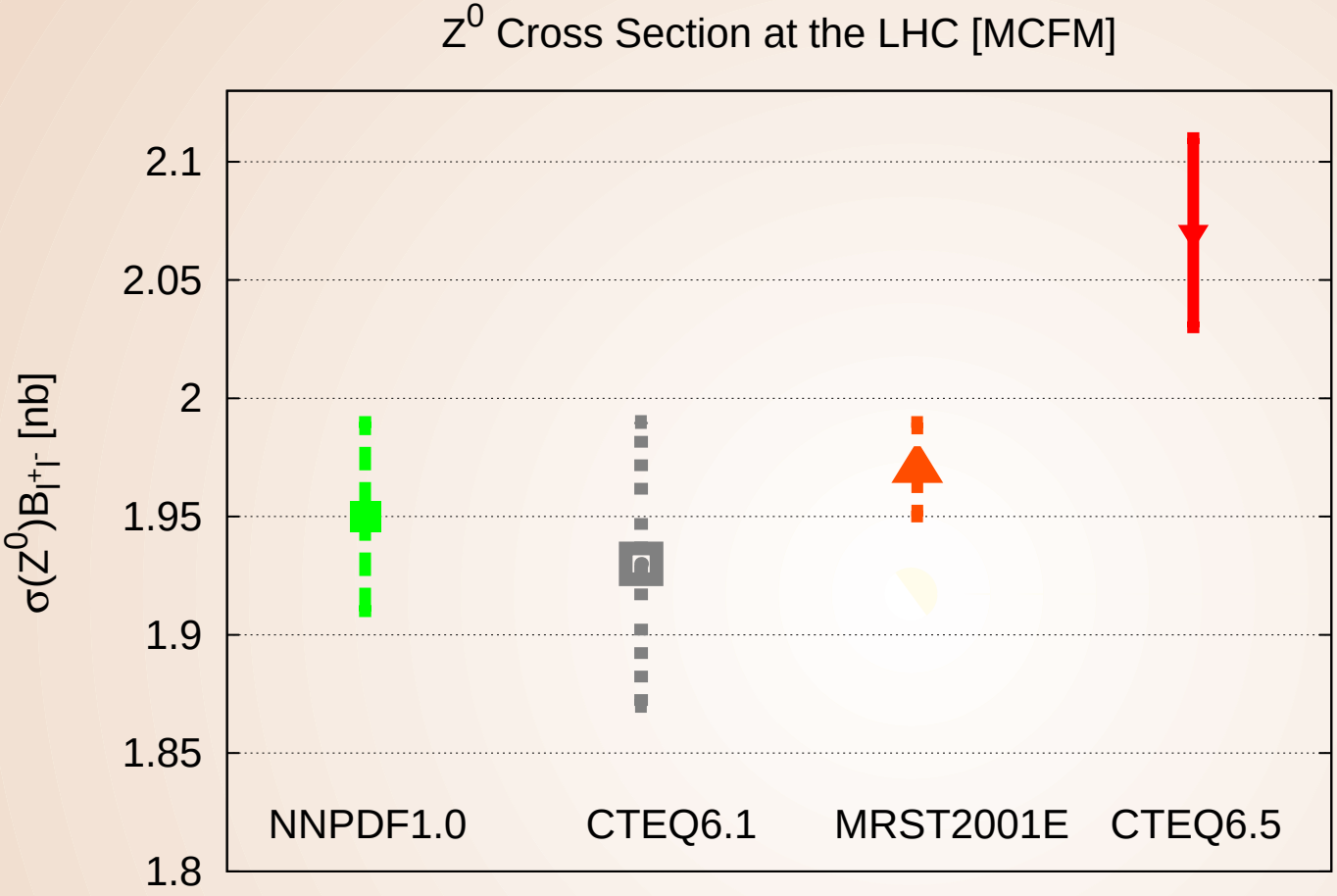
Studies of tolerance in new approaches

Neural Network PDF collaboration (arXiv:0808.1231)



Biases due to the parametrization, Gaussian approximation for errors are reduced; better estimates for the PDF uncertainty may be feasible

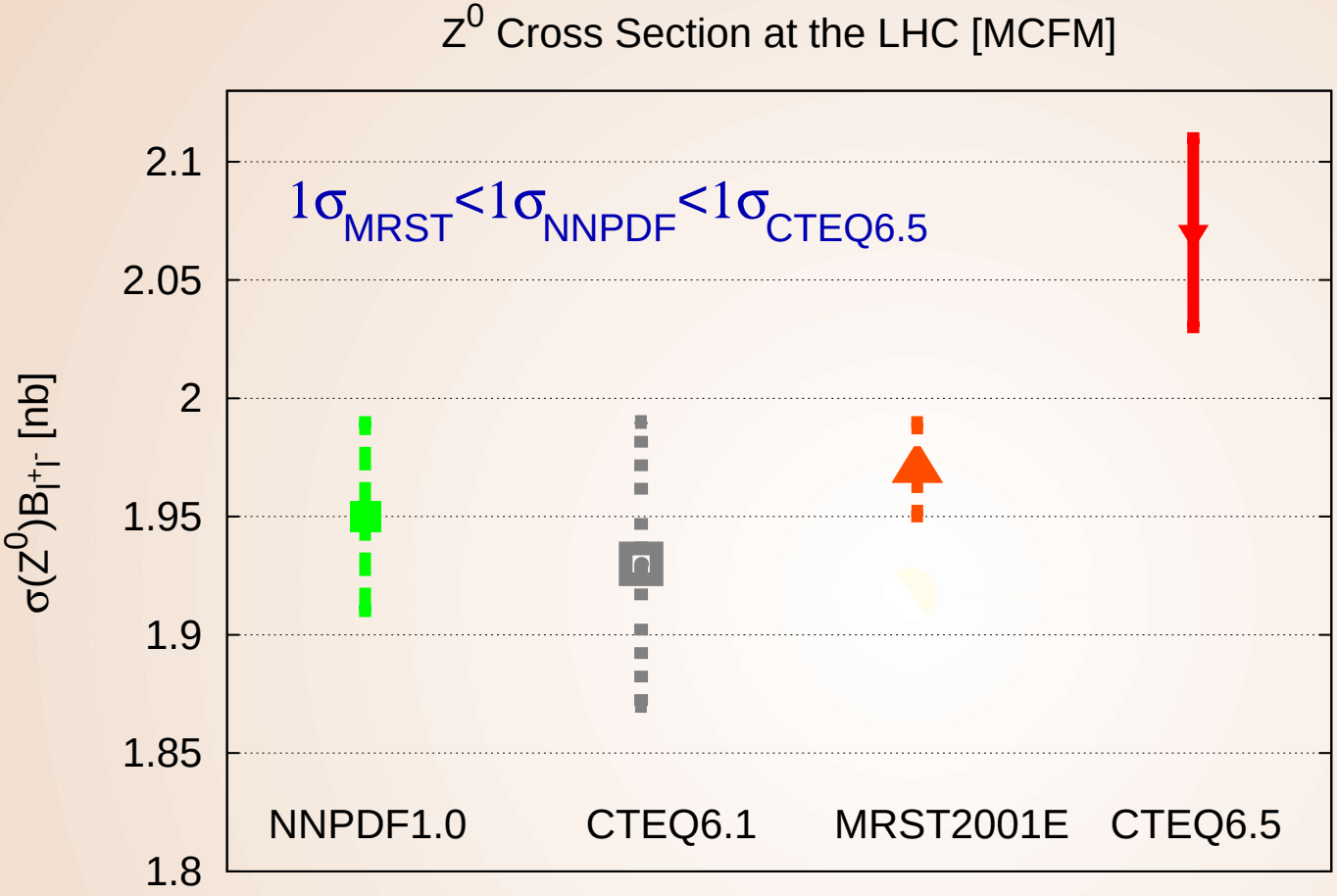
Z boson cross section at the LHC



Year:	2008	2003	2001	2007
Npts:	3161 (DIS)	1120		2714

NNPDF Coll.. arXiv:0808.1231

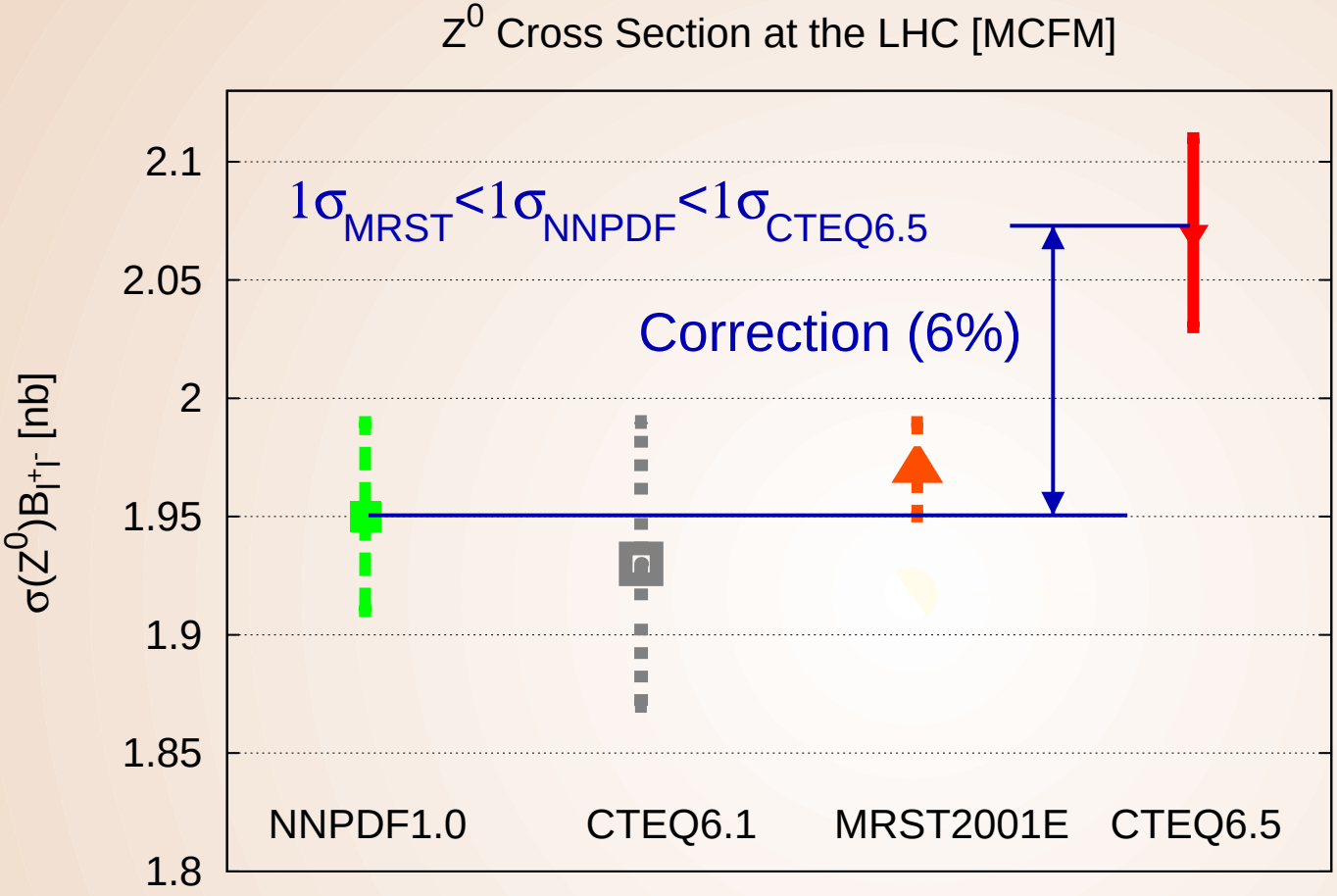
Z boson cross section at the LHC



Year:	2008	2003	2001	2007
Npts:	3161 (DIS)	1120		2714

NNPDF Coll.. arXiv:0808.1231

Z boson cross section at the LHC



Year:	2008	2003	2001	2007
Npts:	3161 (DIS)	1120		2714

NNPDF Coll.. arXiv:0808.1231

Z and W production as "standard candle" processes

Event rates for $pp \to W^{\pm}X$, $pp \to Z^0 X$

- will be measured with accuracy $\delta\sigma/\sigma \sim 1\%$
- will be employed to constrain PDF's and monitor the LHC luminosity $\mathcal{L}$ in real time

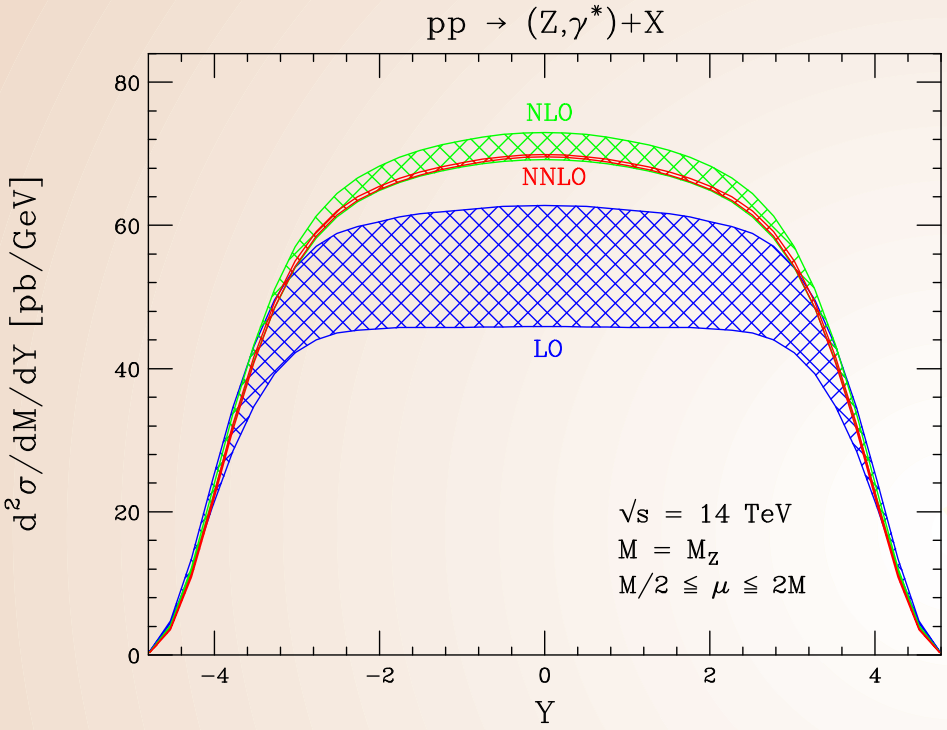
  (Dittmar, Pauss, Zurcher; Khoze, Martin, Orava, Ryskin; Giele, Keller';...)

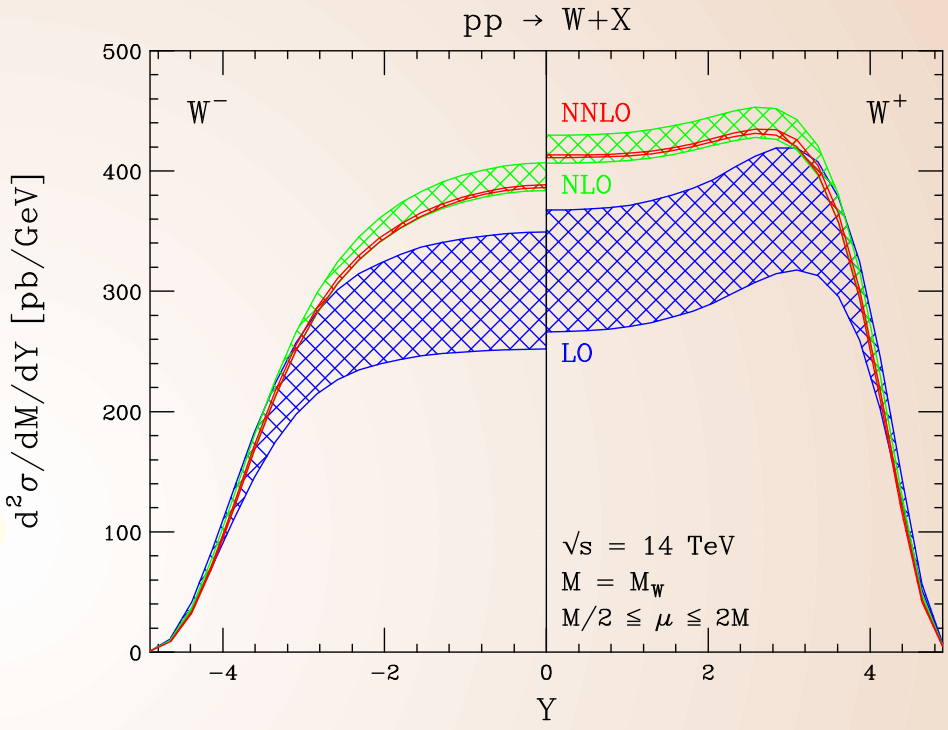
  - other methods will initially give $\delta\mathcal{L} = 10 - 20\%$

Various cross section measurements will be normalized to Z, W cross sections

W and Z rapidity distributions at NNLO

(Anastasiou, Dixon, Melnikov, Petriello, 2003)





- Tiny scale dependence ($\sim 1\%$)
- For $|y| < 2$, NNLO leads to a uniform rescaling

$$\sigma_{NNLO} \approx K_{NNLO} \cdot \sigma_{NLO};\ K^{LHC}_{NNLO} \approx 0.98$$

- Larger corrections at forward rapidities

1. What causes the 6% difference between CTEQ6.5/CTEQ6.6 and CTEQ6.1 Z cross sections at the LHC?

   Answer: improved treatment of c, b scattering in DIS at HERA

2. Which aspects of PDF parametrizations dominate the PDF uncertainty? Which measurements can reduce the PDF uncertainty?

Treatment of heavy quarks in PDF analysis

s, c, b: the least constrained sector of the nucleon structure

- Data from HERA, NuTeV, Tevatron is increasingly sensitive to heavier flavors
- Some theoretical constraints on $Q_{\pm}(x,\mu) = Q(x,\mu) \pm \bar{Q}(x,\mu)$ ($Q = s,\, c,\, b$) can be now released, and new PDF models can be examined
- opportunities for interesting QCD tests
  - QCD factorization with realistic $m_{c,b}$ dependence
  - studies of flavor asymmetries in the quark sea
- impact on BSM searches, general hadronic physics at the LHC, etc.

CTEQ6.5 and CTEQ6.6: advanced treatment of heavy quarks

1. full implementation of the **general-mass "SACOT-χ" scheme**

   - differences in predictions for c, b scattering ($F_2^{c,b}(x,Q^2)$, etc.), EW precision cross sections, as compared to the zero-mass CTEQ6.1

   Tung et al., JHEP 0702, 053 (2007); ***CTEQ6.5, CTEQ6.6***

2. exploration of **free strange PDF's** and/or asymmetric strange sea

$$s_+(x) \neq r\left(\bar{u}(x)+\bar{d}(x)\right), \quad s_-(x) \neq 0,$$
$$\text{where } s_\pm(x) \equiv s(x) \pm \bar{s}(x)$$

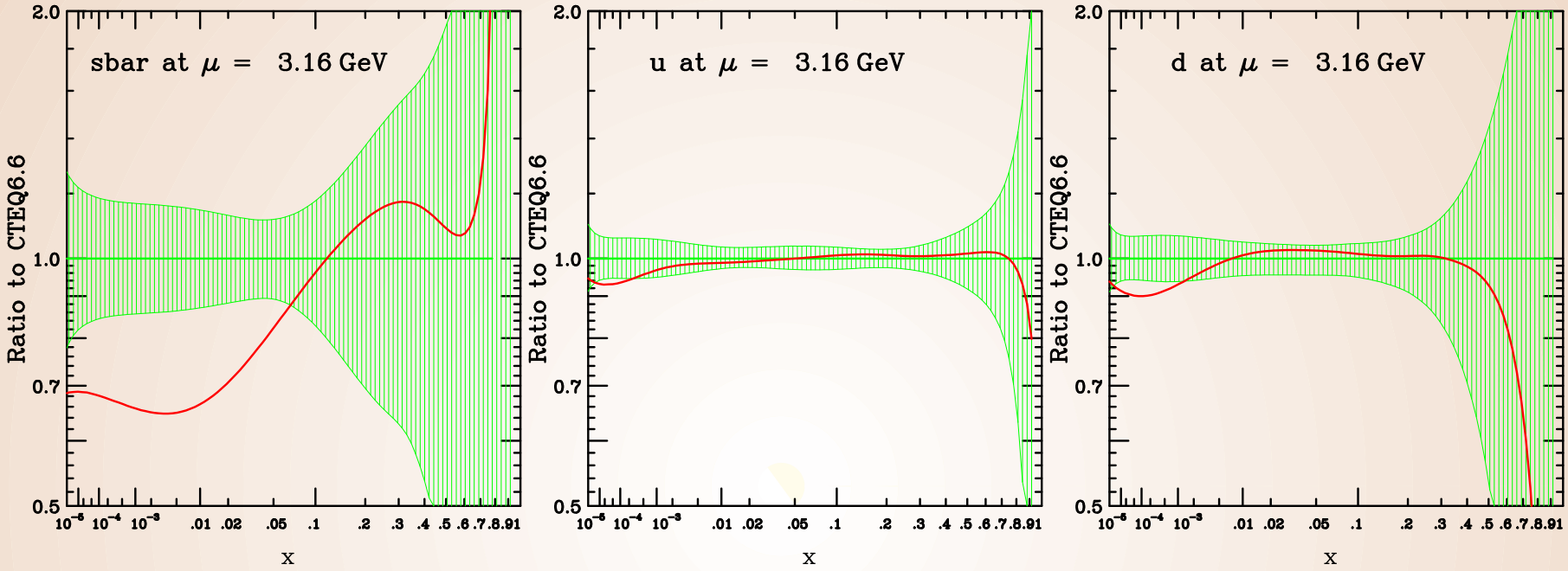
   Lai et al., JHEP 0704, 089 (2007); ***CTEQ6.5S, CTEQ6.6***

3. PDF's with **nonperturbative charm**

   - $c(x,\mu_0 = m_c) \neq 0$ due to low-energy charm excitations (as opposed to $g \to c\bar{c}$ radiative production)

   Pumplin et al., PRD 75, 054029 (2007); ***CTEQ6.5C, 6.6C***

CTEQ6.6 PDF's



dashes: CTEQ6.1M (zero-mass scheme)

- very different strange PDF's
- CTEQ6.6 u, d are above CTEQ6.1 at $x \lesssim 10^{-2}$
  - The result of suppressed charm contribution to $F_2(x, Q)$ at HERA in the GM-VFN scheme

Zero mass variable flavor number scheme (ZM-VFNS)

- Start pdf evolution at charm threshold ($Q=m_c=1.3$ GeV)
  - set c and b distributions to zero at this scale (although can allow for possibility of intrinsic charm/bottom)
    - start b evolution at $Q=m_b$
  - all heavy quarks treated as massless
  - c and b pairs created by gluon splitting
  - adjust running coupling α_s as each flavor threshold is crossed since QCD β function depends on # of active flavors
  - in this approach, only mass effects are due to flavor thresholds and changing of β function
- Most commonly used CTEQ NLO pdf's prior to CTEQ6.5 (such as CTEQ6M, CTEQ6.1) are of this type
- Advantages
  - easy to implement
  - sums large logs of Q^2/m_Q^2 via DLGAP equation
  - asymptotically correct when $Q^2 >> m_Q^2$
- Disadvantages
  - does not treat heavy quark threshold correctly

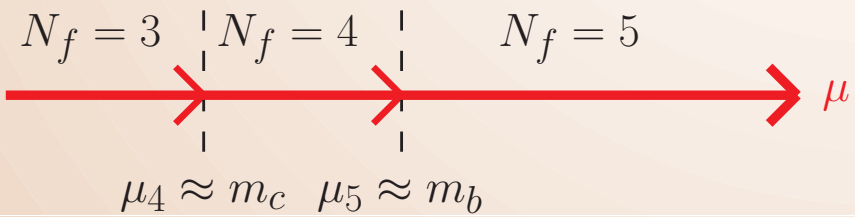
Fixed flavor number scheme

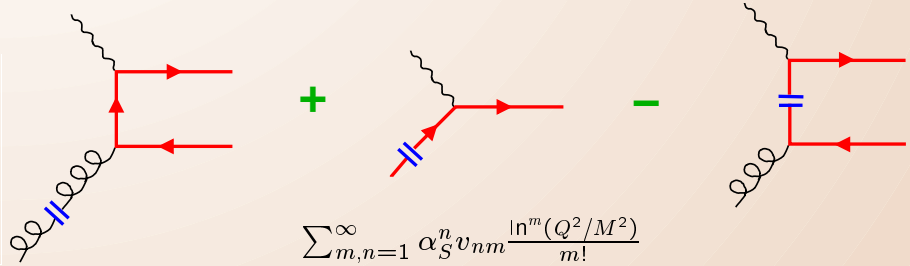
- Calculate heavy quark production from relevant subprocesses such as $\gamma^*g \to Q\bar{Q}$ keeping only light quarks in DGLAP equations
- Only light quarks have pdf's
  - no charm or bottom quark pdf's
- Advantage
  - gets threshold behavior correct
- Disadvantage
  - does not resum potentially large logs of Q^2/m_Q^2

General-mass variable-flavor number scheme

(Aivasis et al.; Chuvakin et al.; Thorne, Roberts; Kniehl et al.; Buza et al.; Cacciari et al.; ...)

- A series of effective fixed-flavor number (FFN) schemes, with N_f (the number of active parton flavors) incremented sequentially at momentum scales $\mu_{N_f} \approx m_{N_f}$
- incorporates essential $m_{c,b}$ dependence near, and away from, heavy-flavor thresholds





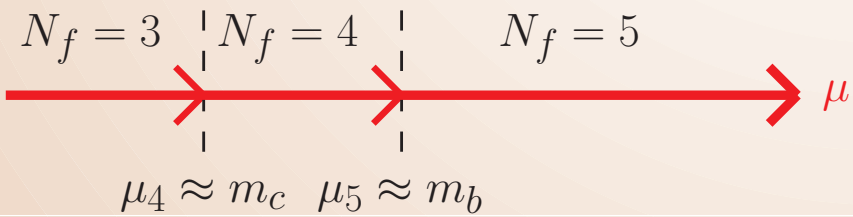
General-mass variable-flavor number scheme

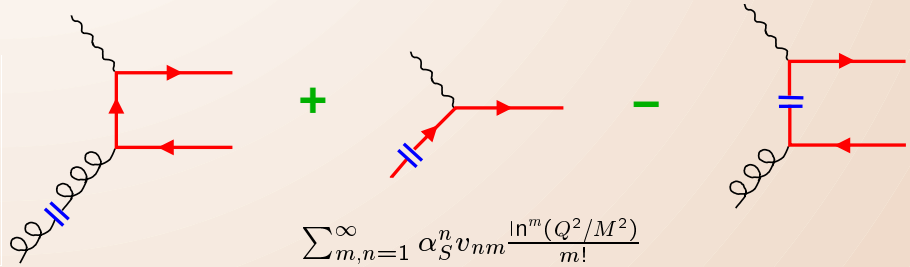
(Aivasis et al.; Chuvakin et al.; Thorne, Roberts; Kniehl et al.; Buza et al.; Cacciari et al.; ...)

Proved for *inclusive DIS* by J. Collins (1998)

$$F_2(x,Q,m_c) = \sum_a \int_\chi^1 \frac{d\xi}{\xi} H_a(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_c}{Q}) f_a(\xi, \frac{\mu}{m_c}) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{Q}\right)$$

- $\lim_{Q\to\infty} H$ exists and is infrared safe
- collinear logarithms $\sum_{k,n=1}^{\infty} \alpha_s^k v_{kn} \ln^n(\mu/m_c)$ are resummed in $f_c(x,\mu/m_c)$
- no terms $\mathcal{O}(m_c/Q)$ in the remainder





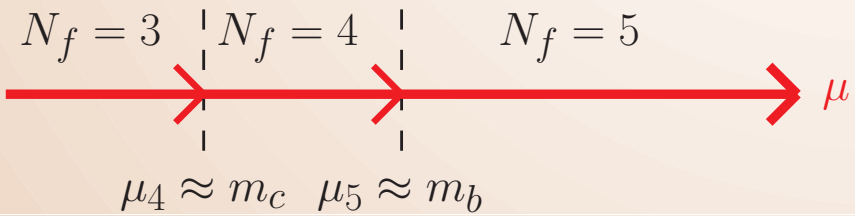
General-mass variable-flavor number scheme

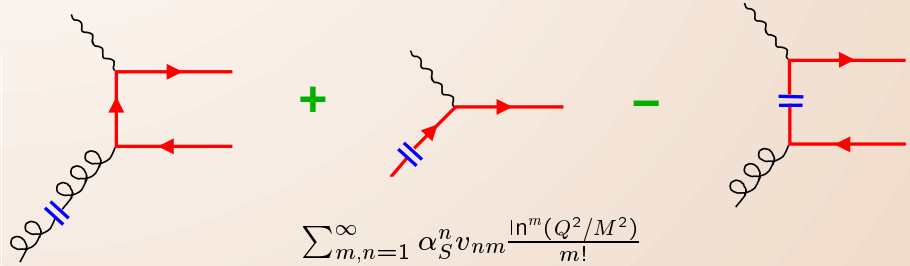
(Aivasis et al.; Chuvakin et al.; Thorne, Roberts; Kniehl et al.; Buza et al.; Cacciari et al.; ...)

Proved for *inclusive DIS* by J. Collins (1998)

$$F_2(x,Q,m_c) = \sum_a \int_\chi^1 \frac{d\xi}{\xi} H_a(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_c}{Q}) f_a(\xi, \frac{\mu}{m_c}) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{Q}\right)$$

- Works most effectively in DIS and Drell-Yan-like processes; practical implementation requires
  1. efficient treatment of mass dependence, rescaling of momentum fractions χ in processes with incoming c, b
  2. physically motivated factorization scale to ensure fast PQCD convergence (e.g., $\mu = Q$ in DIS)

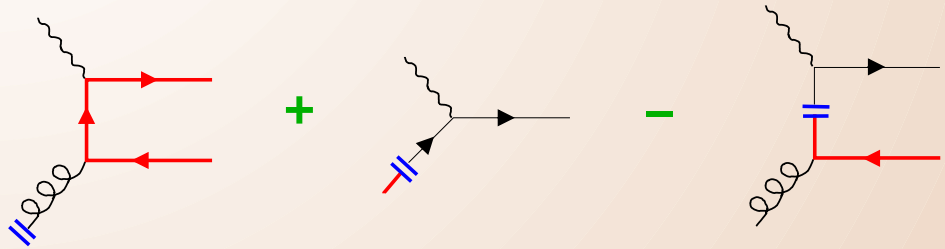




Simplified ACOT (χ) factorization scheme

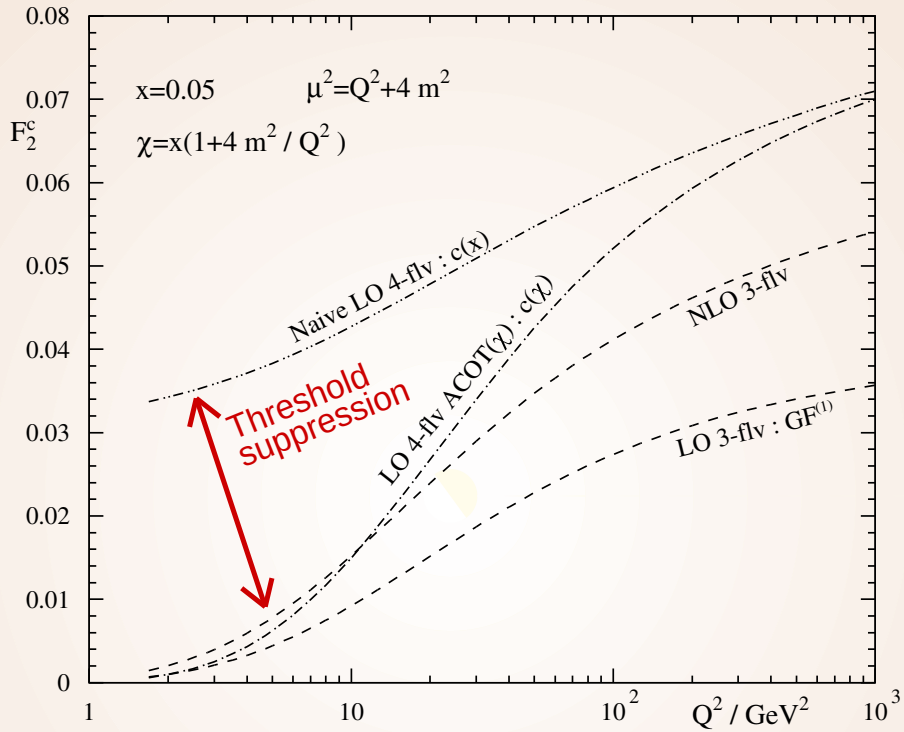
- Defined with $m_c = 0$ in Wilson coefficient functions H with *incoming* charm quarks *(Collins; Kramer, Olness, Soper)*
  - simplifications! close to the full ACOT scheme numerically
- Rescaled momentum fractions χ *(Barnett, Haber, Soper; Tung, Kretzer, Schmidt)*

In neutral-current DIS: $\chi = \begin{cases} x & \text{for incoming } g, q \\ x\left(1 + \frac{4m_c^2}{Q^2}\right) & \text{for incoming } c \end{cases}$



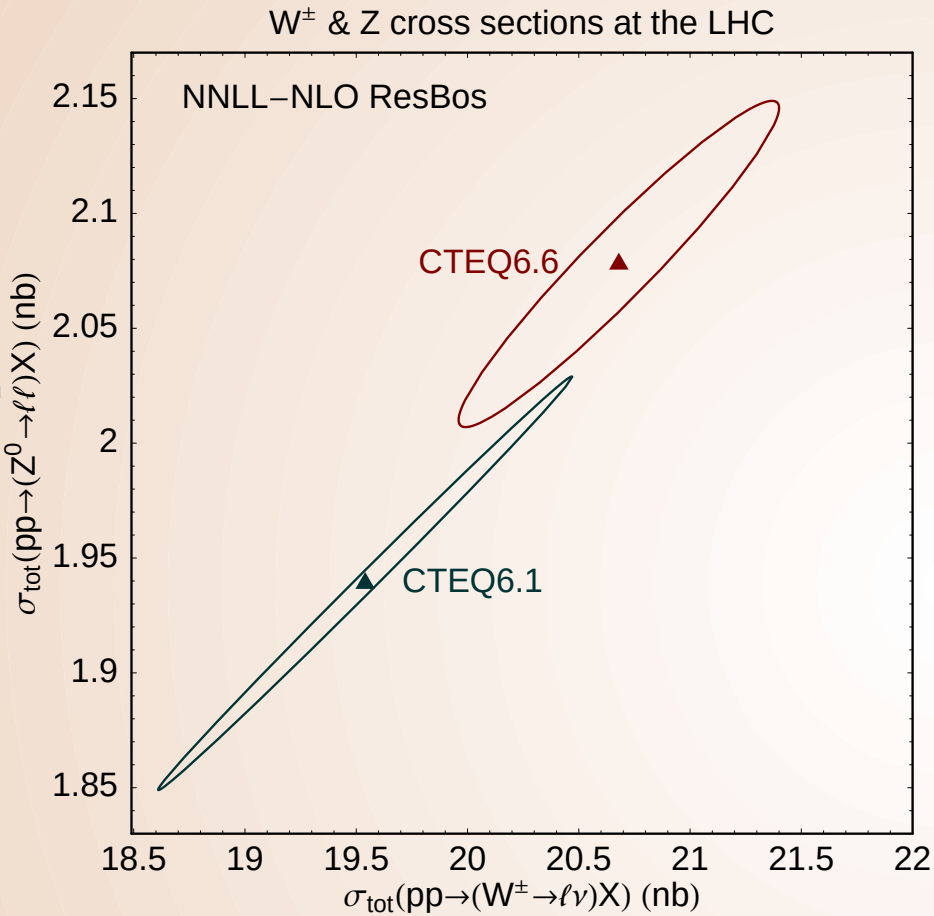
General-mass (ACOT-χ) factorization scheme

(Aivazis et al., 1994; Collins, 1998; Kramer, Olness, Soper, 2000; Tung, Kretzer, Schmidt, 2002; ...)



- Charm Wilson coefficient function is suppressed at $Q \to m_c$
- To keep agreement with DIS F_2 data, u, d, $\bar{u}$, $\bar{d}$ PDF's are enhanced at small x, as compared to the zero-mass (ZM-VFN) scheme

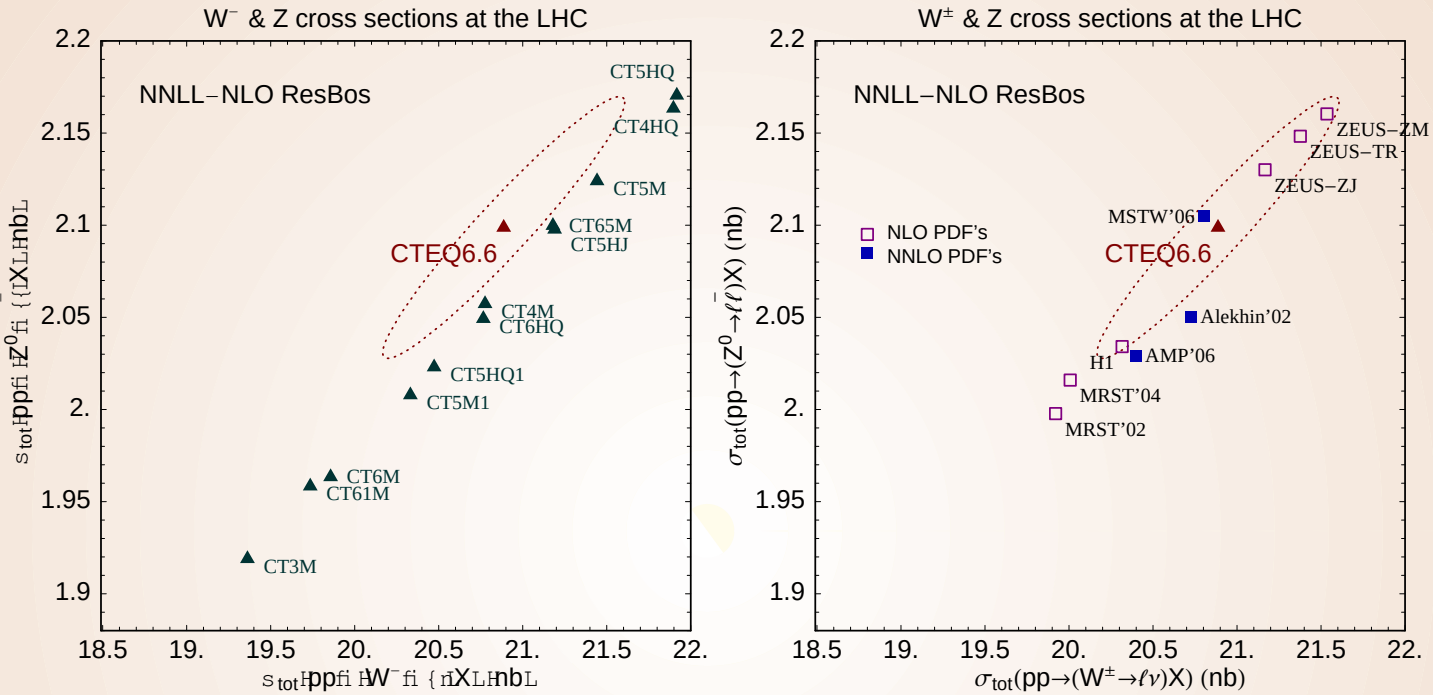
CTEQ6.6 W and Z cross sections at the LHC



- Computation done with NNLL-NLO ResBos *(Balazs, Ladinsky, P.N., Yuan)*, with the goal to estimate *relative* differences due to NLO/NNLO PDFs
- Effect of NNLO hard + NLO EW contributions is nearly independent of PDF's
- Ellipses: the PDF uncertainty for $\Delta\chi^2_{scaled} = 100$ (<90% c.l. for 2-dim dependence)

General-mass CTEQ6.6 predictions are higher by 6-7% compared to zero-mass CTEQ6.1 (enhanced CTEQ6.6 u, d PDF's at $x \sim 0.005$)

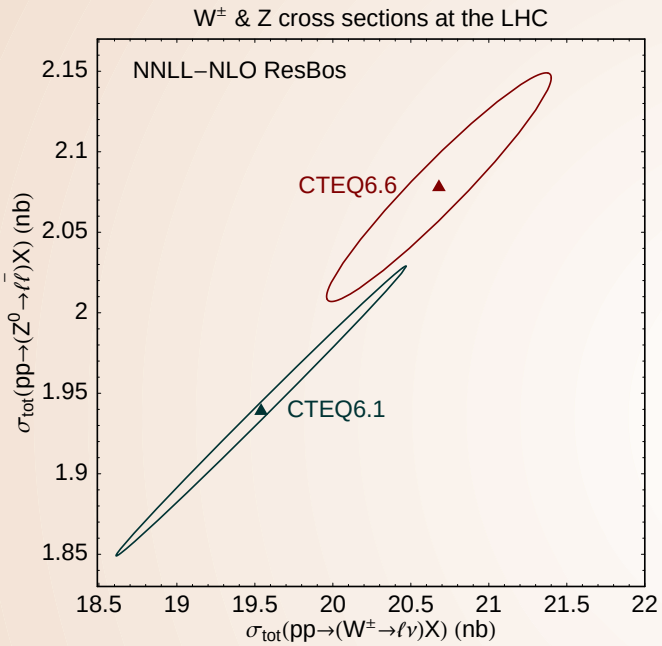
W and Z cross sections at the LHC



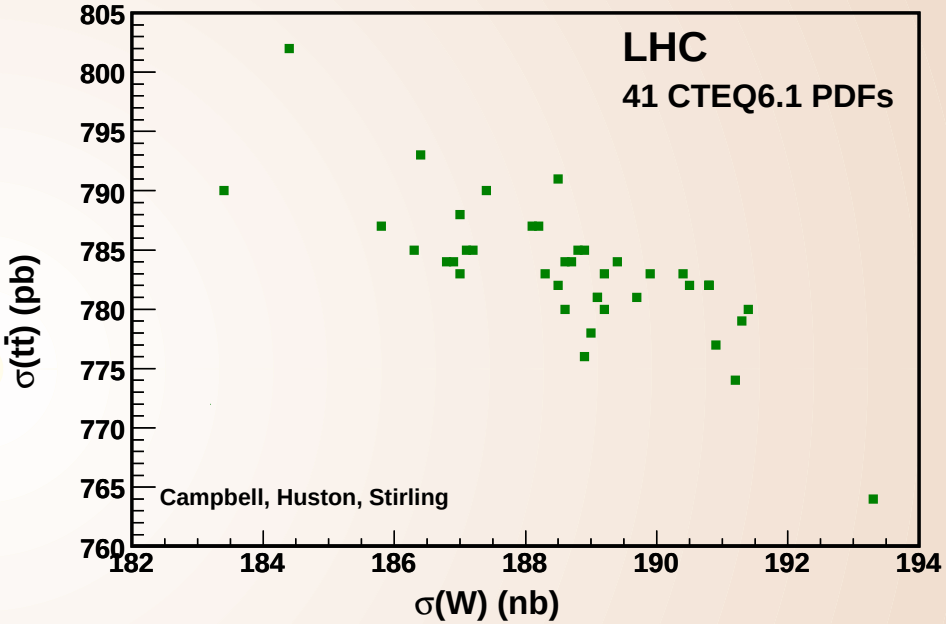
- Such changes in $\sigma_{Z,W}$ exceed NNLO corrections of $\approx -2\%$ or anticipated experimental error of $\sim 1\%$
- MSTW 2006 and 2008 predictions became compatible with the CTEQ6.6 result

PDF-induced correlations in W, Z, $t\bar{t}$ production

$W - Z$: correlated dependence



$W - t\bar{t}$ (or $Z - t\bar{t}$): anti-correlated dependence



- Which parton flavors drive the "experimental" PDF uncertainties and lead to (anti-)correlations between the cross sections?
- The PDF dependence of a cross section ratio σ_1/σ_2 is reduced (enhanced) if σ_1 and σ_2 are correlated (anticorrelated)

Computation of PDF errors in Hessian method

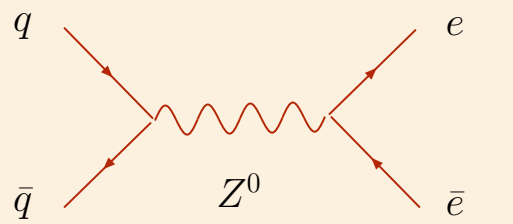
Z production at the LHC

Choose all that apply and select the x range

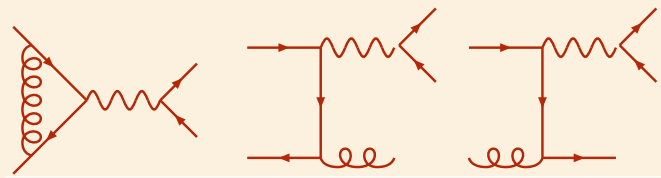
The PDF uncertainty in σ_Z is mostly due to...

1. u, d, $\bar{u}$, $\bar{d}$ PDF's
   at $x < 10^{-2}$ ($x > 10^{-2}$)
2. gluon PDF
   at $x < 10^{-2}$ ($x > 10^{-2}$)
3. s, c, b PDF's
   at $x < 10^{-2}$ ($x > 10^{-2}$)

Leading order

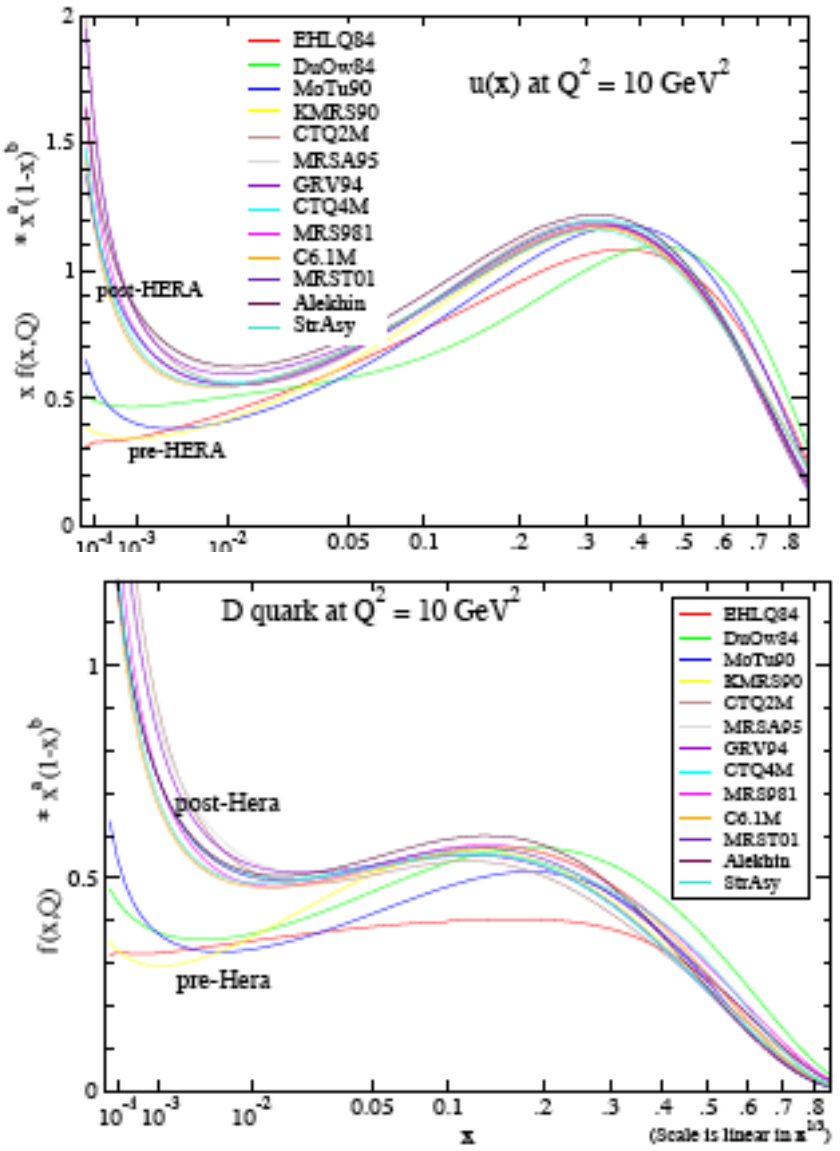


Next-to-leading order



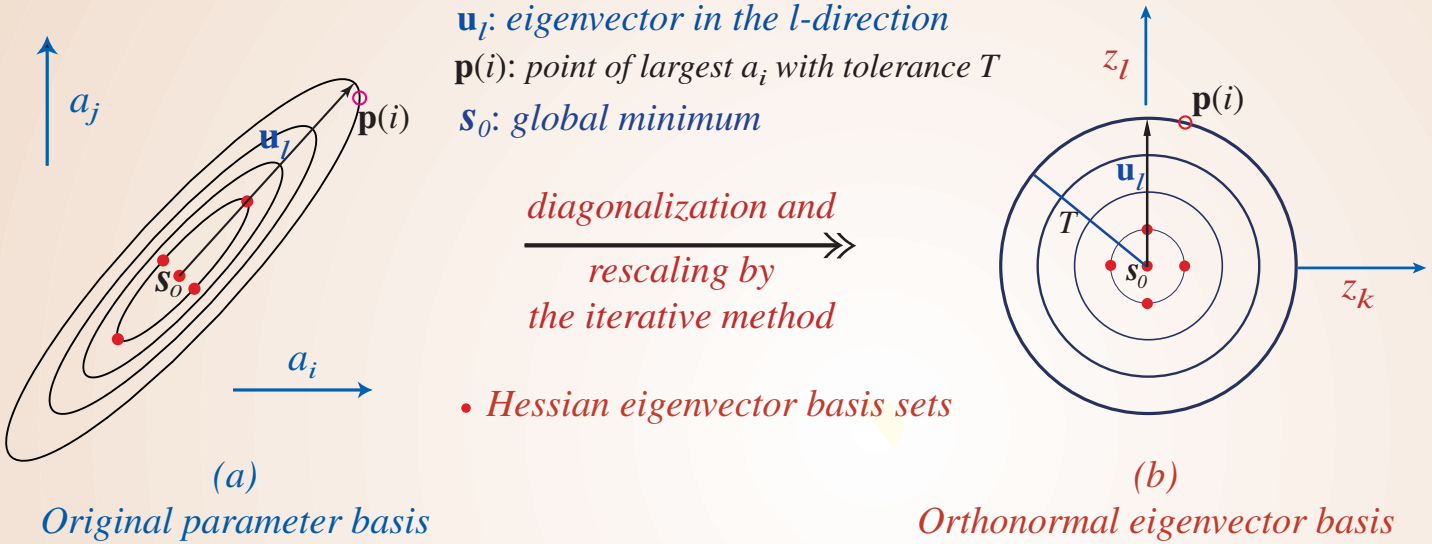
PDF Errors: old way

- Make plots of lots of pdf's (no matter how old) and take spread as a measure of the error
- Can either underestimate or overestimate the error
- Review sources of uncertainty on pdf's
  - data set choice
  - kinematic cuts
  - parametrization choices
  - treatment of heavy quarks
  - order of perturbation theory
  - errors on the data
- There are now more sophisticated techniques to deal with at least the errors due to the experimental data uncertainties



Tolerance hypersphere in the PDF space

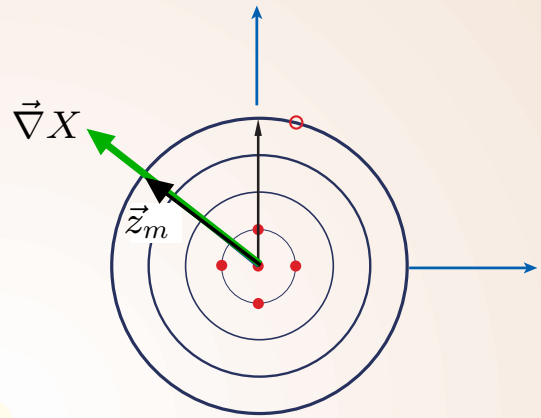
2-dim (i,j) rendition of N-dim (22) PDF parameter space



A hyperellipse $\Delta\chi^2 \leq T^2$ in space of N physical PDF parameters $\{a_i\}$ is mapped onto a hypersphere of radius T in space of N orthonormal PDF parameters $\{z_i\}$

Tolerance hypersphere in the PDF space

2-dim (i,j) rendition of N-dim (22) PDF parameter space



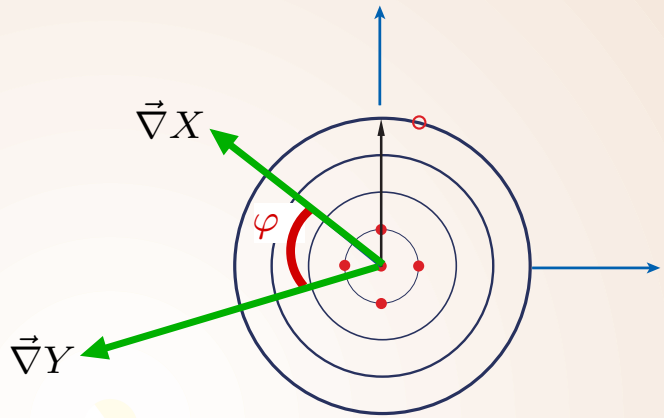
(b)
Orthonormal eigenvector basis

PDF error for a physical observable X is given by

$$\Delta X = \vec{\nabla} X \cdot \vec{z}_m = \left|\vec{\nabla} X\right| = \frac{1}{2}\sqrt{\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)^2}$$

Tolerance hypersphere in the PDF space

2-dim (i,j) rendition of N-dim (22) PDF parameter space



(b)
Orthonormal eigenvector basis

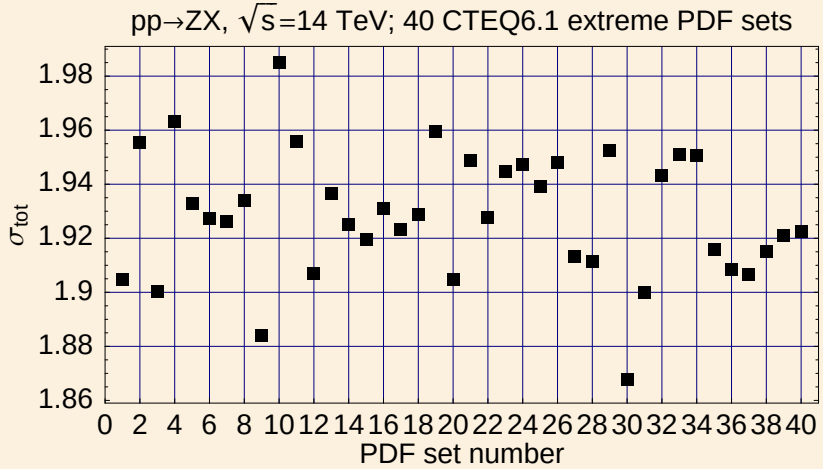
Correlation cosine for observables X and Y:

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

An inefficient application of the error analysis

☺ Compute σ_Z for 40 (now 44) extreme PDF eigensets

☺ Find eigenparameter(s) producing largest variation(s), such as #9, 10, 30



☹ It is not obvious how to relate abstract eigenparameters to physical PDF's $u(x)$, $d(x)$, etc.

Correlation analysis for collider observables

(J. Pumplin et al., PRD 65, 014013 (2002); P.N. and Z. Sullivan, hep-ph/0110378)

A technique based on the Hessian method to relate the PDF uncertainty in physical cross sections to PDF's of specific flavors at known (x, μ)

For $2N$ PDF eigensets and two cross sections X and Y:

$$\Delta X = \frac{1}{2}\sqrt{\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)^2}$$

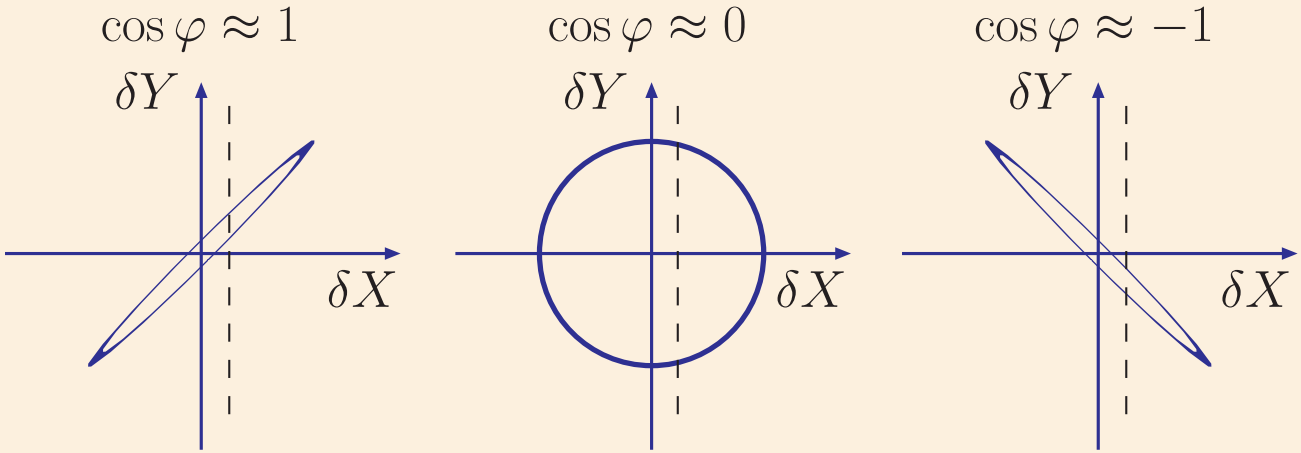
$$\cos\varphi = \frac{1}{4\Delta X\,\Delta Y}\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)\left(Y_i^{(+)} - Y_i^{(-)}\right)$$

$X_i^{(\pm)}$ are maximal (minimal) values of X_i tolerated along the i-th PDF eigenvector direction; $N = 22$ for the CTEQ6.6 set

Correlation angle φ

Determines the parametric form of the $X - Y$ correlation ellipse

$$X = X_0 + \Delta X \cos\theta$$
$$Y = Y_0 + \Delta Y \cos(\theta + \varphi)$$



X_0, Y_0: best-fit values

ΔX, ΔY: PDF errors

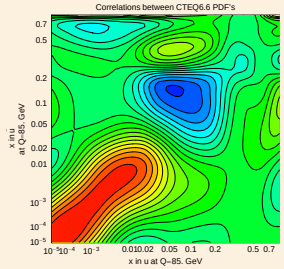
$\cos\varphi \approx \pm 1$: Measurement of X imposes tight constraints on Y

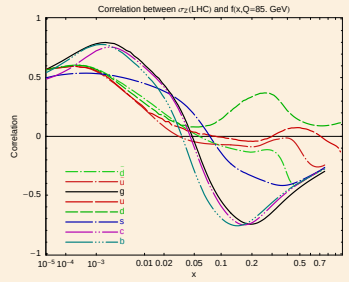
$\cos\varphi \approx 0$: Measurement of X imposes loose constraints on Y

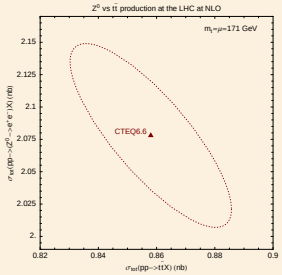
Types of correlations

X and Y can be

- two PDFs $f_1(x_1, Q_1)$ and $f_2(x_2, Q_2)$ (plotted as $\cos\varphi$ vs x_1 & x_2)
- a physical cross section σ and PDF $f(x, Q)$ (plotted as $\cos\varphi$ vs x)
- two cross sections σ_1 and σ_2

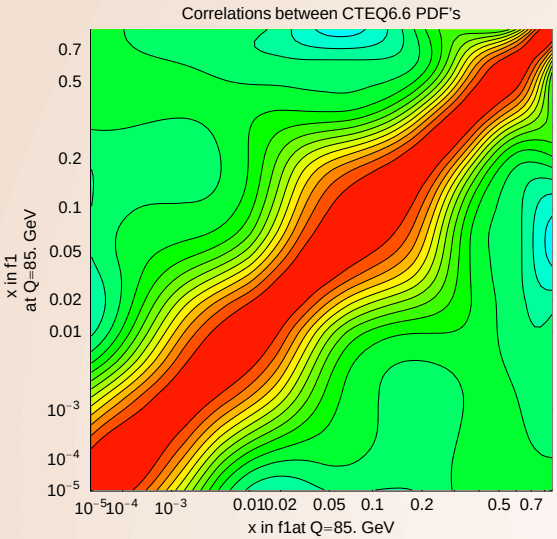




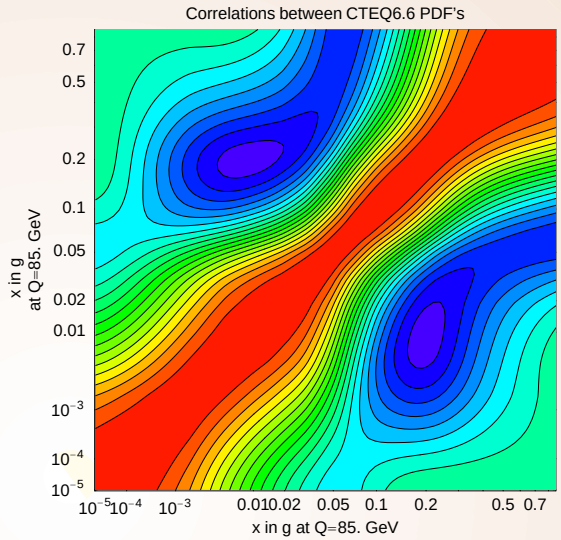


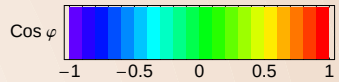
Correlations between $f(x_1, Q)$ and $f(x_2, Q)$ at $Q = 85$ GeV

$u(x_1, Q)$ vs. $u(x_2, Q)$



$g(x_1, Q)$ vs. $g(x_2, Q)$



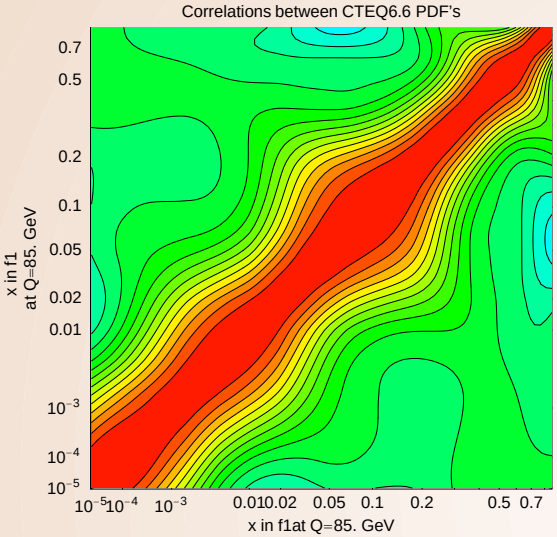


- Momentum sum rule (affects $g(x, Q)$):

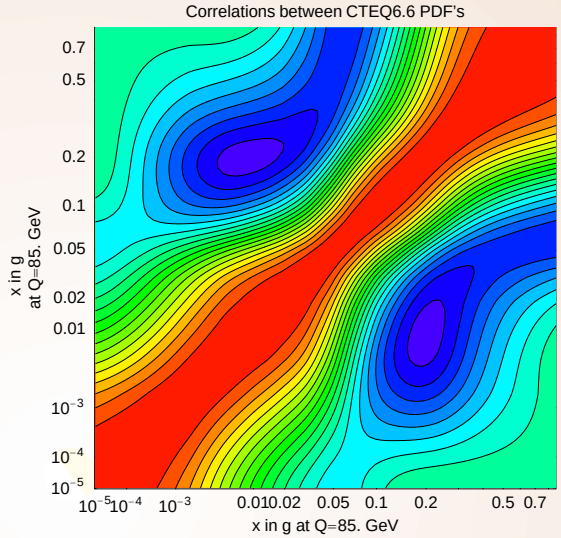
$$\int_0^1 x g(x) dx + \sum_{i=1}^{N_f} \int_0^1 x \left[q_i(x) + \bar{q}_i(x) \right] dx = 1$$

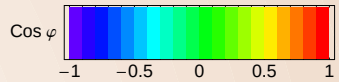
Correlations between $f(x_1, Q)$ and $f(x_2, Q)$ at $Q = 85$ GeV

$u(x_1, Q)$ vs. $u(x_2, Q)$



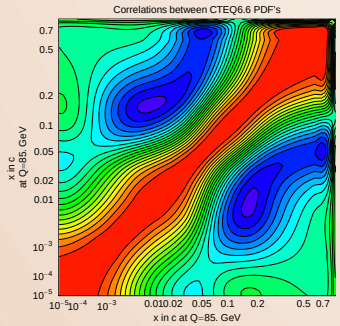
$g(x_1, Q)$ vs. $g(x_2, Q)$



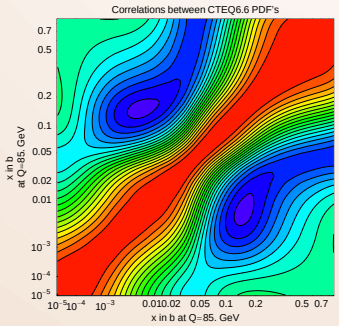


Correlation patterns look similar for g, c, b PDF's (no intrinsic charm here!)

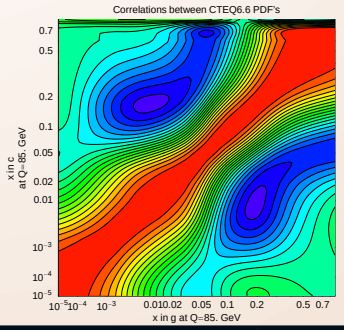
c vs. c



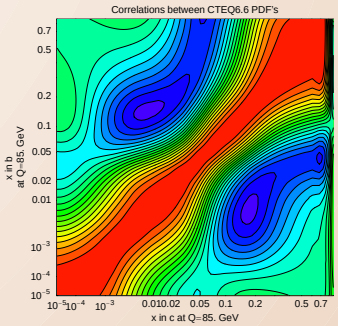
b vs. b



c vs. g



b vs. c



Try it yourself

Correlations between CTEQ6.6 PDF's

http://hep.pa.msu.edu/cteq/public/6.6/pdfcorrs/index.html

CTEQ6.6 webpage

MTA SZTAKI: ... Dictionary | CSCNotesLis...las < TWiki | PatVancouve...las < TWiki | PhysicsAnaly...las < TWiki | Quick guide...nda monitor | http://www....ession.mp3 | Quick guide...nda monitor | Alliance to S... Tax Credits

CTEQ6.6: correlations between parton distribution functions

Main reference | **PDF uncertainty bands** | **CTEQ6.6 *W*, *Z*, ttbar cross sections** | **Additional correlation plots**

A collection of 2-dimensional contour plots showing the correlation cosine cos(phi) for two parton distribution functions (PDF's) $f_a(x,Q)$ obtained in the CTEQ6.6 global analysis. The axes specify momentum fractions x in the two PDF's with specified flavors and factorization scales. The color (or gray shade) of the area is chosen to reflect the value of the correlation cosine at each (x_1, x_2) point according to the scale shown below. Both axes are scaled as $x^{0.2}$.

Flavor of the PDF 1: gluon

Factorization scale of the PDF 1: Q=85 GeV

Flavor of the PDF 2: gluon

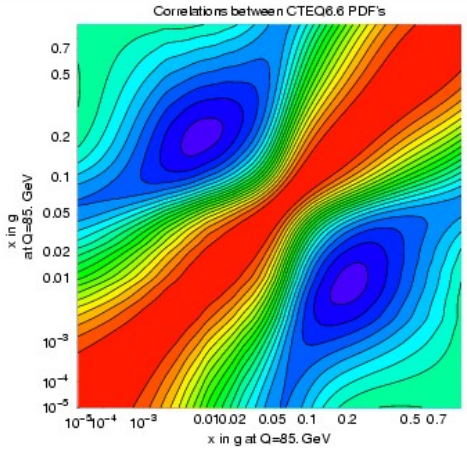
Factorization scale of the PDF 2: Q=85 GeV

Color | Grayscale

−1 −0.5 0 0.5 1

Cos φ

Color scale **eps** **png**

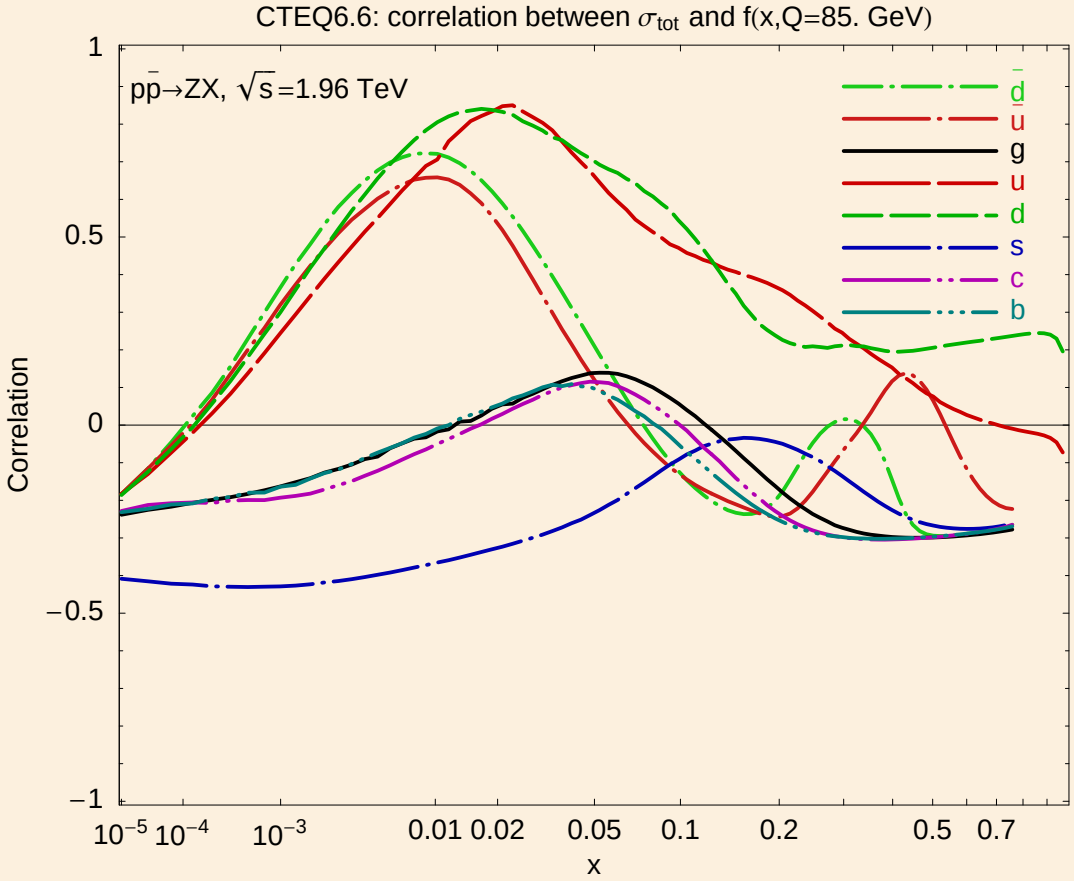


Download **eps** **png**

Reference: arXiV:0802.0007 [hep-ph]

Correlations $\cos\varphi$ between W, Z cross sections and PDF's

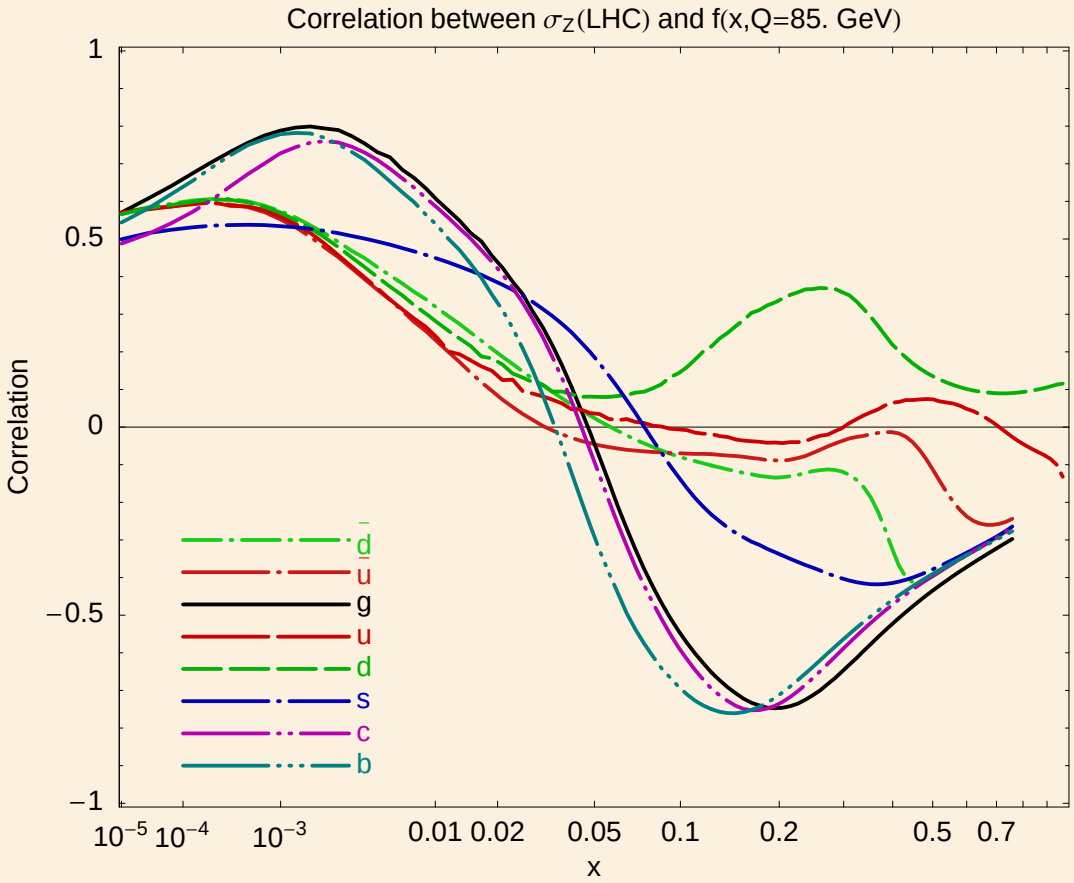
Tevatron Run-2



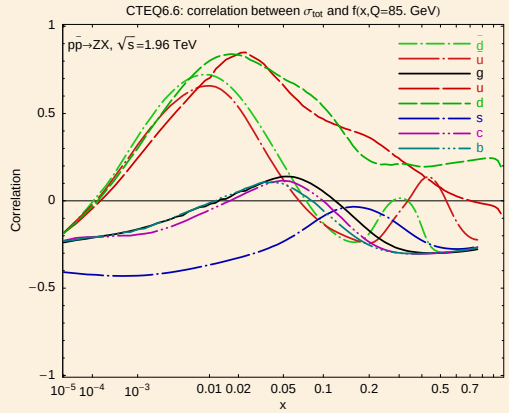
Similar correlations for W production

Correlations $\cos\varphi$ between W, Z cross sections and PDF's

LHC



Tevatron Run-2

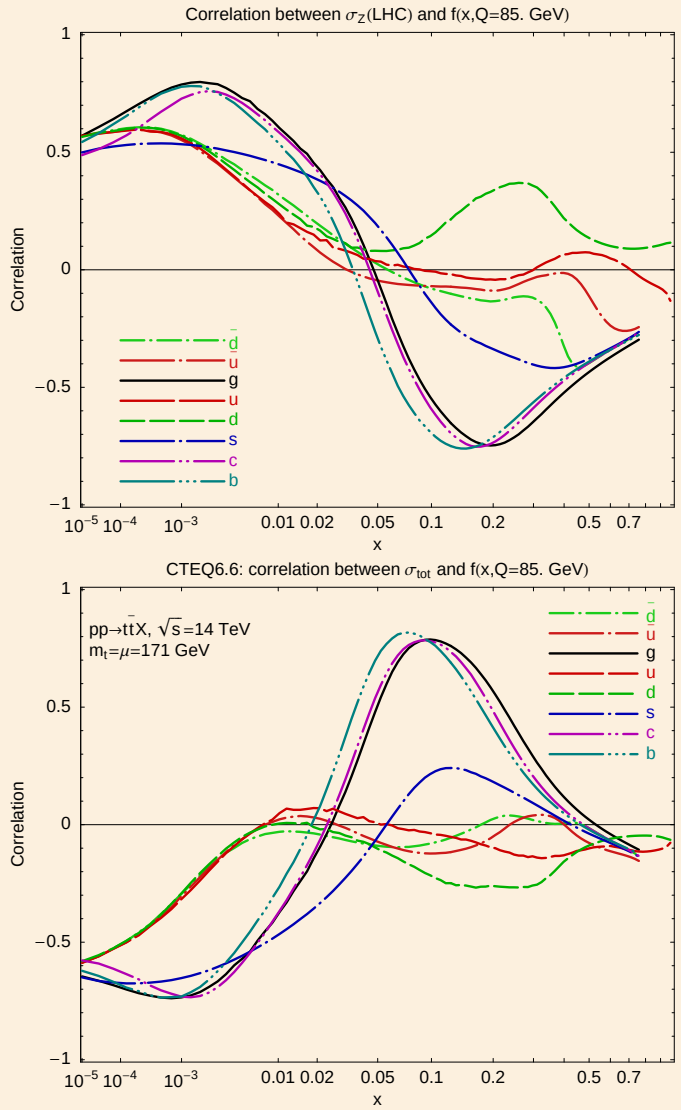


Similar correlations for W production

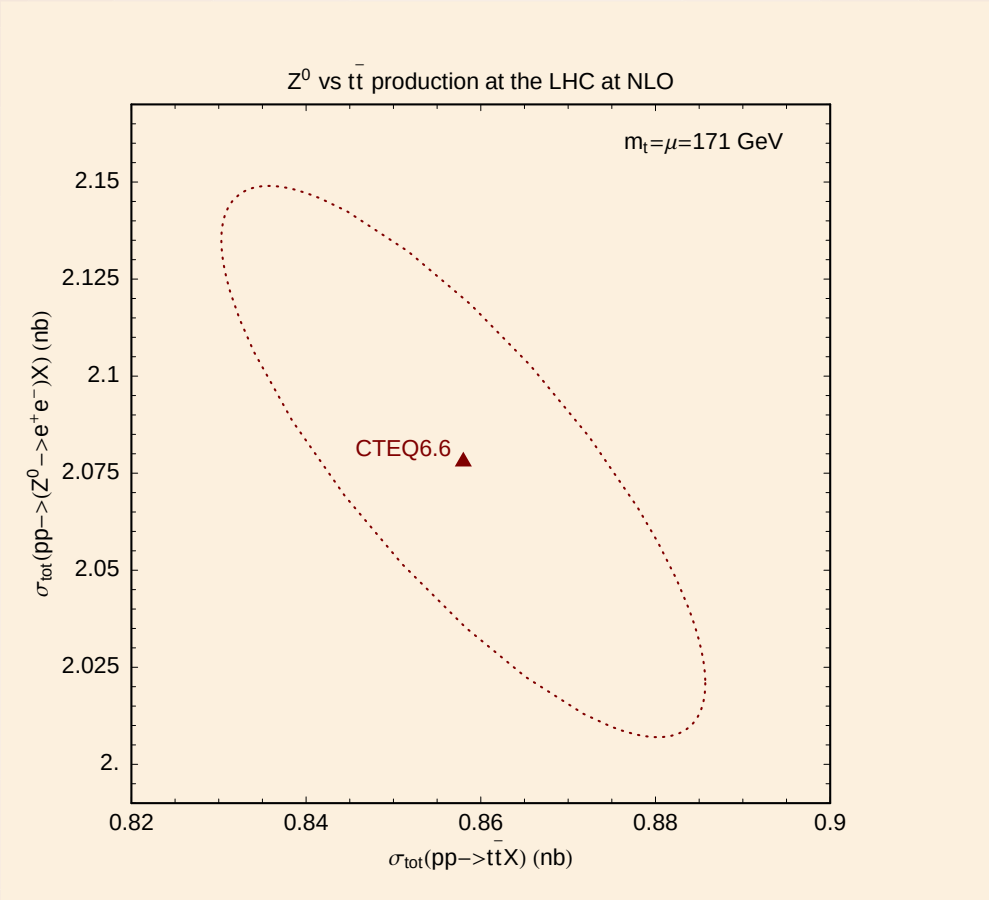
Correlations of Z and $t\bar{t}$ cross sections with PDF's

LHC Z, W cross sections are strongly correlated with $g(x)$, $c(x)$, $b(x)$ at $x \sim 0.005$

$\therefore$ they are strongly anticorrelated with processes sensitive to $g(x)$ at $x \sim 0.1$
($t\bar{t}$, $gg \to H$ for $M_H > 300$ GeV)

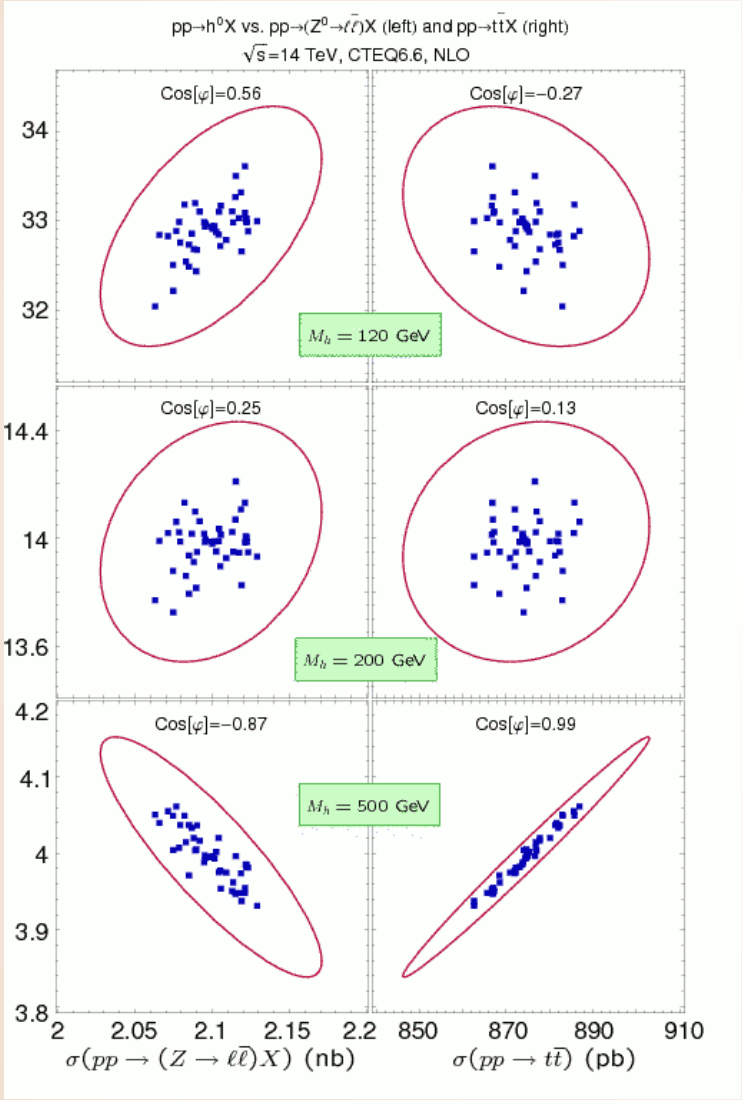


$t\bar{t}$ vs Z cross sections at the LHC



Measurements of $\sigma_{t\bar{t}}$ and σ_Z probe the same (gluon) PDF degrees of freedom at different x values

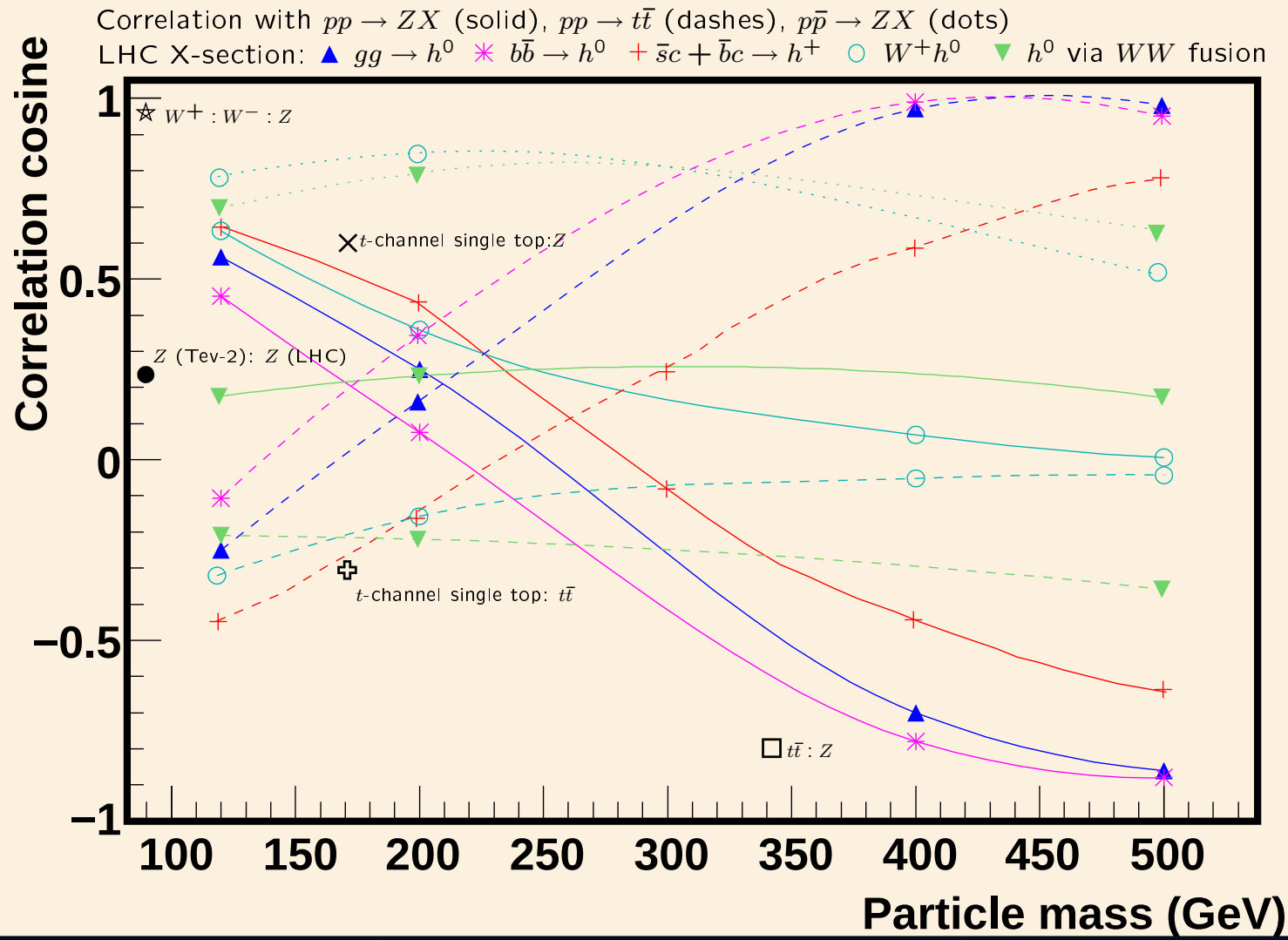
Correlations between $\sigma(gg \to H^0)$, σ_Z, $\sigma_{t\bar{t}}$



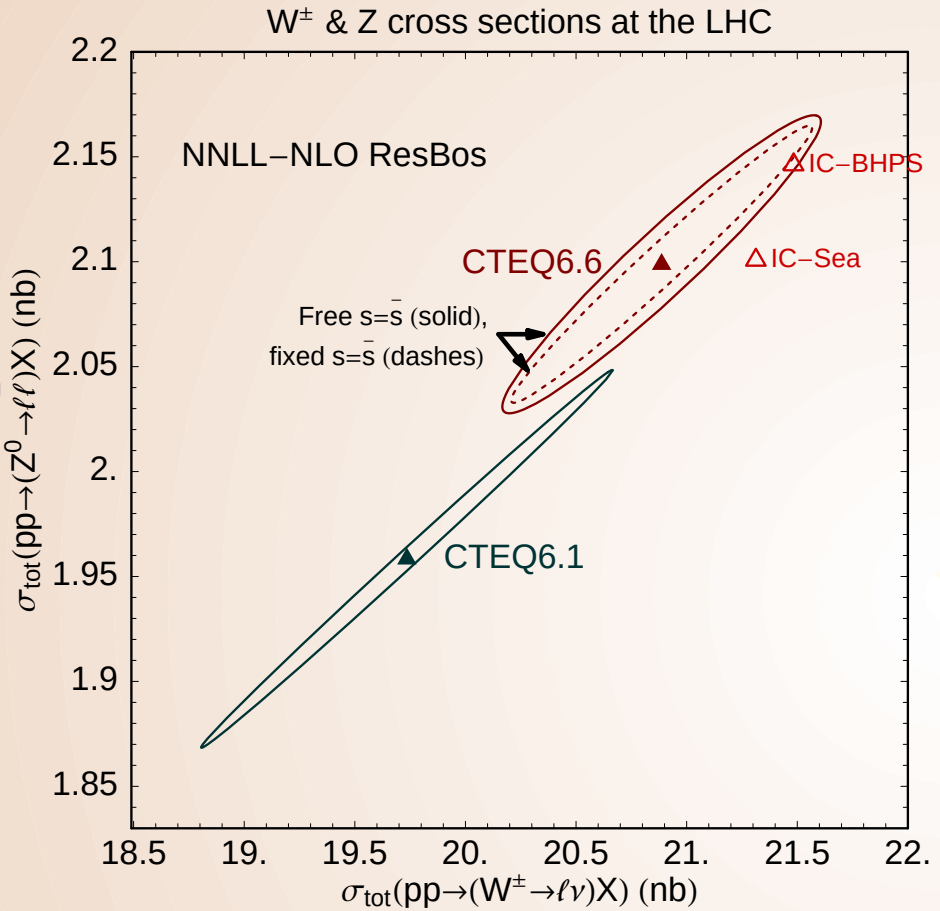
As M_H increases:

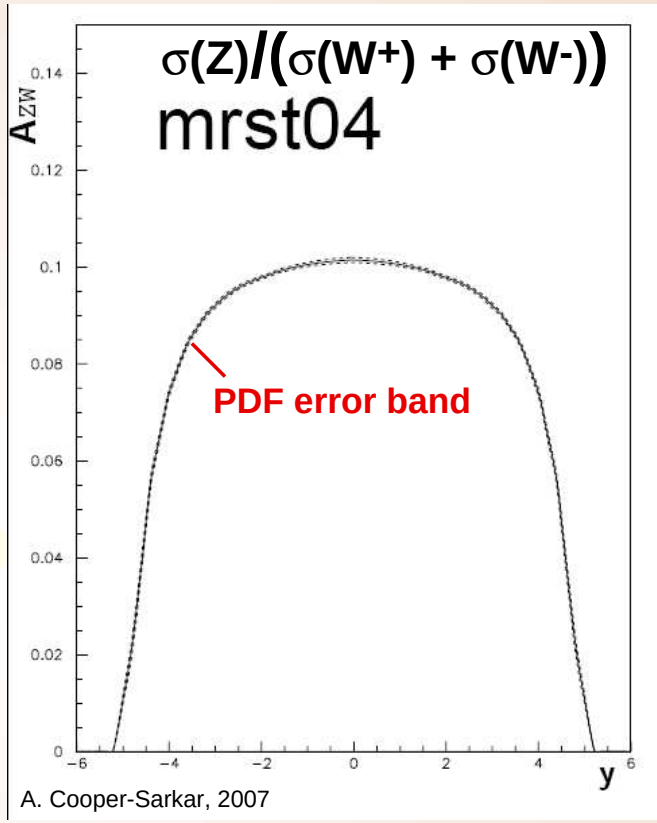
- $\cos\varphi(\sigma_H, \sigma_Z)$ decreases
- $\cos\varphi(\sigma_H, \sigma_{t\bar{t}})$ increases

$\cos\varphi$ for various NLO Higgs production cross sections in SM and MSSM



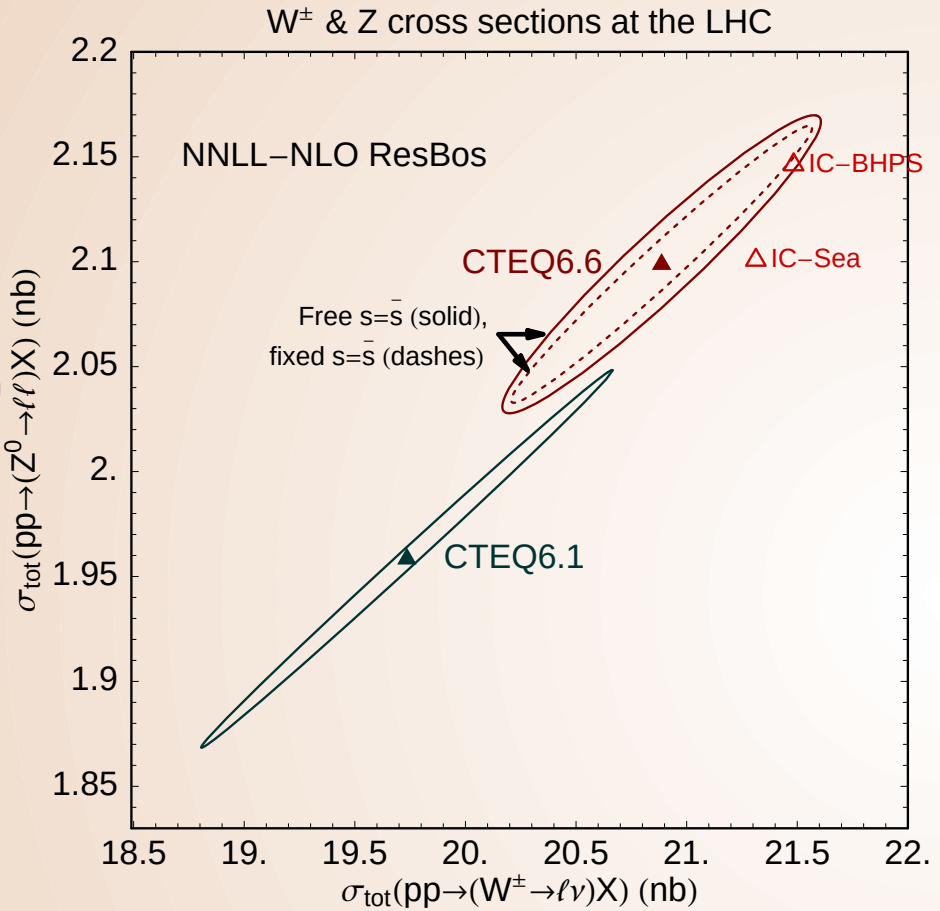
W and Z cross sections and their ratio

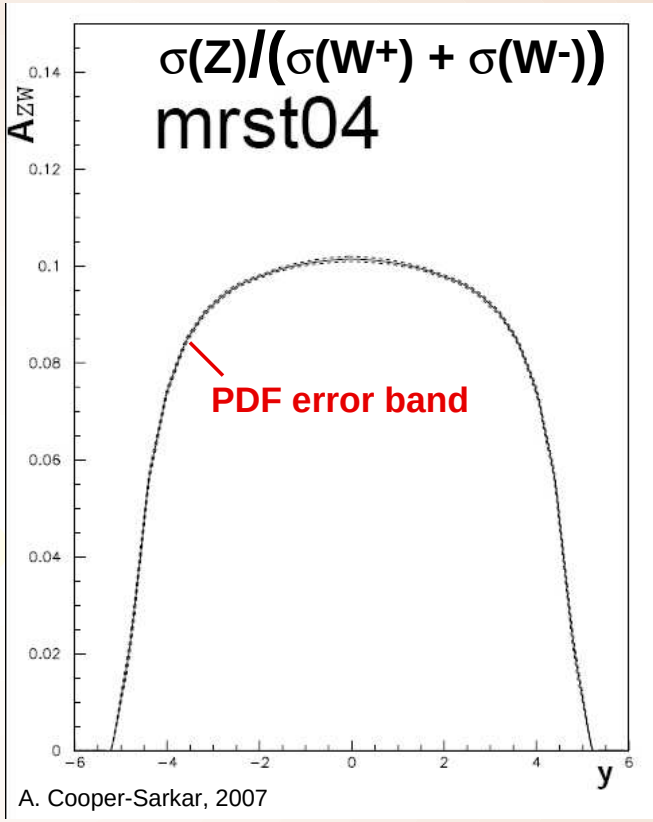




- Radiative contributions have similar structure in $W^{\pm}$ and Z cross sections; cancel well in Xsection ratios
- The PDF uncertainty cancels partially because of differences in s, c, b scattering contributions

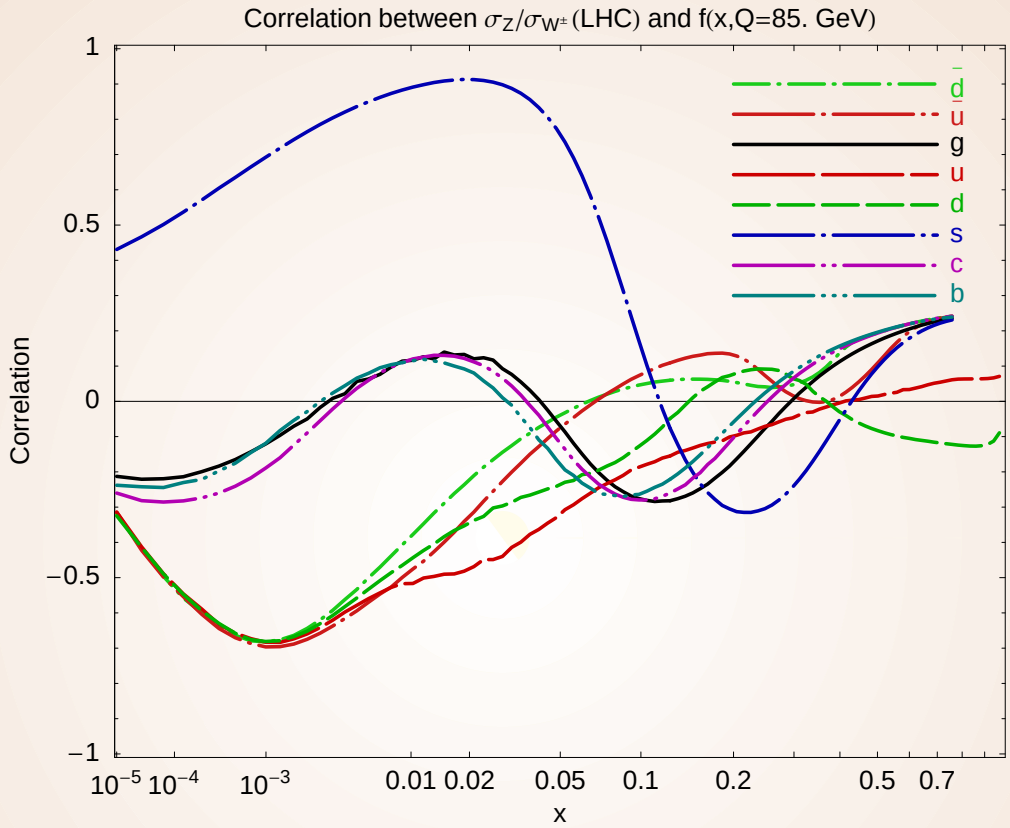
W and Z cross sections and their ratio





- 27% of $\sigma_{NLO}(W^{\pm})$ from $c\bar{s} \to W^{\pm}$, 20% of $\sigma_{NLO}(Z^0)$ from $s\bar{s} \to Z^0$
- non-negligible effects from free strangeness and intrinsic charm (IC) PDF's

σ_Z/σ_W at the LHC



The remaining PDF uncertainty in σ_Z/σ_W is mostly driven by $s(x)$; increases by a factor of 3 compared to CTEQ6.1 as a result of free strangeness in CTEQ6.6

Conclusions

- Narrow PDF error bands can be misleading; theoretical improvements for all fitted experiments and the analysis of the entirety of contributing factors must continue
- CTEQ6.6 PDF's in the general-mass scheme:
  - important differences from ZM-VFNS and some CTEQ6.5 predictions
  - must be used as the standard CTEQ set from now on
- A new technique to study PDF-induced correlations between physical observables
- other ongoing efforts: NNLO/small-x/Q_T resummation in the global fits, impact of Run-2 jet data on the gluon PDF, PDF's for leading-order Monte-Carlos, a ROOT interface for PDF reweighting in Monte-Carlo programs; stay tuned!

Backup slides

Z, W, $t\bar{t}$ cross sections and correlations

Table: Total cross sections σ, PDF-induced errors $\Delta\sigma$, and correlation cosines $\cos\varphi$ for Z^0, $W^\pm$, and $t\bar{t}$ production at the Tevatron Run-2 (Tev2) and LHC, computed with CTEQ6.6 PDFs.

$\sqrt{s}$ (TeV)	Scattering process	$\sigma, \Delta\sigma$ (pb)	Correlation $\cos\varphi$ with Z^0 (Tev2)	$W^\pm$ (Tev2)	Z^0 (LHC)	$W^\pm$ (LHC)
1.96	$p\bar{p} \to (Z^0 \to \ell^+\ell^-)X$	241(8)	1	**0.987**	0.23	0.33
	$p\bar{p} \to (W^\pm \to \ell\nu_\ell)X$	2560(40)	**0.987**	1	0.27	0.37
	$p\bar{p} \to t\bar{t}X$	7.2(5)	-0.03	-0.09	-0.52	-0.52
14	$pp \to (Z^0 \to \ell^+\ell^-)X$	2080(70)	0.23	0.27	1	**0.956**
	$pp \to (W^\pm \to \ell\nu)X$	20880(740)	0.33	0.37	**0.956**	1
	$pp \to (W^+ \to \ell^+\nu_\ell)X$	12070(410)	0.32	0.36	**0.928**	**0.988**
	$pp \to (W^- \to \ell^-\bar{\nu}_\ell)X$	8810(330)	0.33	0.38	**0.960**	**0.981**
	$pp \to t\bar{t}X$	860(30)	-0.14	-0.13	**-0.80**	**-0.74**

Correlations with single-top cross sections

Table: Correlation cosines $\cos\varphi$ between single-top, W, Z, and $t\bar{t}$ cross sections at the Tevatron Run-2 (Tev2) and LHC, computed with CTEQ6.6 PDFs.

Single-top production channel	Correlation $\cos\varphi$ with Z^0 (Tev2)	$W^\pm$ (Tev2)	$t\bar{t}$ (Tev2)	Z^0 (LHC)	$W^\pm$ (LHC)	$t\bar{t}$ (LHC)
$t-$channel (Tev2)	-0.18	-0.22	**0.81**	**-0.82**	**-0.79**	0.56
$t-$channel (LHC)	0.09	0.14	-0.64	0.56	0.53	-0.42
$s-$channel (Tev2)	**0.83**	**0.79**	0.18	0.22	0.27	-0.3
$s-$channel (LHC)	**0.81**	**0.85**	-0.42	0.6	0.68	-0.33

$t\bar{t}$ production as a standard candle process

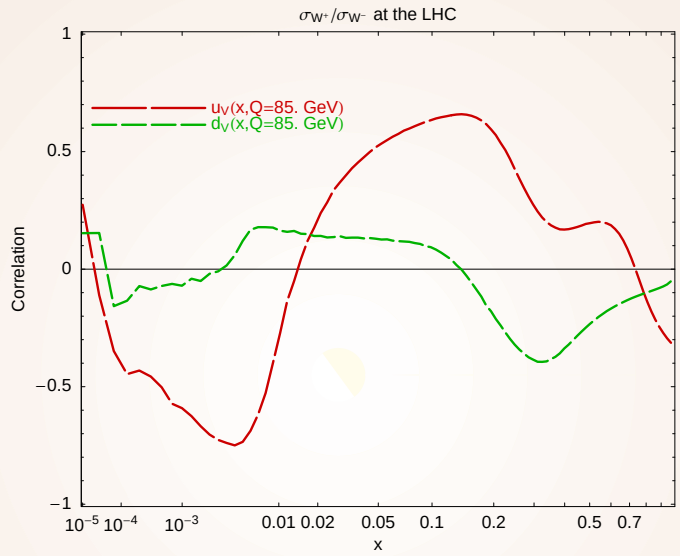
Uncertainties in $\sigma_{t\bar{t}}$ for $m_t = 171$ GeV

Type	Current	Projected	Assumptions
Scale dependence	11% (NLO)	$\sim 3-5\%$? (NNLO+resum.)	$m_t/2 \leq \mu \leq 2m_t$
PDF dependence	2%	1%?	1σ c.l.
m_t dependence	5% $\delta m_t = 2$ GeV	$< 3\%$ $\delta m_t = 1$ GeV	
Total (theory)	12%	$\sim 5\%$	
Experiment	8% (CDF)	5%?	

- Measurements of $\sigma_{t\bar{t}}$ with accuracy $\sim 5\%$ may be within reach; useful for monitoring of $\mathcal{L}_{LHC}$ in the first years, normalization of cross sections sensitive to large-x glue scattering, as well as for new physics searches *(reviewed by T. Han in arXiv:0804.3178)*

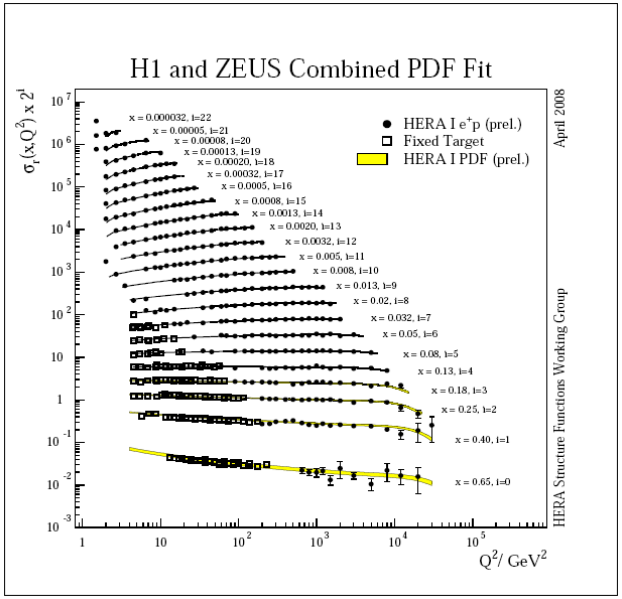
Updated theory estimates in Cacciari et al., arXiv:0804.2800; Moch, Uwer, arXiv:0804.1476

$\sigma(W^+)/\sigma(W^-)$



$\sigma(W^+)/\sigma(W^-) = 1.36 + 0.016$ (CTEQ6.6), 1.36 (MSTW'06NNLO), 1.35 (MRST'04NLO)

Combined HERA-1 data set on neutral-current DIS cross sections



New HERA-I PDF fit predictions vs. H1/ZEUS combined data for NC e^+p.

Total uncertainties on the PDF fit predictions are included but can barely be resolved.

Averaging of the data leads to cross-calibration of H1 and ZEUS measurements, drastic reduction of systematic uncertainties

Gang Li, HERALHC workshop, 2008

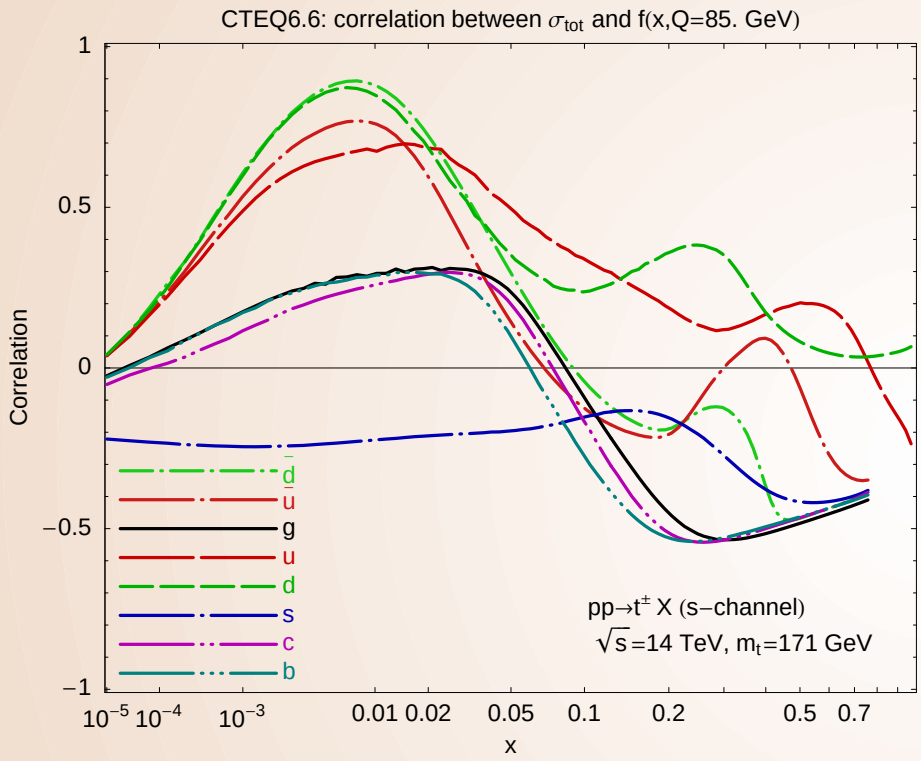
Origin of the differences with older PDF sets

1. Corrections of wrong or outdated assumptions

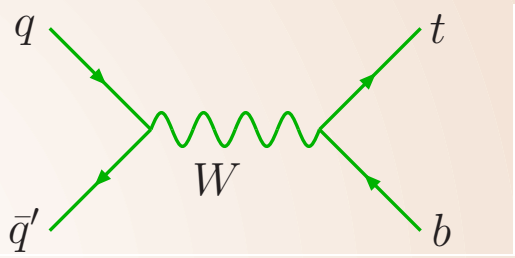
- inclusion of (N)NLO QCD hard scattering contributions
- improved treatment of heavy-quark (s, c, b) scattering
- relaxation of ad hoc constraints on PDF parametrizations
- improved numerical approximations
- ...

Origin of the differences with older PDF sets

An example of a small correlation with the gluon



Single-top production (NLO)



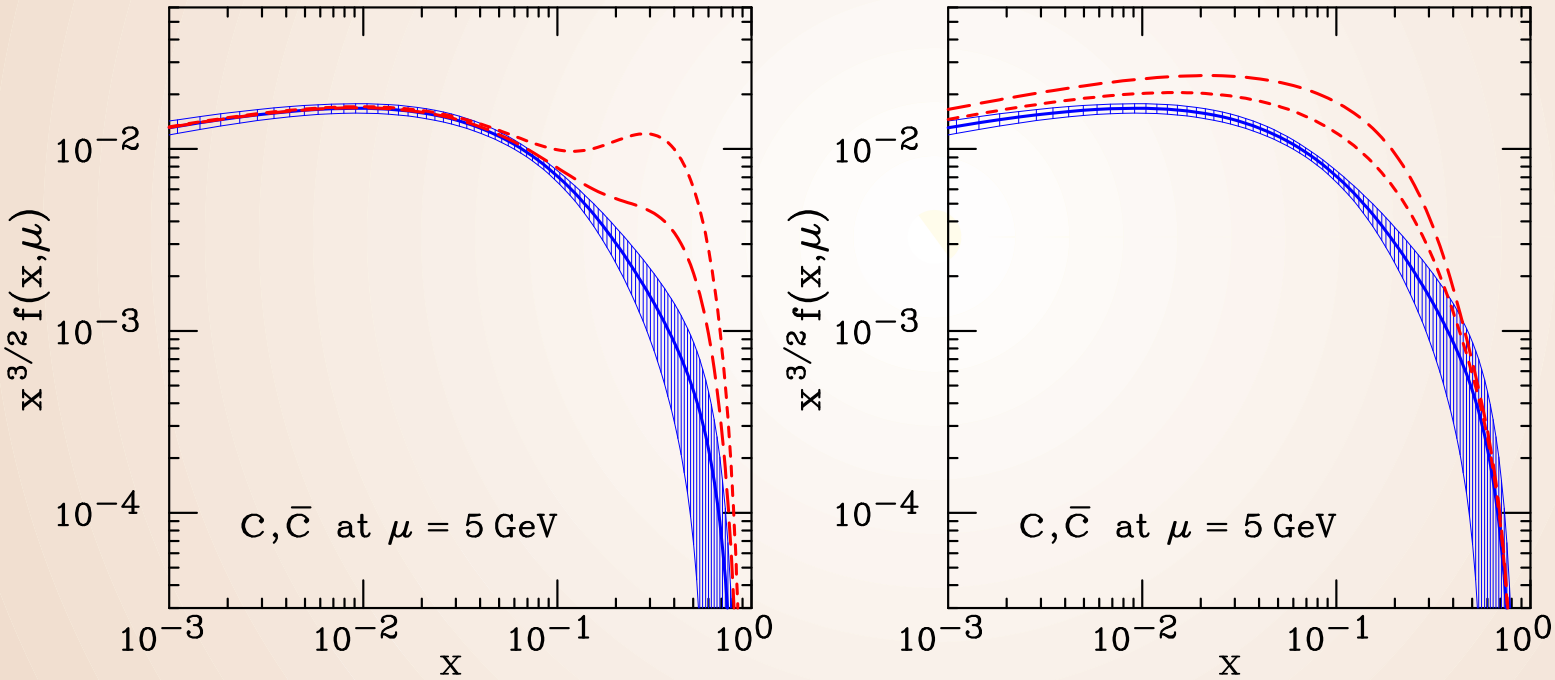
- typical $x \sim 0.01$
- mostly correlated with u, d PDF's

PDF uncertainties in W, Z *total* cross sections are irrelevant for some quark scattering processes (single-top, Z', ...)

Special PDF's with nonperturbative charm

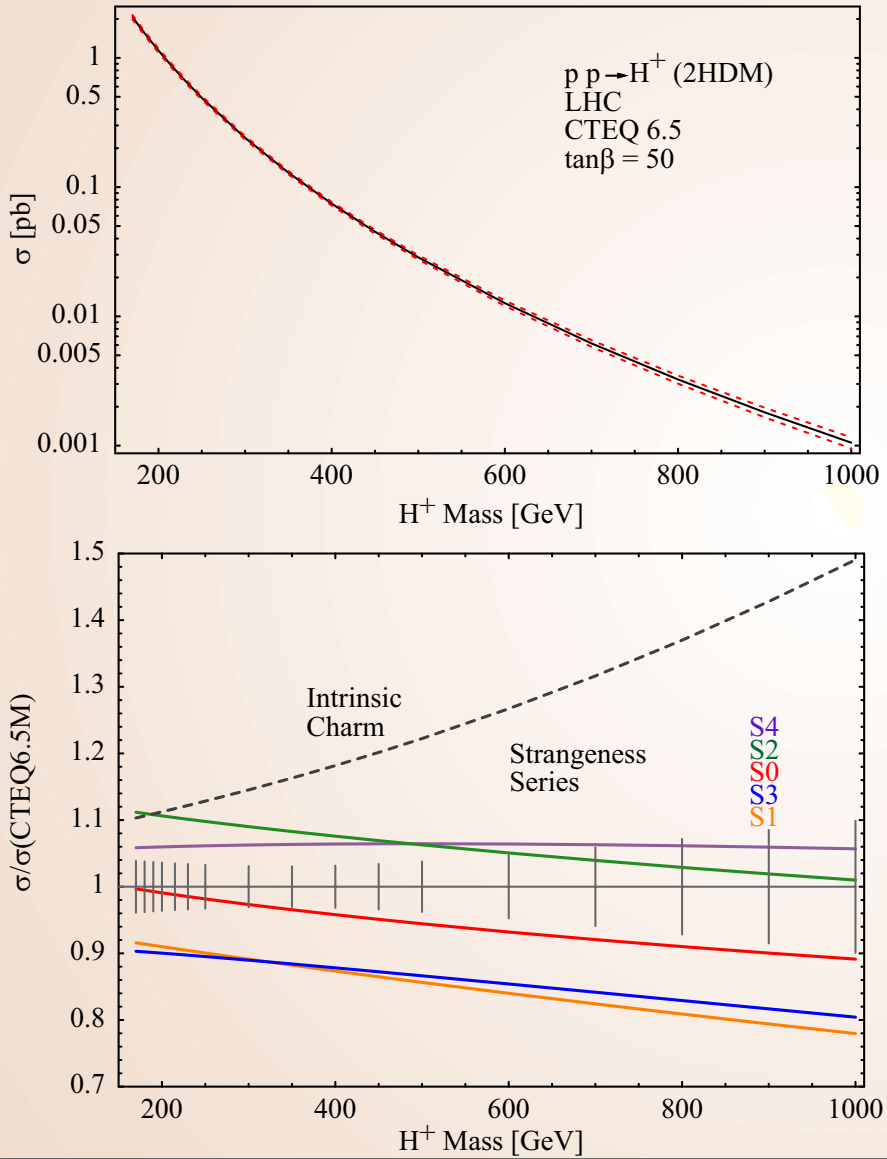
(Pumplin et al., 2007; updated in CTEQ6.6C)

Three models responsible for intrinsic charm generation (light-cone, meson-cloud, and phenomenological sea-like), with $\langle x \rangle_{c+\bar{c}}$ up to 3.5% at scale Q_0



The enhancement in $c(x, Q)$ persists at all practical Q, can be observed at the Tevatron and LHC

$c\bar{s} + c\bar{b} \to H^+$ in 2-Higgs doublet model at the LHC



CTEQ6.6M uncertainty band covers most of the CTEQ6.5 uncertainty due to strangeness

"Maximum-strength" sea-like IC leads to large enhancement $\Rightarrow$ new measurements ($p\bar{p} \to ZcX$?) are needed to constrain it!