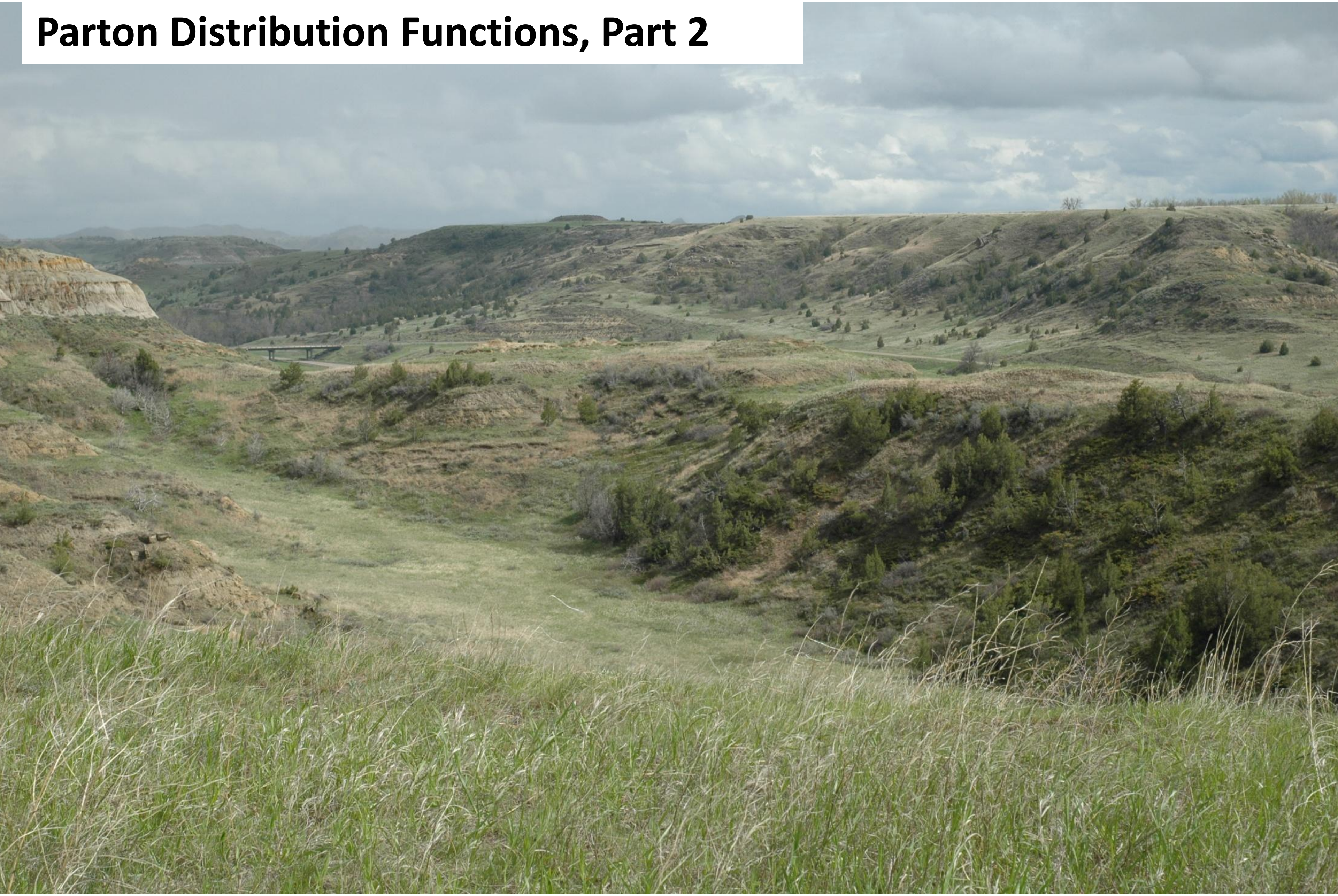
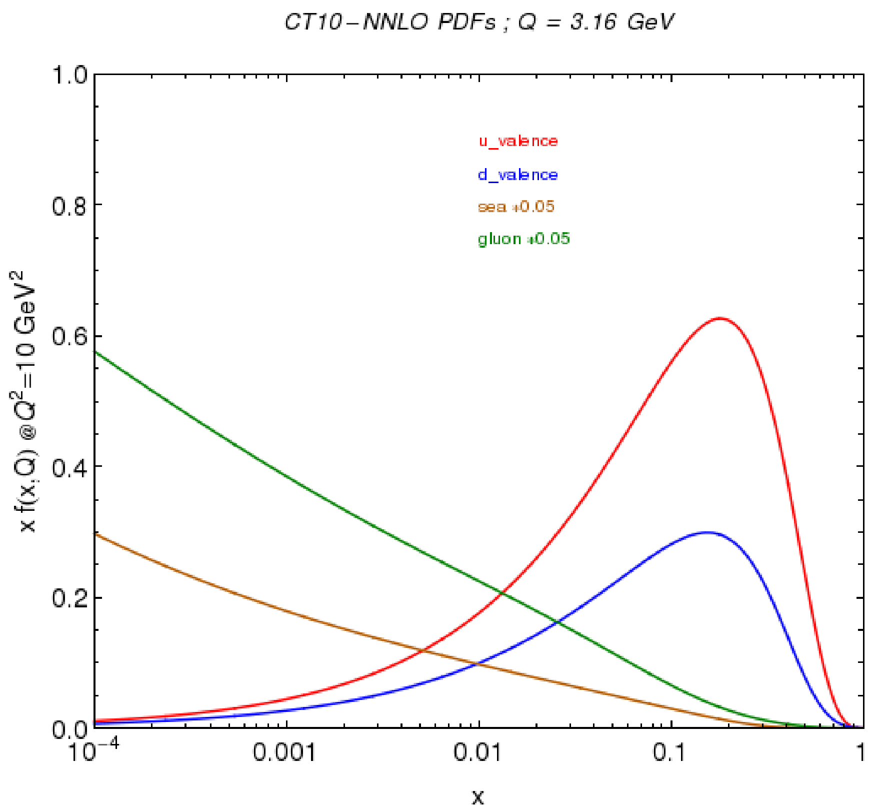
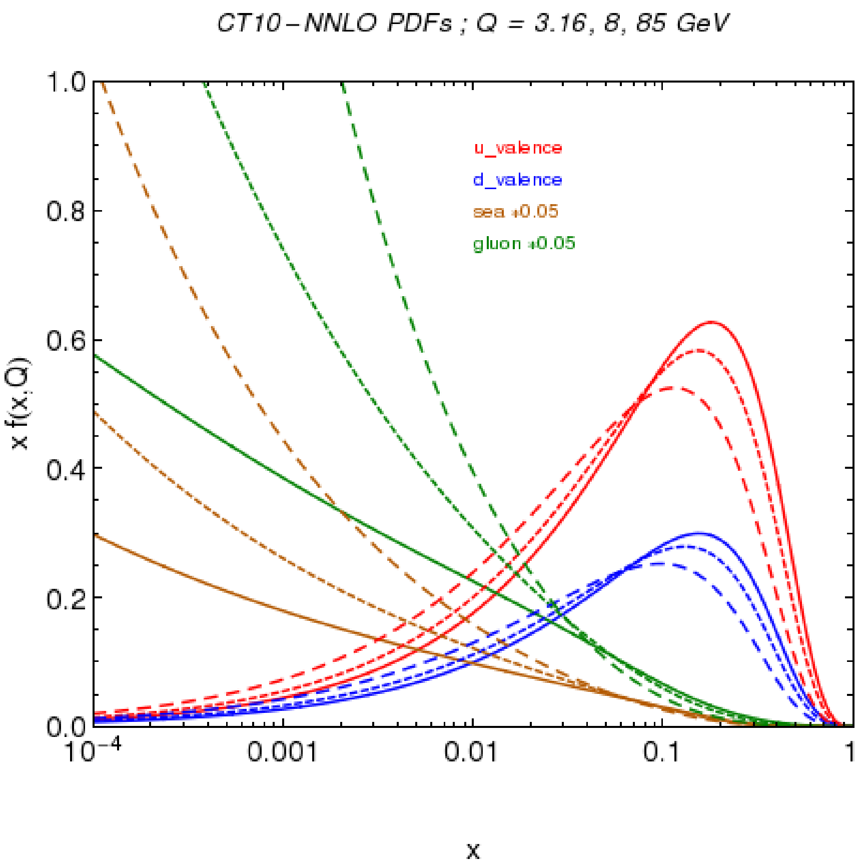
Parton Distribution Functions, Part 2

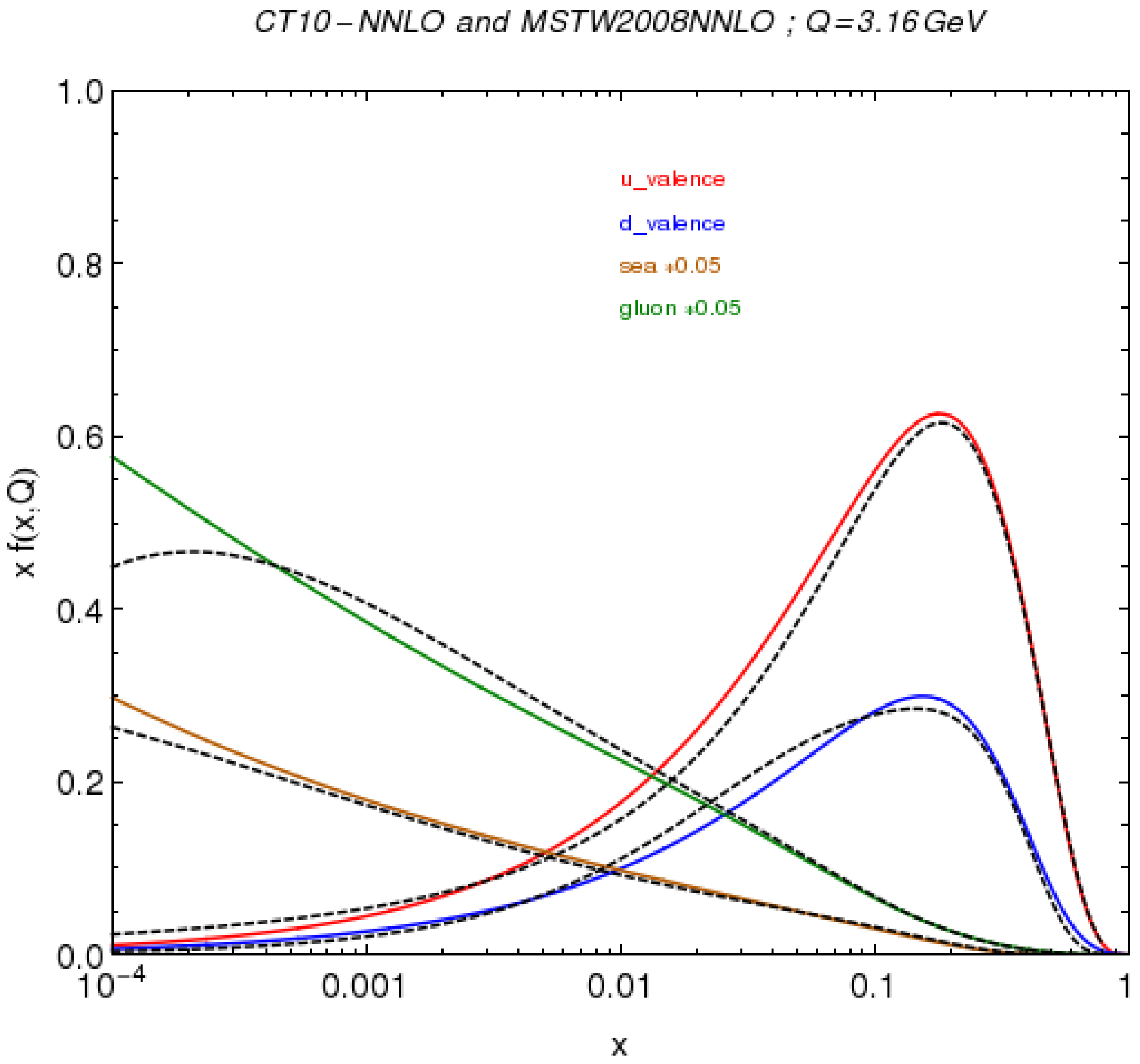


CT10-NNLO Parton Distribution Functions

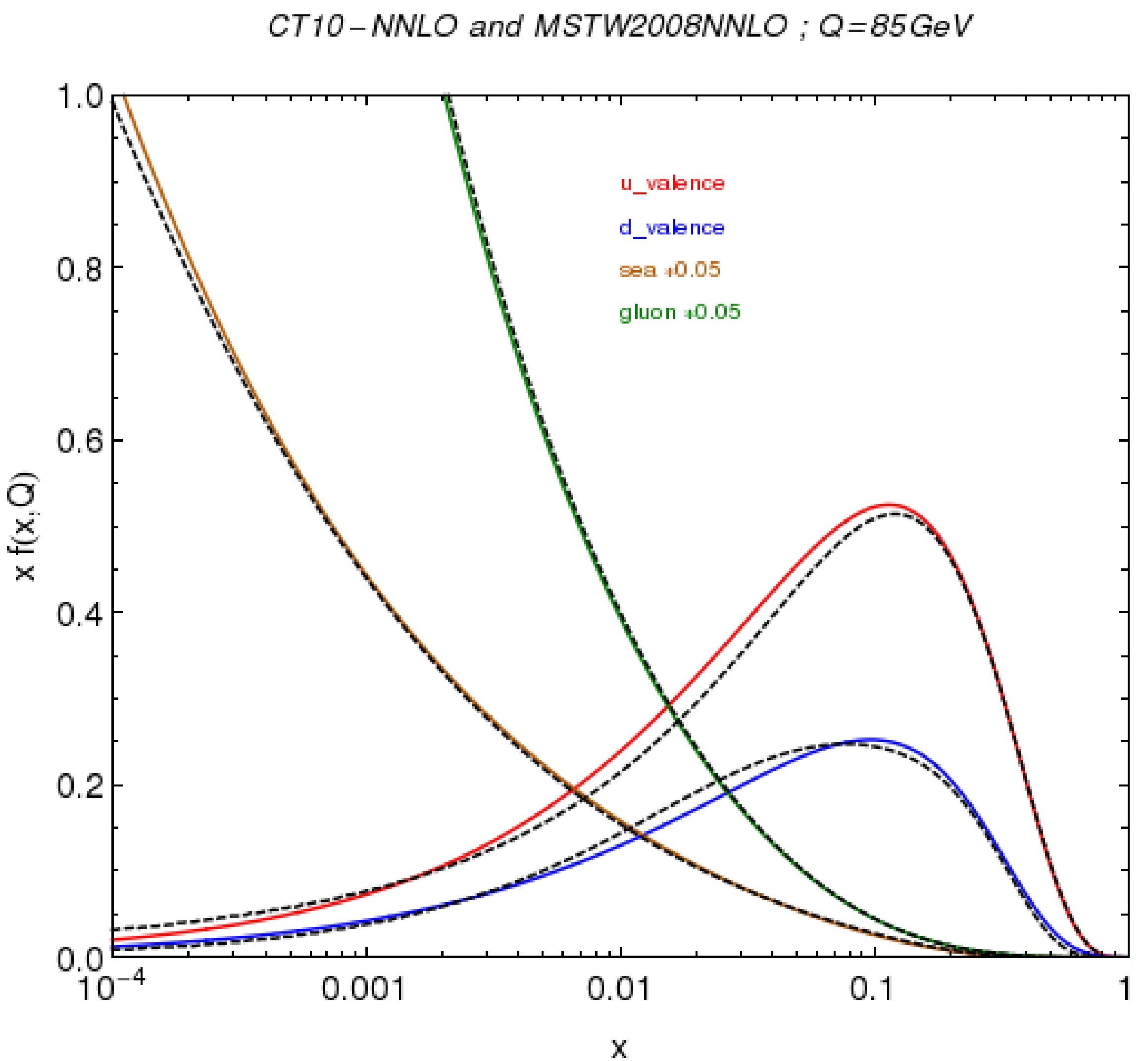




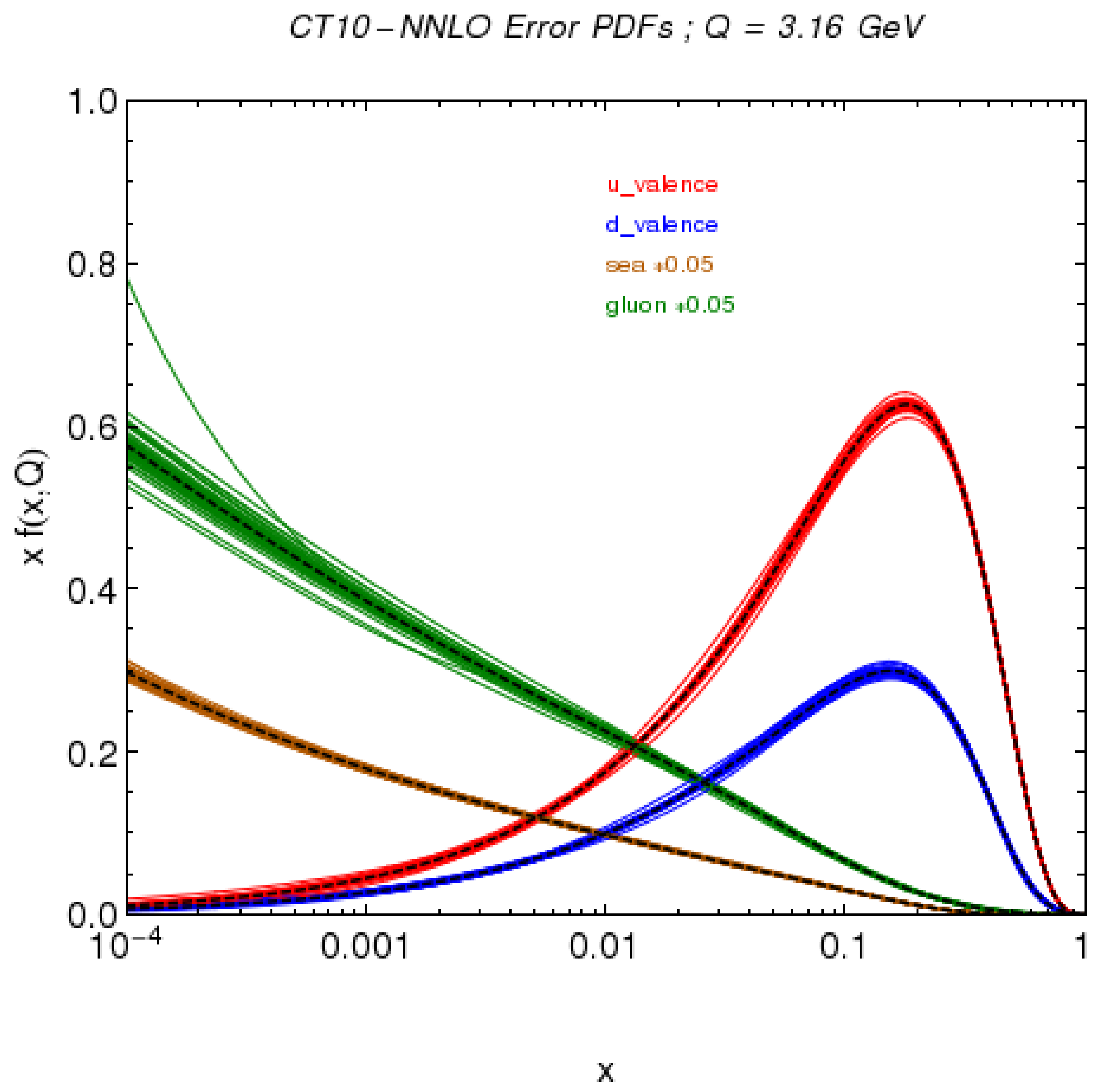
CT10-NNLO compared to MSTW2008 (NNLO)



CT10-NNLO compared to MSTW2008 (NNLO)



CT10-NNLO error pdfs



D. Uncertainties of Parton Distribution Functions

1. "Errors" and Uncertainties

2. Propagation of Experimental "Errors"

3. PDF Uncertainties

4. The role of α_s in Global Analysis

5. Implications for LHC Physics

Lecture 2: Errors and Uncertainties in the Global Analysis of QCD

Mark Twain

… from "Chapters from My Autobiography", published in 1906 …

"Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: ***There are three kinds of lies: lies, damned lies, and statistics."***

"Errors" and Uncertainties

$$\sigma_{ep} = \text{PDF} \otimes C$$

Data ; We're trying to determine this ; Calculation

How accurately can we determine the PDFs?

The accuracy is limited by ...

- Experimental "errors" — statistical; systematic
- Theory "errors" — LO, NLO, NNLO; choice of momentum; value of α_S
- Parametrization errors

Parametrization

In the CTEQ Global Analysis, we parametrize the PDFs ***$f_i(x,Q)$*** at a low Q scale, Q_0 = 1.3 GeV. For example,

$$q_v(x,Q_0) = a_0 x^{a_1} (1-x)^{a_2} \exp\{ a_3 x + a_4 x^2 + a_5 \sqrt{x} \}$$

(q = u or d ; $q_v = q - \bar{q}$)

{ The a_i's are different for different flavors}

Potentially, 6 x (2+2+1+1)=36 parameters.

(CT10-NNLO has 25 parameters.)

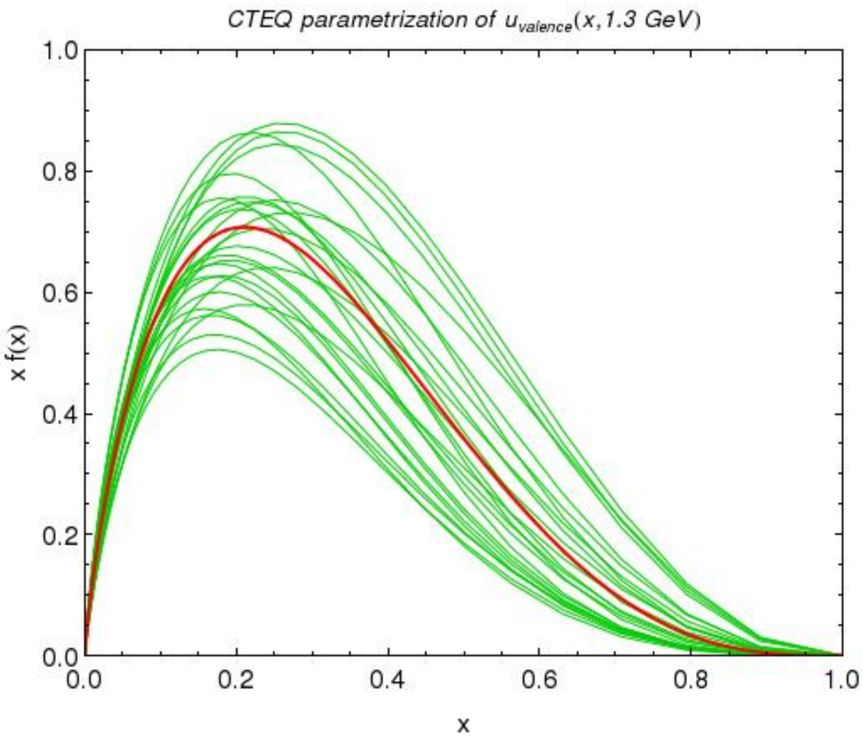
Then ***$f_i(x,Q)$*** is DETERMINED for Q > Q_0 by the renormalization group evolution equations (DGLAP).

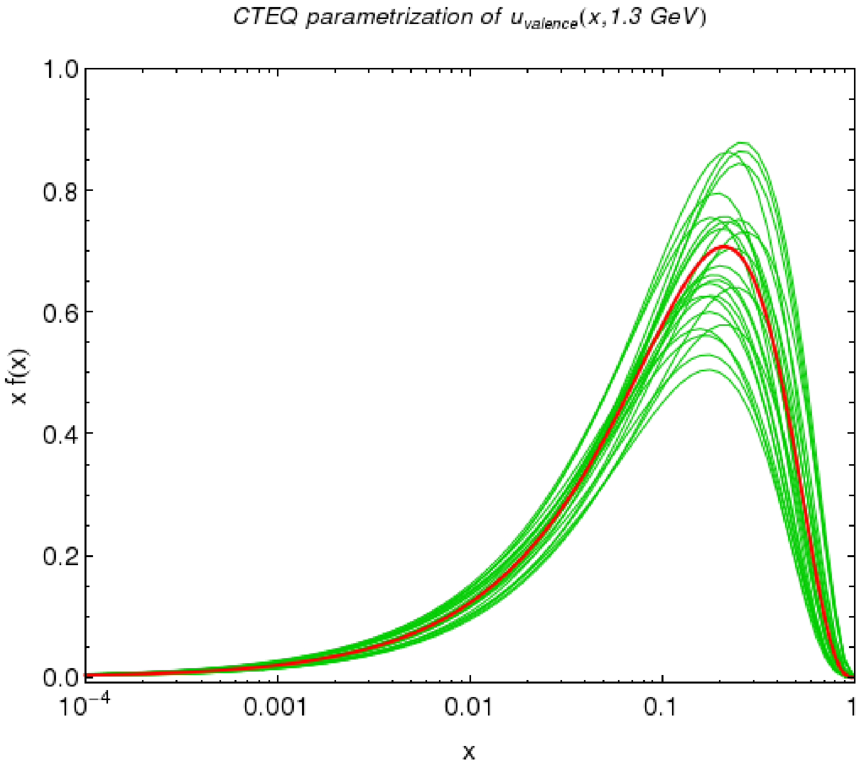
Find the parameter values ***$\{a_{i0} \ldots a_{i5}\}$*** such that theory and data agree most closely.

Parametrization

$$q_v(x,Q_0) = a_0 x^{a_1} (1-x)^{a_2} \exp\{\ a_3 x + a_4 x^2 + a_5 \sqrt{x}\ \}$$

We hope that the parametrization if sufficiently general that the best choice of parameters will fit the "true" function...





25 random choices of the 5 parameters

Propagation of Experimental "Errors"

Consider the data from a single experiment. The experimental collaboration will publish a set of measurements, $\{M_i\,;\, i = 1, 2, 3,\ldots, N_{dp}\}$.

Each measurement M_i has several parts,

$$M_i = \{ D_i\,;\, \sigma_{0i}\,;\, \{\sigma_{\nu i}\,,\ \nu = 1, 2, 3, \ldots, N_{sy}\}\}$$

$\boldsymbol{D_i}$ = central data value; average of many events.

σ_{0i} = correlated errors; 1 standard deviation of the statistical error for D_i.

$\{\boldsymbol{\sigma_{\nu i}}\}$ = correlated errors; $\sigma_{\nu i}$ = 1 standard deviation of the systematic error type ν; N_{sy} = the number of sources of systematic error.

So, $$D_i = X_i + \sigma_{0i} r_{0i} + \sum_{\nu=1}^{N_{sy}} \sigma_{\nu i} r_\nu$$ where X_i = the "true value".

(r_{0i} and r_ν are normalized random variables)

χ^2 analysis

The simplest case is $N_{sy} = 0$ (no systematic errors).
So let's start by assuming $N_{sy} = 0$. (Later we'll include systematic errors.)

Define

$$\chi^2(a) = \sum_{i=1}^{N_{dp}} \frac{[D_i - T_i(a)]^2}{\sigma_{0i}^2}$$

where $T_i(a)$ = the theoretical value for measurement M_i, which depends on a set of theoretical parameters $\{a_1, a_2, a_3, \ldots, a_D\}$.

Minimize $\chi^2(a)$; $\partial\chi^2 / \partial a_k = 0$ for $k = 1, 2, 3, \ldots, D$.

which yields the central fit $\{ a_1^{(0)}, a_2^{(0)}, a_3^{(0)}, \ldots, a_D^{(0)} \}$

The variations of χ^2 around the minimum tells us the uncertainty $\{ \delta a_1, \delta a_2, \delta a_3, \ldots, \delta a_D \}$

Hessian Analysis

In a small neighborhood of $\{a^{(0)}\}$,

$$\chi^2(a) = \chi^2(a^{(0)}) + \sum_{j,k=1}^{N_{sy}} \tfrac{1}{2} \left[\frac{\partial^2 \chi^2}{\partial a_j \partial a_k}\right]_0 (a_j - a_j^{(0)})(a_k - a_k^{(0)})$$

the Hessian matrix, H_{jk}

Ludwig Otto Hesse

Now, make a prediction based on the theory, for some ***other*** quantity Q:

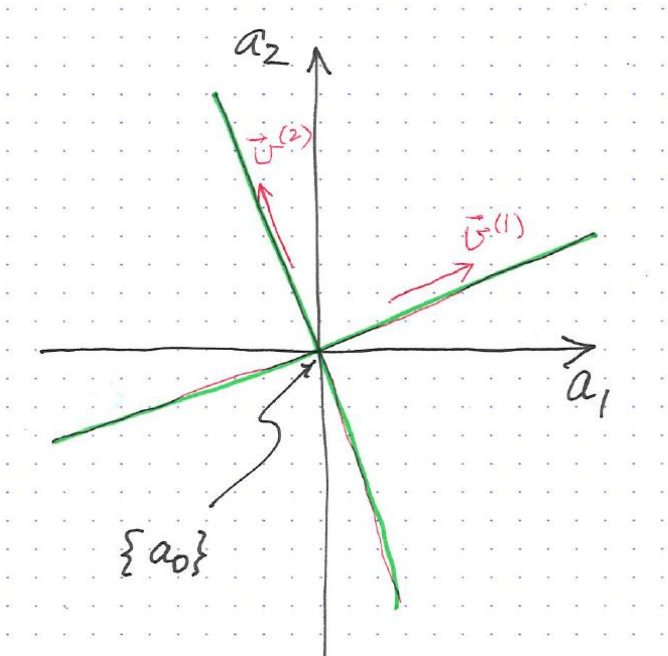
$$Q(a) = Q(a^{(0)}) + \sum_{j=1}^{D} \frac{\partial Q}{\partial a_j} (a_j - a_j^{(0)}) + \ldots$$

$$Q_{\text{prediction}} = Q(a^{(0)}) \pm \delta Q$$

$$(\delta Q)^2 = \sum_{j,k=1}^{D} (a_j - a_j^{(0)})(a_k - a_k^{(0)}) \frac{\partial Q}{da_j} \frac{\partial Q}{da_k}$$

the allowed variation, δa_j

Parameter variations in the ***eigenvector basis***



Eigenvectors of the Hessian,

$$\sum_{k=1}^{D} H_{jk} v_k^{(m)} = \lambda^{(m)} v_j^{(m)}$$

(m = 1, 2, 3, ..., D)

The distance along the eigenvector $\boldsymbol{v}^{(m)}$ is allowed to vary by an amount of order $\pm(1/\lambda^{(m)})^{1/2}$. ***("tolerance")***

Then we have D × 2 alternative fits of the theory parameters, and one central fit.

Systematic errors

We have

$$D_i = X_i + \sum_{\nu=1}^{N_{sy}} \sigma_{\nu i} r_\nu + \sigma_{0i} r_{0i}$$

$\sigma_{\nu i}$ = standard deviation of systematic error ν, of D_i

σ_{0i} = standard deviation of the uncorrelated error, of D_i

(r_ν and r_{0i} are unknown random variables, of order 1)

We want $\{T_i(a)\}$ to fit $\{X_i\}$ as closely as possible, from the knowledge we have.

$$\chi^2(a;\{r_\nu\}) = \sum_{i=1}^{N_{dp}} \frac{[D_i - \sum_{\nu=1}^{N_{sy}} \sigma_{\nu i} - T_i(a)]^2}{\sigma_{oi}^2} + \sum_{\nu=1}^{N_{sy}} r_\nu^2$$

Minimize this χ^2 with respect to both $\{a_k\}$ and $\{r_\nu\}$.

I.e., we fit the theory $T_i(a)$ to "optimally shifted data",

$$\text{shifted data} = D_i - \sum_{\nu=1}^{N_{sy}} \sigma_{\nu i} \bar{r}_\nu \quad ; \quad \bar{r}_\nu = \text{optimal normalized shift}$$

Results of the Global Analysis of QCD

$\{ f_i^{(0)} (x,Q^2) \}$ = the central fit (i = 0,±1,±2,±3, ±4, ±5)

$\{ f_i^{(k)} (x,Q^2) \}$ = the error PDFs; alternative fits
(k = 1, 2, 3, … ,50)

k = 1,2	Eigenvector 1	±directions
k = 3,4	Eigenvector 2	±directions
	…	
k = 2n-1,2n	Eigenvector n	±directions
	…	

… accessible at the CT10-NNLO web site,
or the Durham Parton Distribution Generator

PDF Uncertainty

Consider the prediction of a quantity *Q*. Ideally,

$$(\delta Q)^2 = \sum_{n=1}^{25} [Q(a_{2n-1}) - Q(a_0)]^2$$

$$(\delta Q)^2 = \sum_{n=1}^{25} [Q(a_{2n}) - Q(a_0)]^2$$

$$(\delta Q)^2 = \sum_{n=1}^{25} [(Q(a_{2n}) - Q(a_{2n-1})) / 2]^2$$

symmetric errors

This is the "Master Formula" for symmetric errors. But the behavior in the neighborhood of the minimum is not perfectly quadratic. So, instead,

$$(\delta^+ Q)^2 = \sum_{n=1}^{25} [\max(Q_n^+ - Q_0, Q_n^- - Q_0, 0)]^2$$

$$(\delta^- Q)^2 = \sum_{n=1}^{25} [\max(Q_0 - Q_n^+, Q_0 - Q_n^-, 0)]^2$$

asymmetric errors

Democracy among Experiments

$$\chi^2_{global}(a) = \sum_{\text{experiments}} \{ \chi^2_{\text{expt.}}(a) + \mathbf{P}_{\text{expt.}} \}$$

… from the previous page, including the correlated shifts $\{r_\nu\}_{\text{expt}}$.

$\mathbf{P}$ is a penalty that prevents ***its*** experiment from deviating too much (90% C.L.) from the theory.

- Minimize χ^2_{global}.
- Diagonalize the ***full*** Hessian Matrix.
- Final Result: $f_i^{(0)}(x,Q^2)$ and 50 variations $\{f_i^{(k)}(x,Q^2)\}$

…the ***eigenvector-basis PDF variations.***

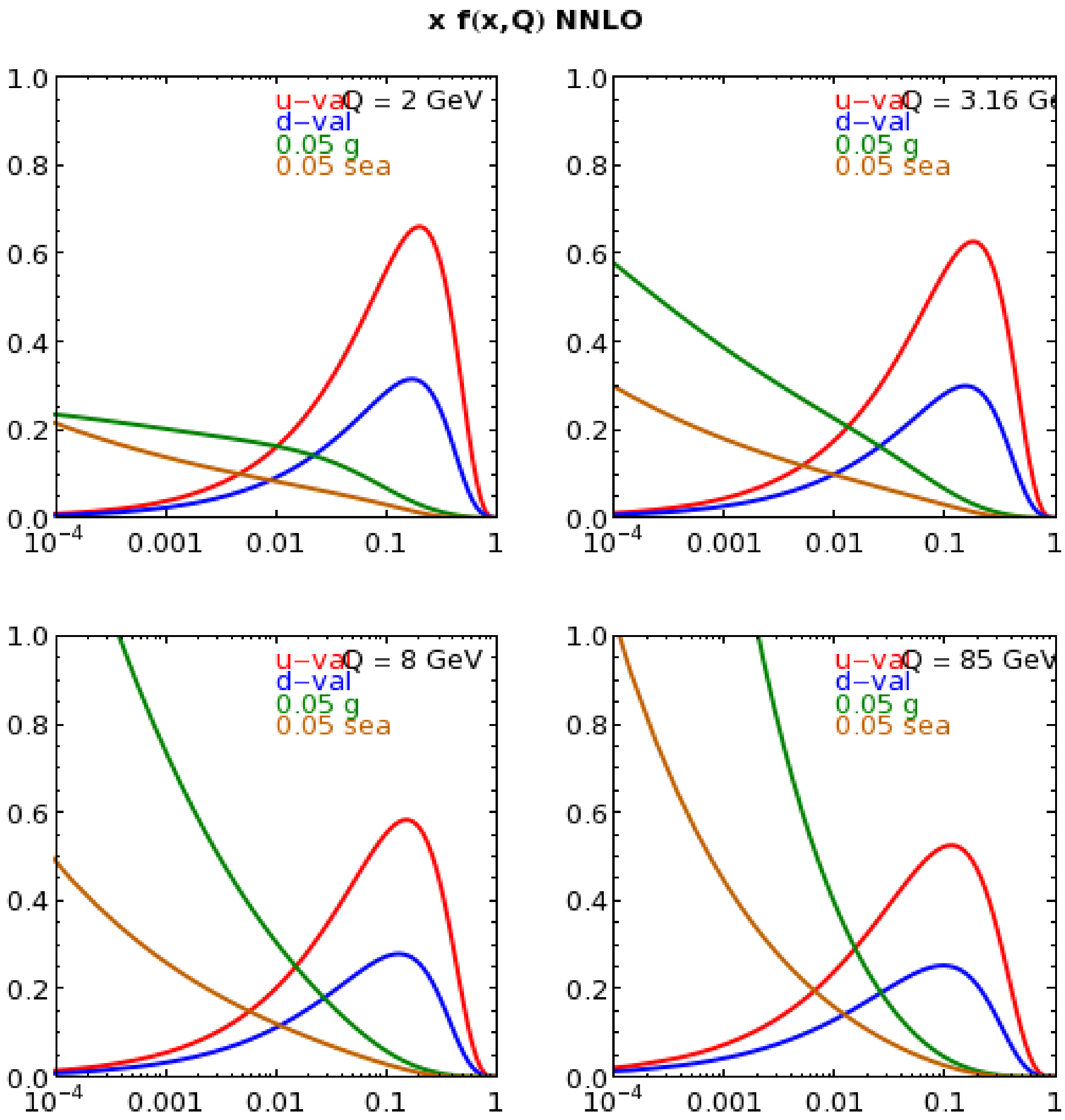
Egalitarianism among Experiments

Experimental Data Used in the CT10 Analysis

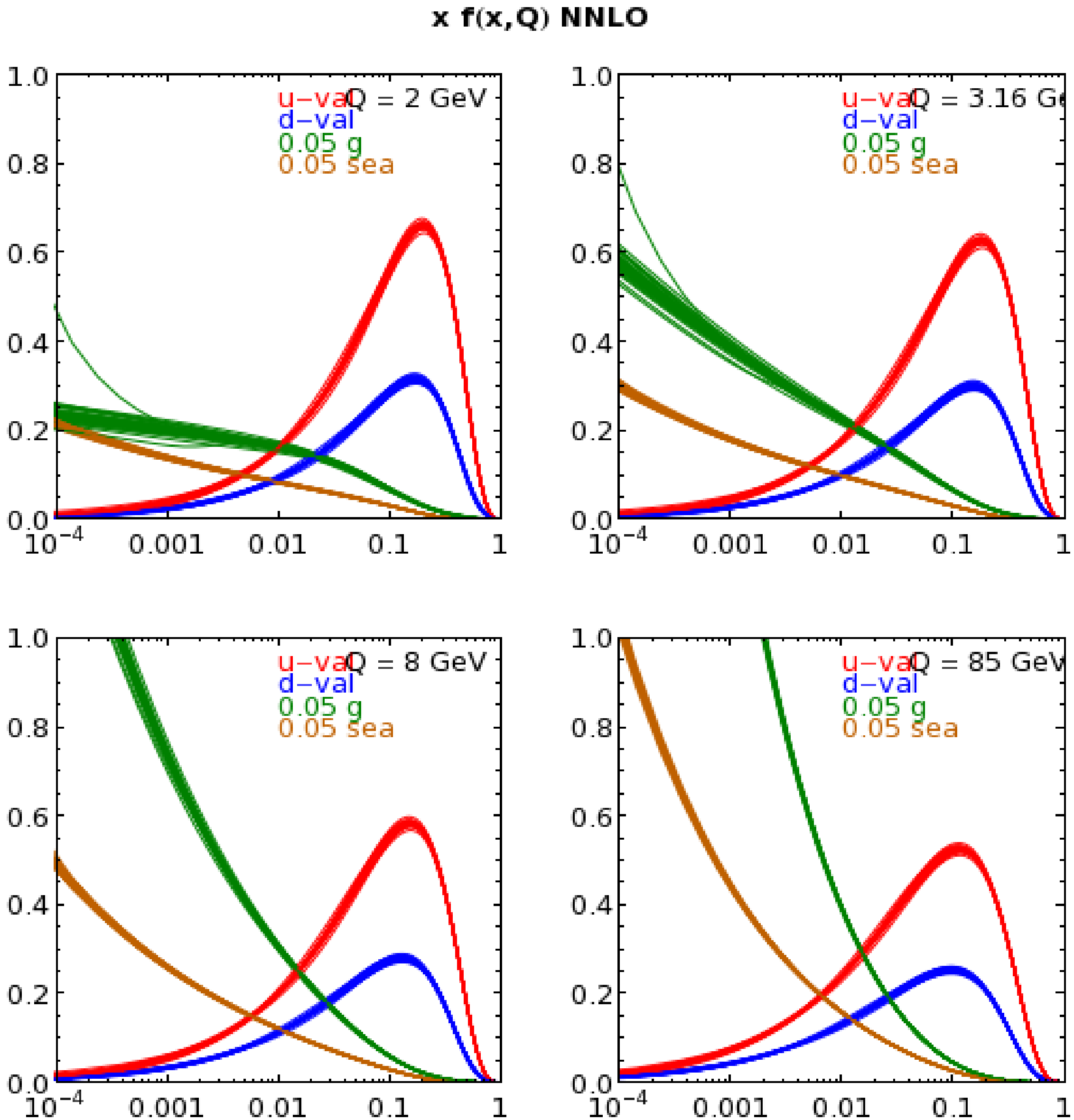
CT10-NNLO Table		Ndp	Chi^2	Nsy	
1/ 159	HERA1X0	579	617.	114	Combined HERA1 NC+CC DIS (2009)
2/ 101	BcdF2pCor	339	392.	5	BCDMS collaboration
3/ 102	BcdF2dCor	251	291.	5	BCDMS collaboration
4/ 103	NmcF2pCor	201	333.	11	NMC collaboration
5/ 104	NmcRatCor	123	151.	5	NMC collaboration
6/ 108	cdhswf2	85	70.5	0	P Berge et al Z Phys C49 187 (1991)
7/ 109	cdhswf3	96	77.9	0	P Berge et al Z Phys C49 187 (1991)
8/ 110	ccfrf2.mi	69	67.8	5	Yang&Bodek model-independent
9/ 111	ccfrf3.md	86	34.8	0	Shaevitz&Seligman model-dependent processed by SK
10/ 201	e605	119	95.7	0	DY Q^3 dSig/dQ dy proton on heavy target
11/ 203	e866f	15	9.7	0	E866 experiment: pd / 2pp
12/ 225	cdfLasy	11	13.4		W production: decay lepton asymmetry CDF Run-1
13/ 140	HN+67F2c	8	9.3	0	H1 neutral current charm
14/ 143	HN+90X0c	10	16.3	8	H1 neutral current charm
15/ 156	ZN+67F2c	18	13.4	0	ZEUS neutral current charm
16/ 157	ZN+80F2c	27	16.7	0	ZEUS neutral current charm
17/ 124	NuTvNuChXN	38	29.6	0	NuTev Neutrino Dimuon Reduced xSec
18/ 125	NuTvNbChXN	33	28.4	0	NuTev Neutrino Dimuon Reduced xSec
19/ 126	CcfrNuChXN	40	48.0	0	Ccfr Neutrino Dimuon Reduced xSec
20/ 127	CcfrNbChXN	38	26.4	0	Ccfr Neutrino Dimuon Reduced xSec
21/ 204	e866ppxf	184	234.	0	E866 experiment: DY pp: Q^3 dSig/dQ dxf
22/ 260	ZyD02a	28	15.6	6	Z rapidity dist. (D0 TeV II-a)
23/ 261	ZyCDF2	29	46.5	6	Z rapidity dist. (CDF TeV II)
24/ 227	cdfLasy2	11	11.4	0	W production: decay lepton asymmetry CDF Run-2
25/ 231	d02Easy1	12	26.0	0	W production: decay elec asymmetry D0 Run-2 Pt>25
26/ 234	d02Masy1	9	14.8	0	W production: decay muon asymmetry D0 Run-2 Pt>20
27/ 504	cdf2jtCor2	72	101.	24	(run II: cor.err; ptmin & ptmax)
28/ 514	d02jtCor2	110	114.	23	(run II: cor.err; ptmin & ptmax)

next ……

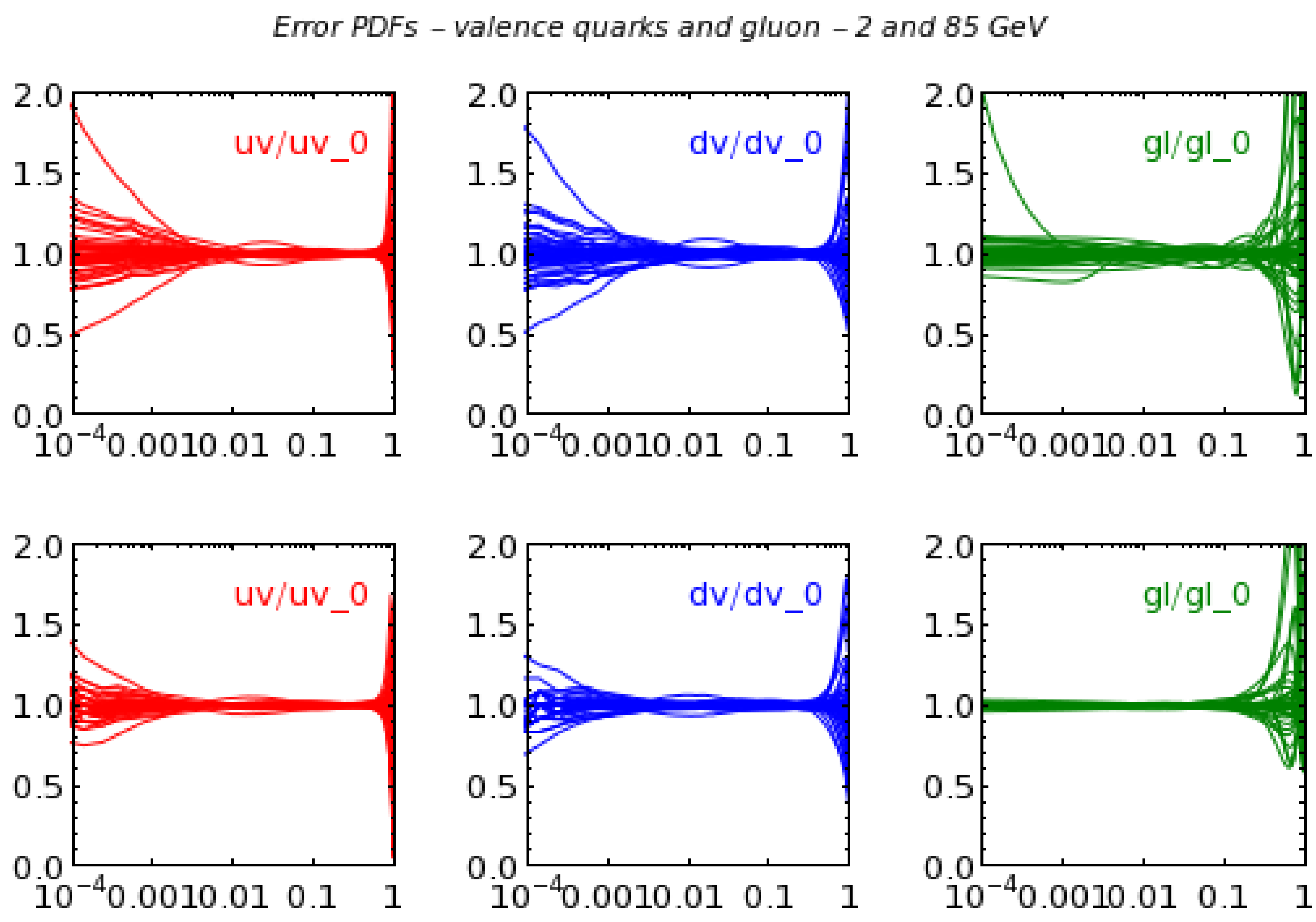
Uncertainties of Parton Distribution Functions



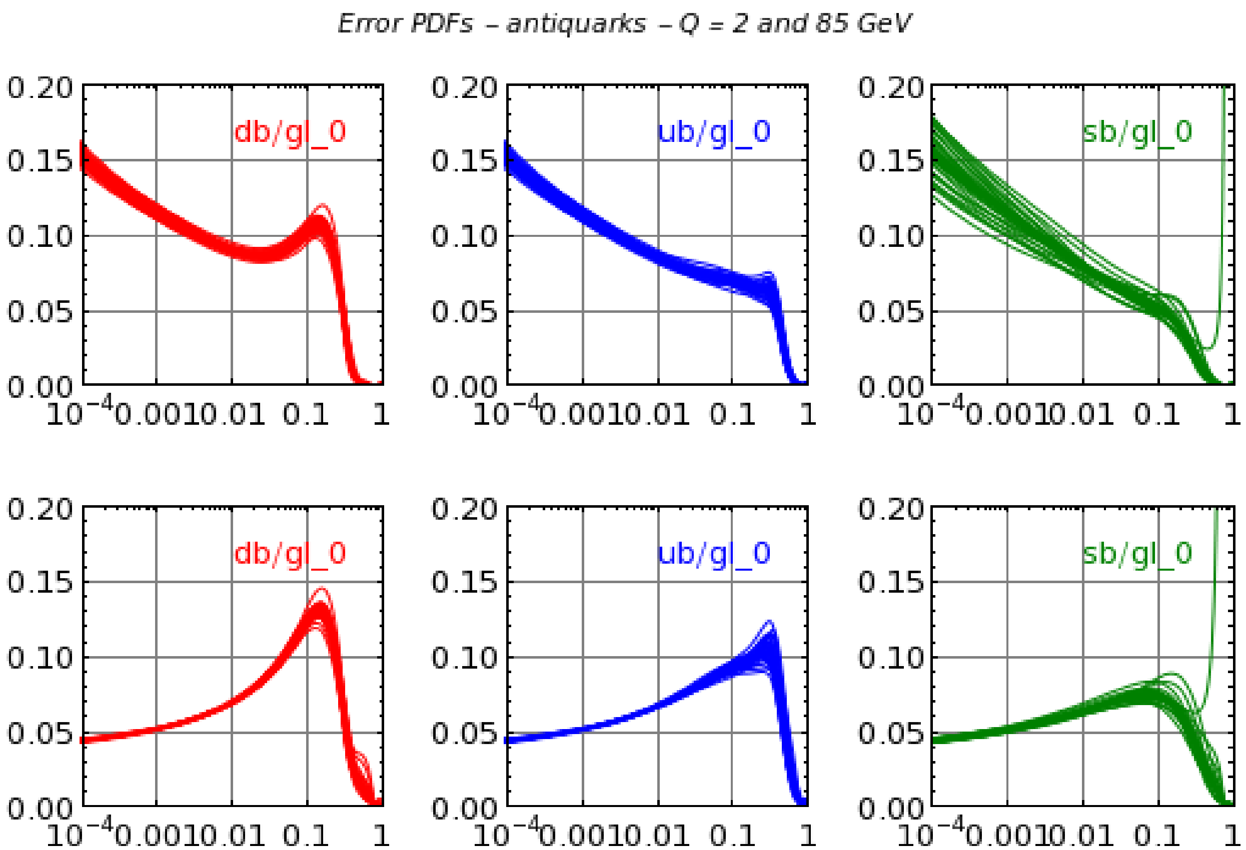
CT10-NNLO Error PDFs



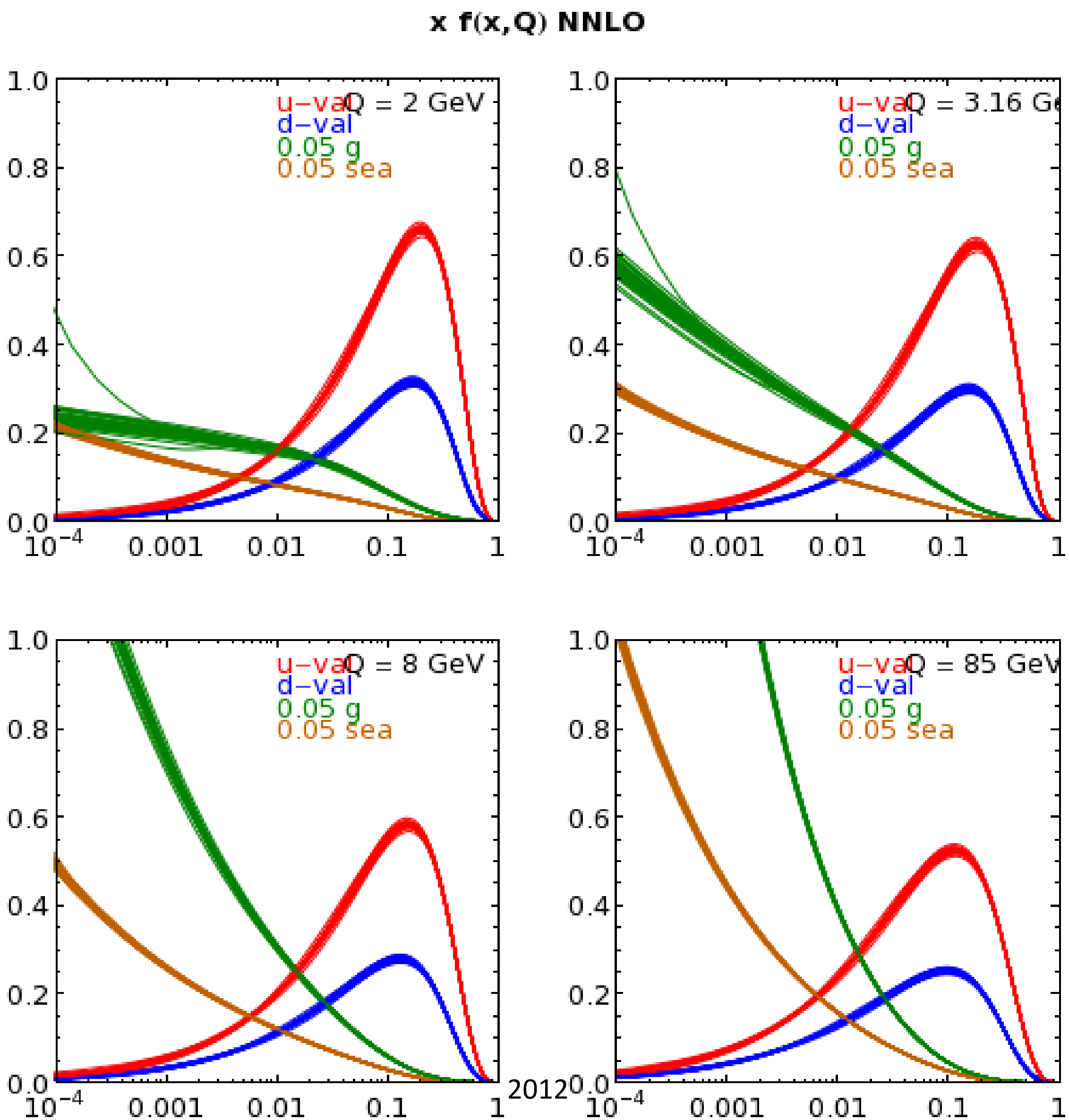
CT10-NNLO Error PDFs
-- ratio plots for valence quarks and gluon



CT10-NNLO Error PDFs
-- ratio plots : antiquarks / gluon



E. Applications to LHC Physics



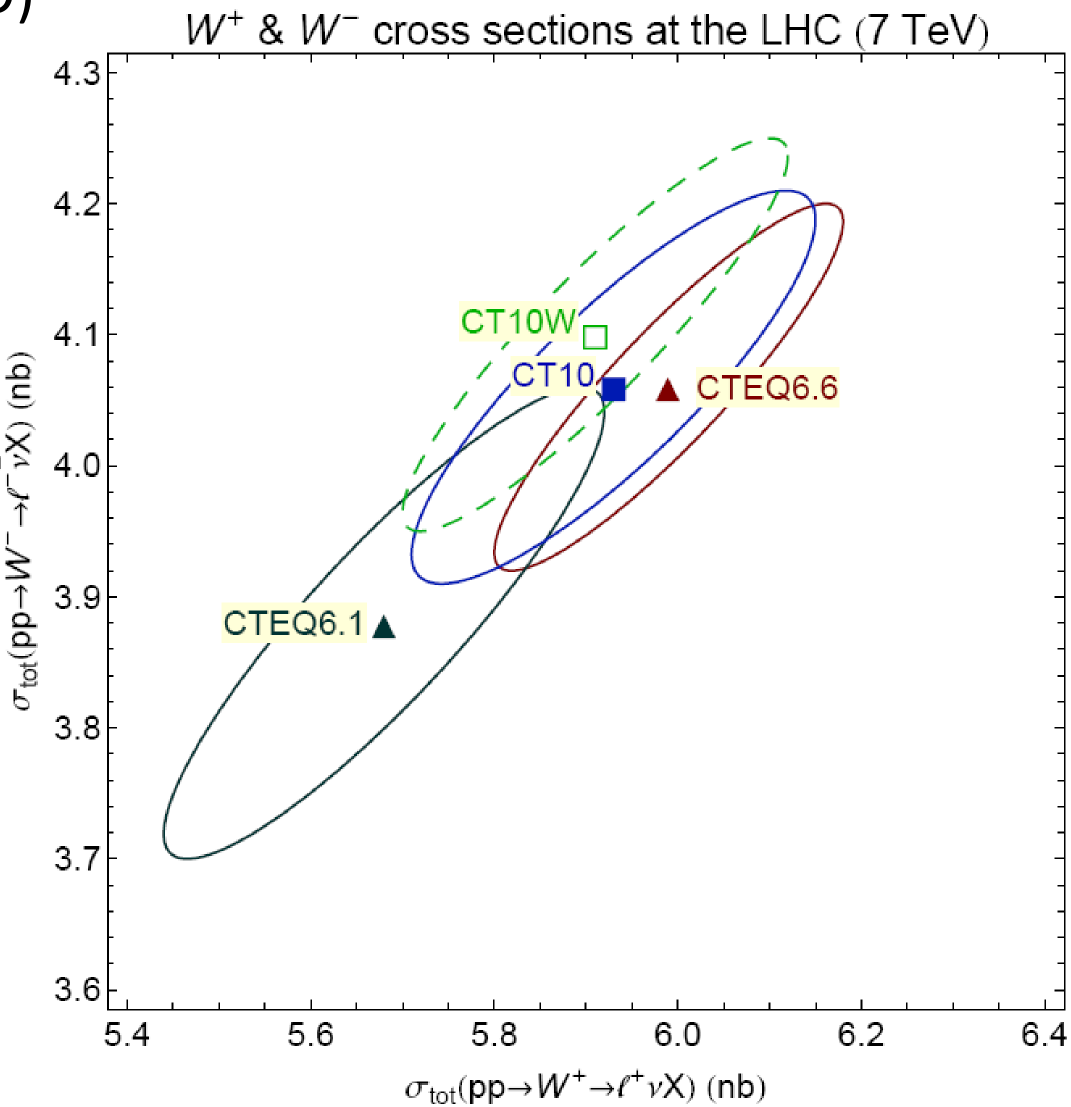
LHC predictions

- The search for New Physics will require high precision comparisons between standard model predictions and experimental measurements.
- We seek discrepancies between standard model theory and data.
- The CT10-NNLO PDFs – central fit and eigenvector basis variations – will be used in the theory predictions.
- Please understand the importance of the Master Formula!

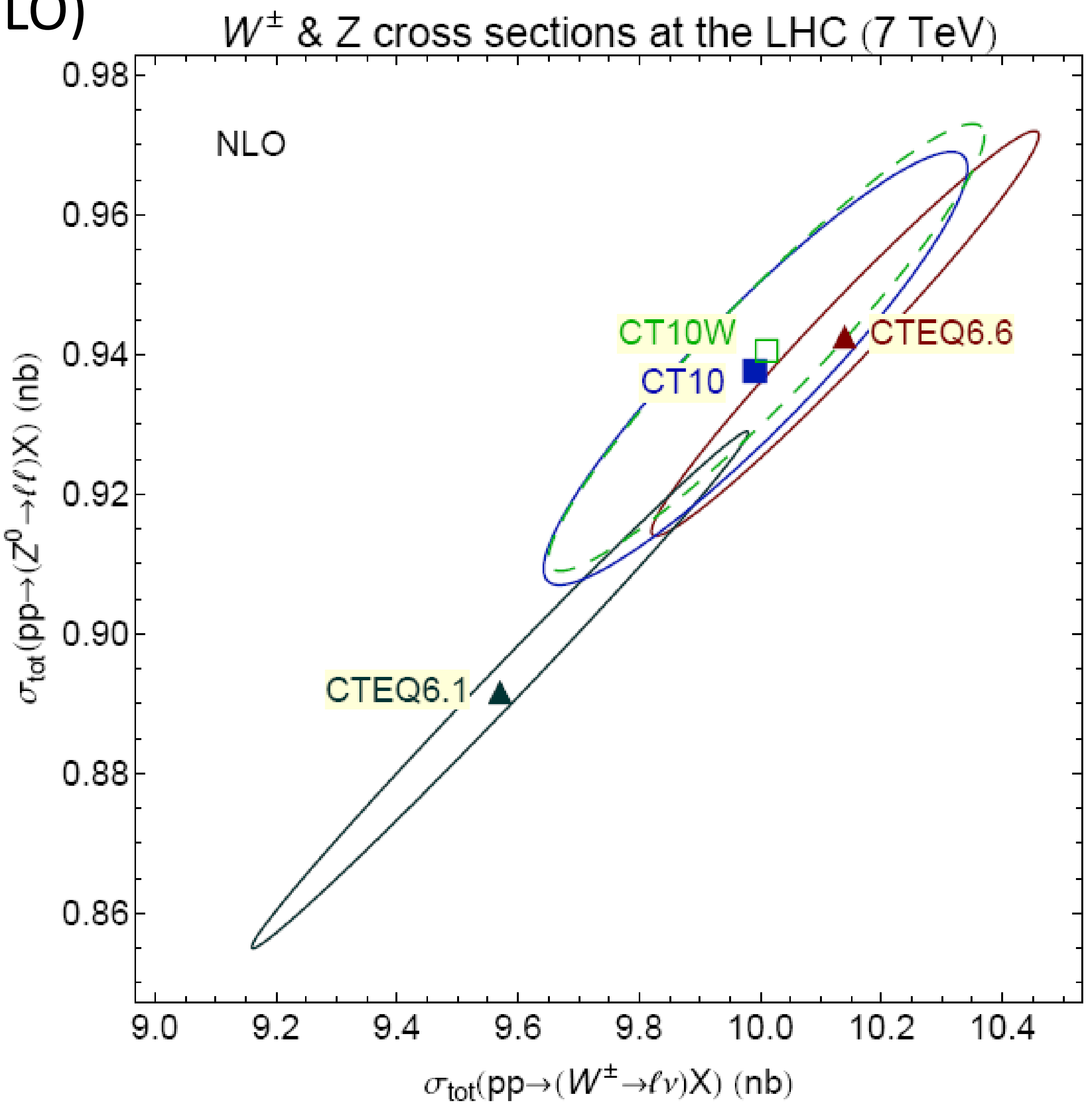
$$Q_{central} = Q(0)$$

$$(\Delta Q)^2 = \tfrac{1}{4}\sum_{n=1}^{D}[Q(2n-1) - Q(2n)]^2$$

… from the CT10 paper(NLO)



… from the CT10 paper(NLO)

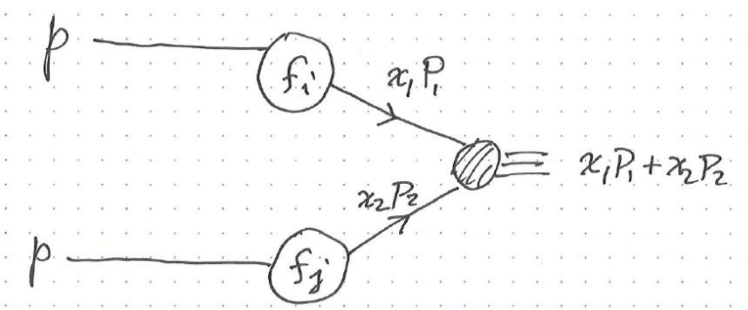


Total cross sections for production of electroweak bosons
(Dr. Zhao Li)

Boson	Collider	unit	CT10 (NLO)	CT10-NNLO	MSTW2008
W+	LHC14	nb	12.2 ± 0.5	12.7 ± 0.5	12.4 ± 0.2
W+	LHC7	nb	6.0 ± 0.2	6.3 ± 0.2	6.2 ± 0.1
W+	Tevatron	nb	1.35 ± 0.05	1.38 ± 0.05	1.38 ± 0.02
W-	LHC14	nb	8.9 ± 0.4	9.4 ± 0.4	9.3 ± 0.2
W-	LHC7	nb	4.10 ± 0.15	4.29 ± 0.16	4.31 ± 0.07
Z	LHC14	nb	2.07 ± 0.08	2.17 ± 0.08	2.13 ± 0.03
Z	LHC7	nb	0.96 ± 0.03	1.00 ± 0.03	0.99 ± 0.02
Z	Tevatron	pb	260 ± 9	263 ± 8	261 ± 5
H(sm)	LHC14	pb	101 ± 9	99 ± 8	102 ± 7
H(sm)	LHC7	pb	31.2 ± 1.9	29.7 ± 1.7	29.8 ± 1.3
H(sm)	Tevatron	pb	1.77 ± 0.12	1.77 ± 0.12	1.80 ± 0.11

Luminosity Functions

Consider production of a high-mass state at the LHC.



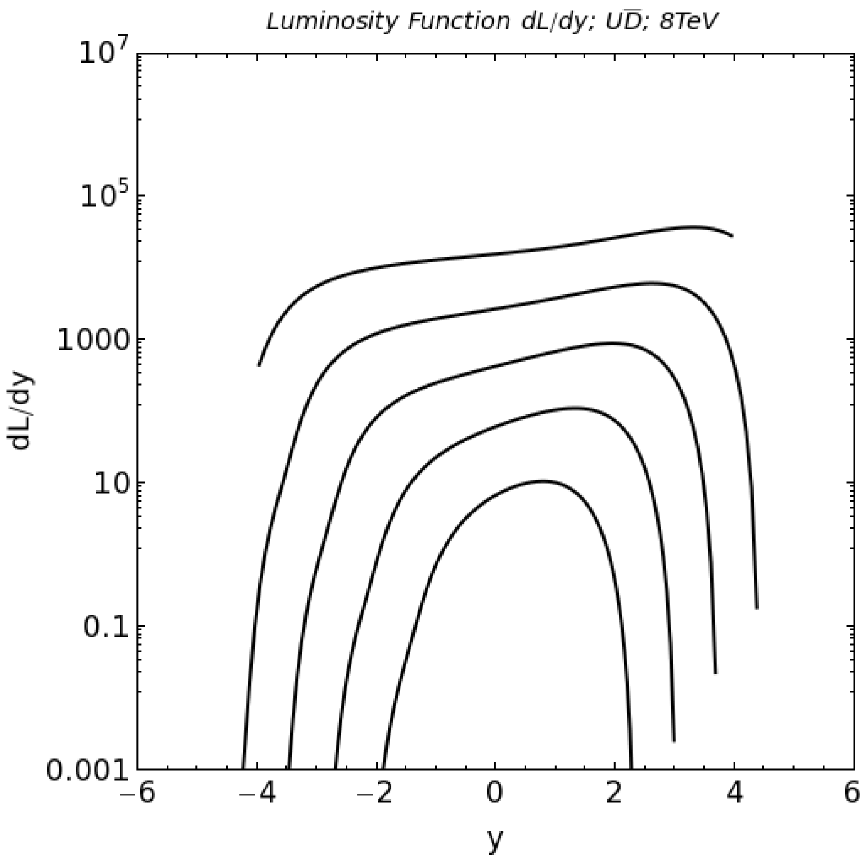
$$M = (x_1 P_1 + x_2 P_2)^2 \approx 2 x_1 x_2 P_1 \cdot P_2$$
$$\approx x_1 x_2 s \quad \text{where } s = (P_1 + P_2)^2.$$

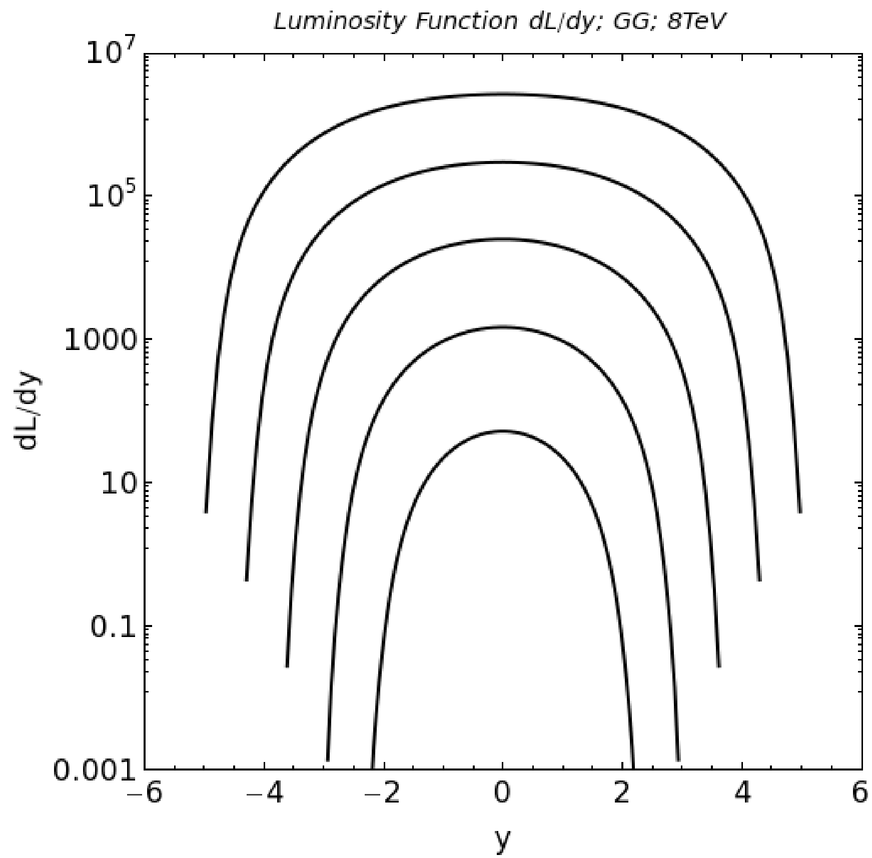
$$L_{ij}(M) = \int_0^1 dx_1 \int_0^1 dx_2\, f_i(x_1, M)\, f_j(x_2, M)\, \delta(x_1 x_2 - M^2/s)$$

$$x_1 = e^y\, \hat{M}/\sqrt{s} \quad \text{and} \quad x_2 = e^{-y}\, \hat{M}/\sqrt{s}$$

$$L_{ij}(M) = \int_{-y_M}^{y_M} dy\, f_i(e^y\, M/\sqrt{s}, M)\ f_j(e^{-y}\, \hat{M}/\sqrt{s}, M)$$

$$y_M = \ln(\sqrt{s}/M)$$





M = 50, 100, 200, 400, 800 GeV

The luminosity function is a measure of the rate of production --- the contribution of PDFs to that rate --- for states with the corresponding quantum numbers. The uncertainty of the L.F. is a good estimate of the PDF uncertainty for that process.