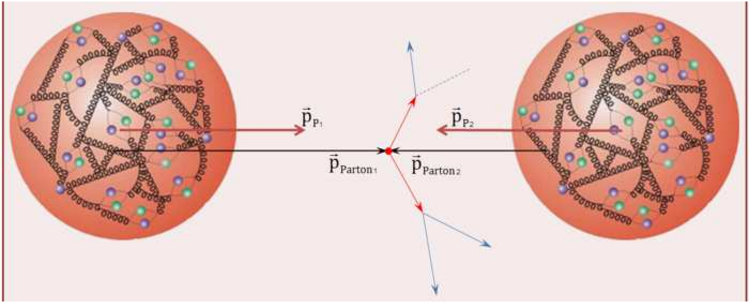
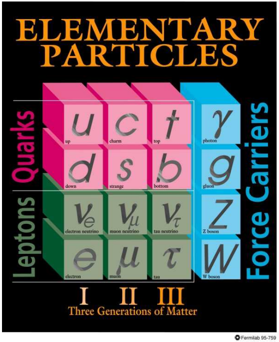
Recent developments in PDF

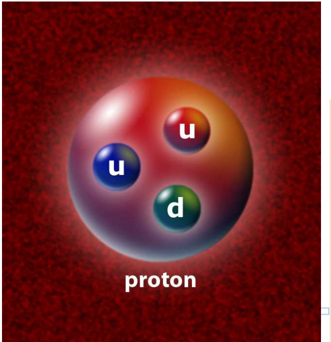
Tie-Jiun Hou
Northeastern University

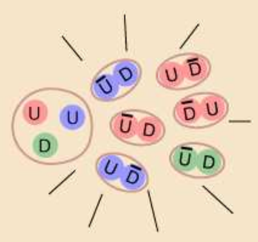
July 24, 2019
QCD summer school at Harbin

Proton is constructed by quark and gluon

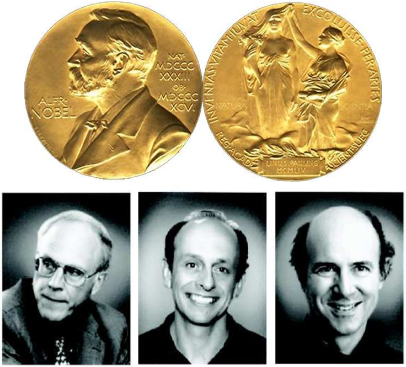




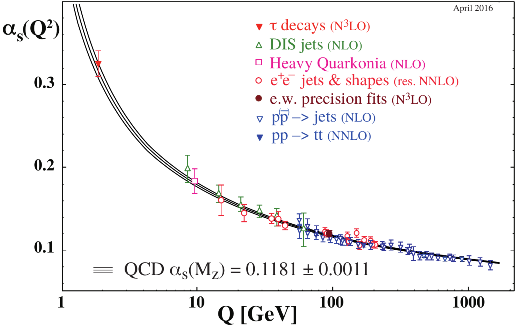




Nobel Prize 2004:
Asympotic freedom in strong interaction

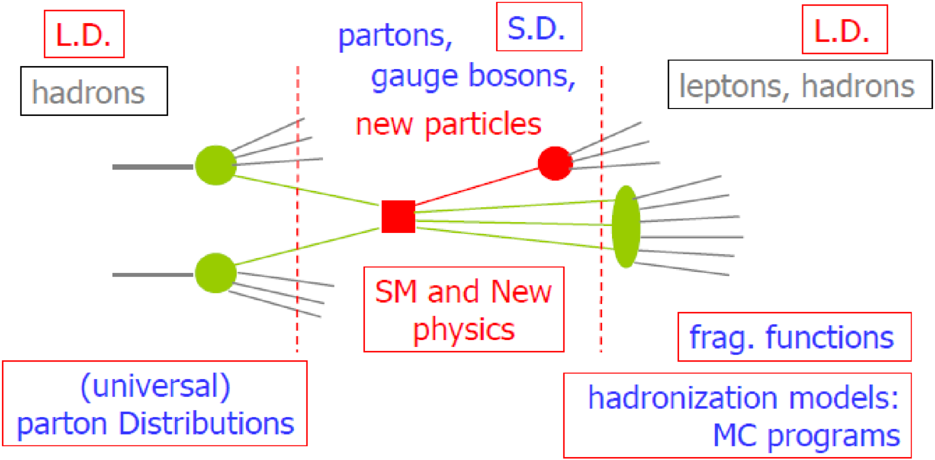


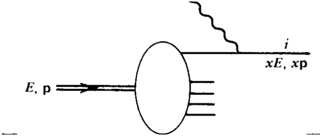
David J. Gross, H. David Politzer and Frank Wilczek



Short distance $\rightarrow$ perturbative

Long distance $\rightarrow$ non-perturbative

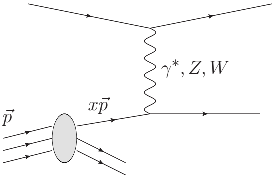




Parton distribution function (PDF) $f_{j/A}(x, Q)$ discribe the possibility to find a parton j, i.e. quark and gluon, in a nucleaon A.

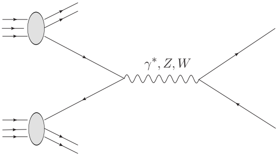
Factorize the long distance into PDF $f(x,Q)$

DIS process:



$$\sigma = f_{i/p}(x, \mu_f, \{a\}) \otimes \hat{\sigma}_i(\mu_f)$$

Drell-Yan process:

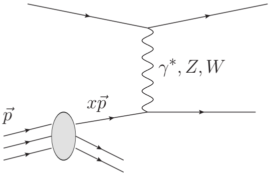


$$\sigma = f_{i/p_1}(x, \mu_f, \{a\}) \otimes f_{j/p_2}(x, \mu_f, \{a\}) \otimes \hat{\sigma}_{ij}(\mu_f)$$

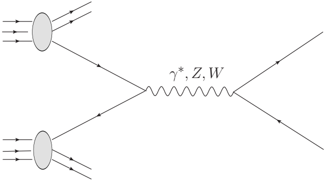
Long distance and short distance physics are factorized by $\mu_f \gg \Lambda_{QCD}$.

PDF $f(x,Q)$ is universal

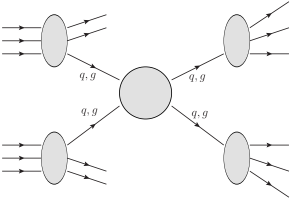
DIS



Drell-Yan

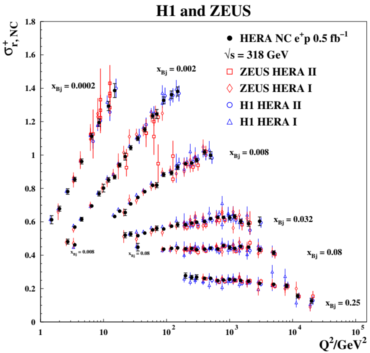


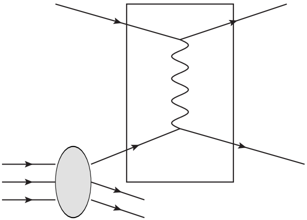
jet



PDF is determined by comparing data and hard cross section

$$\sigma = f(x, Q^2, \{a\}) \otimes \hat{\sigma}$$





PDF	Hard part
LO	LO
NLO	NLO
NNLO	NNLO

PDF evolve

$$\frac{\partial q_i(x,\mu^2)}{\partial \ln \mu^2} = P^v_{qq} \otimes q_i + P^v_{q\bar{q}} \otimes \bar{q}_{\bar{i}} + P^s_{qq} \otimes \sum_k^{N_f} q_k + P^s_{q\bar{q}} \otimes \sum_k^{N_f} \bar{q}_{\bar{k}} + P_{qg} \otimes g$$

$$\frac{\partial \bar{q}_{\bar{i}}(x,\mu^2)}{\partial \ln \mu^2} = P^v_{q\bar{q}} \otimes q_i + P^v_{qq} \otimes \bar{q}_{\bar{i}} + P^s_{q\bar{q}} \otimes \sum_k^{N_f} q_k + P^s_{qq} \otimes \sum_k^{N_f} \bar{q}_{\bar{k}} + P_{qg} \otimes g$$

$$\frac{\partial g(x,\mu^2)}{\partial \ln \mu^2} = P_{gq} \otimes \sum_k^{N_f} (q_k + \bar{q}_k) + P_{gg} \otimes g$$

DGLAP equations tell us how the PDFs evolve from low energy scale, the input energy scale, to high energy scale, the energy scale of interaction. But it does not tell us it's x-dependency. PDFs $f(x, Q_0)$ at input energy scale Q_0 is determined by Data.

PDF has order

$$\sigma = f_{i/p}(x, \mu_f, \{a\}) \otimes \hat{\sigma}_i(\mu_f)$$

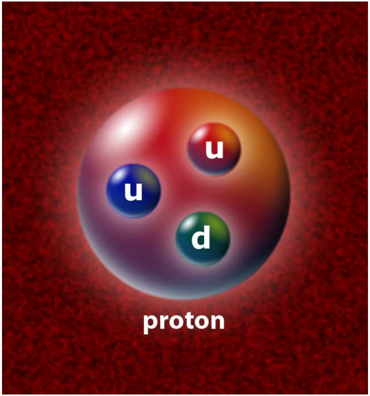
$$\hat{\sigma}(x, \alpha_s) = \alpha_s \hat{\sigma}^{(1)}(x) + \alpha_s^2 \hat{\sigma}^{(2)}(x) + \alpha_s^3 \hat{\sigma}^{(3)}(x) + \ldots$$

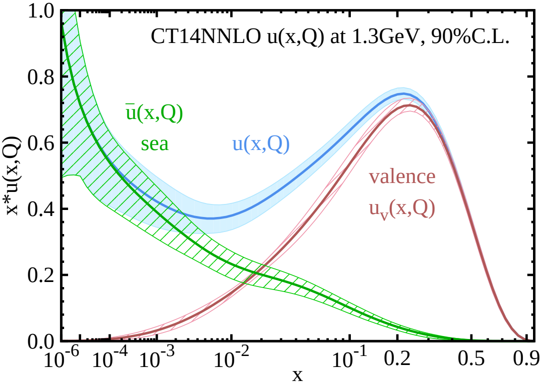
$$\frac{\partial f(x, \mu^2)}{\partial \ln \mu^2} = P(x) \otimes f(x)$$

$$P(x, \alpha_s) = \alpha_s P^{(1)}(x) + \alpha_s^2 P^{(2)}(x) + \alpha_s^3 P^{(3)}(x) + \ldots$$

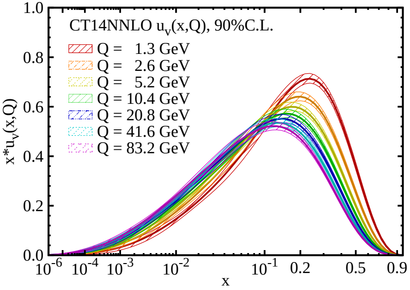
The order of PDFs is determined by the order of splitting function and the hard core calculation.

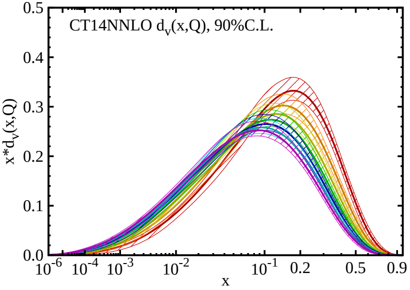
PDF has valence and sea contribution



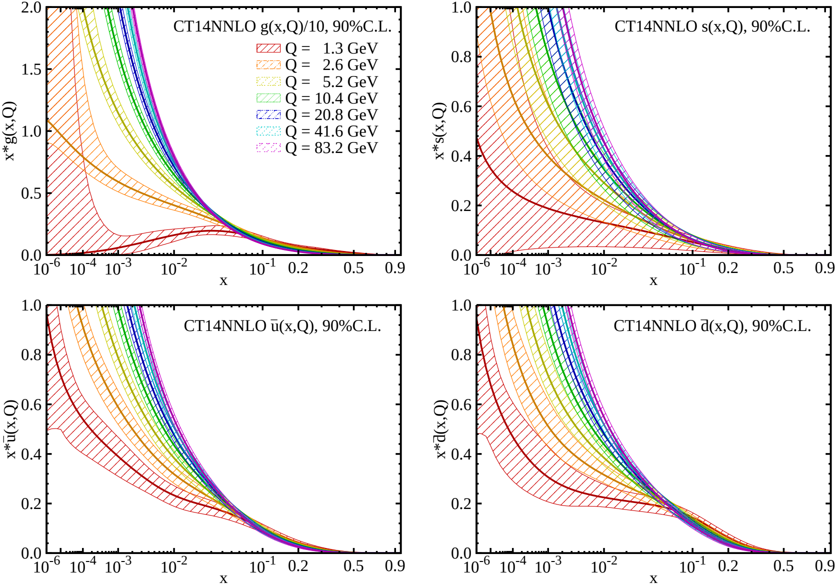


Evolution of PDFs: valence quark

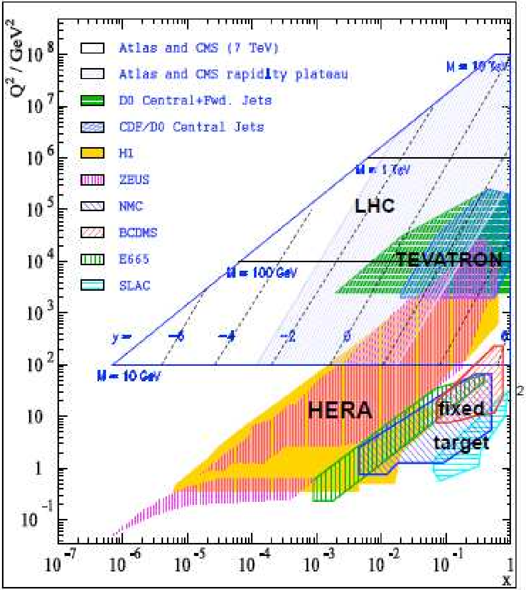




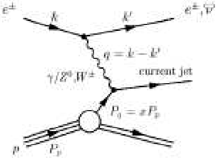
Evolution of PDFs: sea quark and gluon



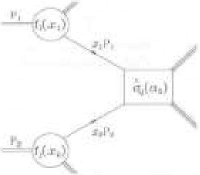
Experimental access to the proton structure



HERA: low and medium x



LHC: important constraints on g(x), flavour separation



Fixed Target: high x, nuclear PDFs

Parametrization of PDFs

CT global analysis takes $Q_0 = 1.3\text{GeV} \gg \Lambda_{QCD}$, and assume

$$xf_a(x, Q_0, \{a_1, a_2, ...\}) = x^{a_1}(1-x)^{a_2}P_a(x)$$

- $x \to 0$: $f \propto x^{a_1}$, Regge-like behavior
- $x \to 1$: $f \propto (1-x)^{a_2}$, quark counting rules
- $P(x; a_3, a_4, ...)$: affects intermediate x; In CT14, Bernstein polynomial is applied.

How many flavor of PDFs should we fit?

$$Q_0 = 1.3\text{GeV} \gg \Lambda_{QCD}$$

$\rightarrow$ all flavors under charm quark mass m_c.

$$g, \quad u_v, \quad d_v, \quad s, \quad \bar{u}_s = u_s, \quad \bar{d}_s = d_s$$

Where $u = u_v + u_s$, and $d = d_v + d_s$.

Why not $Q_0 > m_c$? what about intrinsic charm?

We will back to this.

u $\quad I(J^P) = \frac{1}{2}(\frac{1}{2}^+)$

$m_u = 2.2^{+0.6}_{-0.4}$ MeV $\quad$ Charge $= \frac{2}{3}\ e$ $\quad I_z = +\frac{1}{2}$
$m_u/m_d = 0.38$–0.58

---

d $\quad I(J^P) = \frac{1}{2}(\frac{1}{2}^+)$

$m_d = 4.7^{+0.5}_{-0.4}$ MeV $\quad$ Charge $= -\frac{1}{3}\ e$ $\quad I_z = -\frac{1}{2}$
$m_s/m_d = 17$–22
$\overline{m} = (m_u+m_d)/2 = 3.5^{+0.7}_{-0.3}$ MeV

---

s $\quad I(J^P) = 0(\frac{1}{2}^+)$

$m_s = 96^{+8}_{-4}$ MeV $\quad$ Charge $= -\frac{1}{3}\ e$ $\quad$ Strangeness $= -1$
$m_s\ /\ ((m_u + m_d)/2) = 27.3 \pm 0.7$

---

c $\quad I(J^P) = 0(\frac{1}{2}^+)$

$m_c = 1.27 \pm 0.03$ GeV $\quad$ Charge $= \frac{2}{3}\ e$ $\quad$ Charm $= +1$
$m_c/m_s = 11.72 \pm 0.25$
$m_b/m_c = 4.53 \pm 0.05$
$m_b - m_c = 3.45 \pm 0.05$ GeV

---

b $\quad I(J^P) = 0(\frac{1}{2}^+)$

Charge $= -\frac{1}{3}\ e$ $\quad$ Bottom $= -1$

$m_b(\overline{\text{MS}}) = 4.18^{+0.04}_{-0.03}$ GeV
$m_b(1\text{S}) = 4.66^{+0.04}_{-0.03}$ GeV

---

t $\quad I(J^P) = 0(\frac{1}{2}^+)$

Charge $= \frac{2}{3}\ e$ $\quad$ Top $= +1$

Mass (direct measurements) $m = 173.21 \pm 0.51 \pm 0.71$ GeV [a,b]
Mass ($\overline{\text{MS}}$ from cross-section measurements) $m = 160^{+5}_{-4}$ GeV [a]
Mass (Pole from cross-section measurements) $m = 174.2 \pm 1.4$ GeV

Requirements for PDF parametrization

- Valence quark number sum rule

$$\int_0^1 [u(x) - \bar{u}(x)]dx = 2, \int_0^1 [d(x) - \bar{d}(x)]dx = 1$$

$$\int_0^1 [s(x) - \bar{s}(x)]dx = 0$$

Where $u = u_v + \bar{u}$, $d = d_v + \bar{d}$. $(s(x) - \bar{s}(x))$ can be non-zero.

- Momentum sum rule

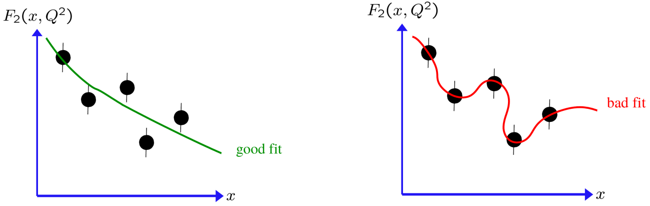
$$\sum_{a=q,\bar{q},g} \int_0^1 x f_{a/p}(x,Q)dx = 1$$

Where

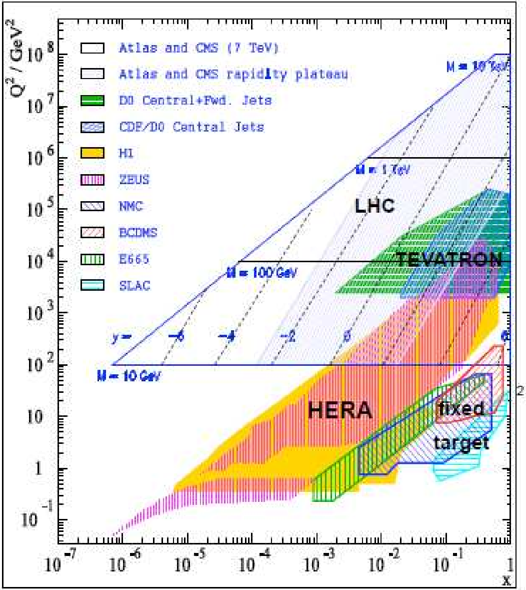
$$\int_0^1 x\, g(x,Q)dx \sim 0.45$$

Requirements for PDF parametrization

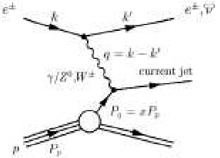
- A valid PDF set must not produce unphysical predictions for observables
  - Any concievable hadron cross section σ must be non-negative: $\sigma > 0$. This is typically realized by requiring $f_{a/p}(x, Q) > 0$.
  - Any cross section asymmetry A must lie in the range $-1 \leq A \leq 1$. This constrains the range of allowed PDF parametrizations.
- PDF parametrization for $f_{i/p}(x, Q)$ must be ”flexible just enough” to reach agreement with the data, without reproducing random fluctuation.



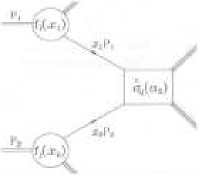
Experimental access to the proton structure



HERA: low and medium x

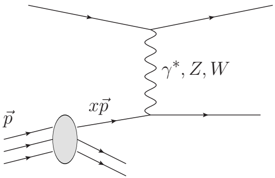


LHC: important constraints on g(x), flavour separation

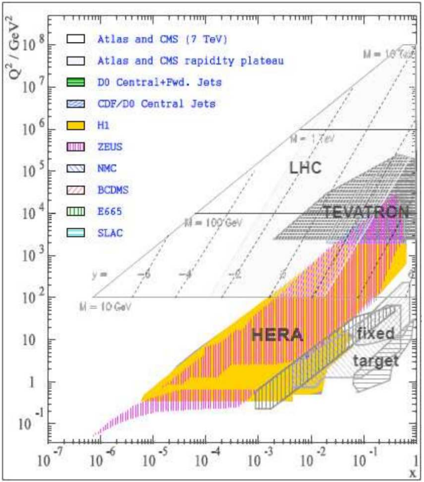


Fixed Target: high x, nuclear PDFs

inclusive Deep Inelastic Scattering(DIS)

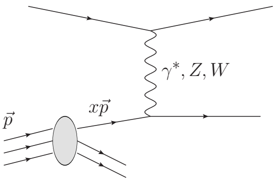


HERA:
Neutral-current $e^{\pm}p \to e^{\pm}X$
Charged-current $ep \to \nu X$



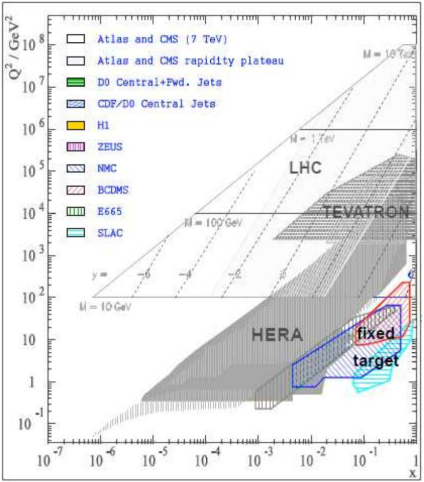
P. Nodolsky

inclusive Deep Inelastic Scattering(DIS)



HERA:
Neutral-current $e^{\pm}p \to e^{\pm}X$
Charged-current $ep \to \nu X$

Fixed-target:
eN,νN, νN scattering

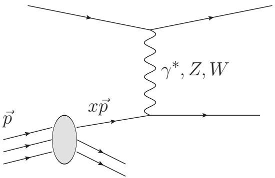


P. Nodolsky

Neutral-current *ep* DIS

Kinematics:

- $s = (p_e + p_p)^2$: total energy
- $Q^2 = -q^2 = -(p_e - p'_e)^2$: momentum transfer
- $x = Q^2/(2p_p \cdot q)$: Bjorken scaling variable
- $y = Q^2/(xs)$: inelasticity
- $W^2 = Q^2(1-x)/x$: energy of the hadronic final state



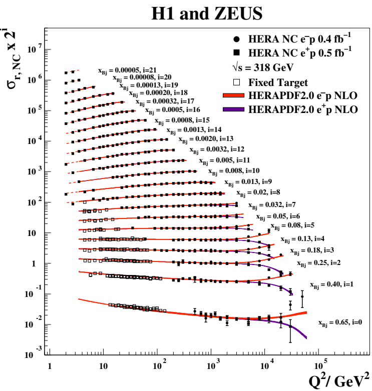
$$\frac{d^2\sigma^{\mathrm{NC},\ell^\pm}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[Y_+F_2^{NC} \mp Y_-xF_3^{NC} - y^2F_L^{NC}\right]$$

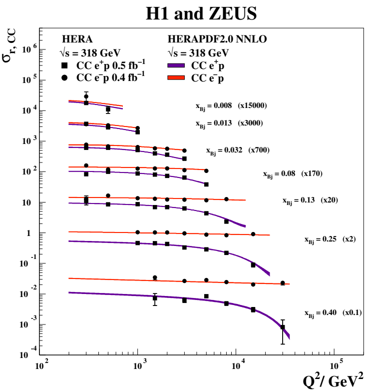
With $Y_\pm = 1 \pm (1-y)^2$. The data is fitted either in the form of structure function or reduced cross section.

P. Nodolsky

Final combined DIS cross sections at HERA(1506.06042)

41 data sets on NC and CC DIS from H1 and ZEUS are combined into 1 set. 2927 data points are combined into 1307 data points. 165 correlated systematic errors are reanalyzed and calibrated.





PDF combinations in DIS at the lowest order

Neutral current $l^{\pm}p$:

$$F_2^{l^{\pm}p}(x,Q^2) = \frac{4}{9}(u+\bar{u}+c+\bar{c}) + \frac{1}{9}(d+\bar{d}+s+\bar{s}+b+\bar{b})$$

- PDFs are weighted by the fractional EM quark coupling $e_i^2 = 4/9$ or $1/9$.
- 4 times more sensitivity to u and c than to d, s, and b
- No sensitivity to the gluon at this order

Charged current (νN) DIS :

$$F_2^{\nu N}(x,Q^2) = x \sum_{i=u,d,s,\ldots} (q_i + \bar{q}_i)$$

$$xF_3^{\nu N}(x,Q^2) = x \sum_{i=u,d,s,\ldots} (q_i - \bar{q}_i)$$

DIS at next-to-leading order (NLO) and beyond

Logarithmic corrections to Bjorken scaling (Q dependence of $F_2(x, Q^2)$) are sensitive to the gluon PDF through DGLAP equations,

$$\frac{df_{i/p}(x,\mu)}{d\ln\mu} = \sum_{j=g,q,\bar{q}..} P_{i/j} \otimes f_{j/p}(x,\mu)$$

Thus, when examined at NLO, the DIS data constrains

- $\sum_i (q_i + \bar{q}_i)$ in an amazingly large range $10^{-5} < x < 0.5$.
- u and d at $10^{-2} < x < 0.3$.
- $g(x, Q)$ at $x < 0.1$.

DIS cannot fully separate quarks from antiquarks, or s, c, b contributions from u and d contributions; more so because of systematic effects in fixed-target DIS experiments (higher-order terms, nuclear corrections,...)

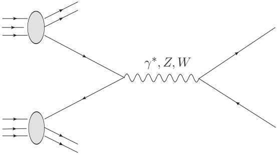
P. Nodolsky

Z and W production and charge asymmetry

$$\frac{d\sigma^Z}{dy} \propto \frac{4}{9}[u_A\bar{u}_B + \bar{u}_A u_B] + \frac{1}{9}[d_A\bar{d}_B + \bar{d}_A d_B] + \dots$$

$$\frac{d\sigma^{W^+}}{dy} \propto u_A\bar{d}_B + \bar{d}_A u_B + \dots$$

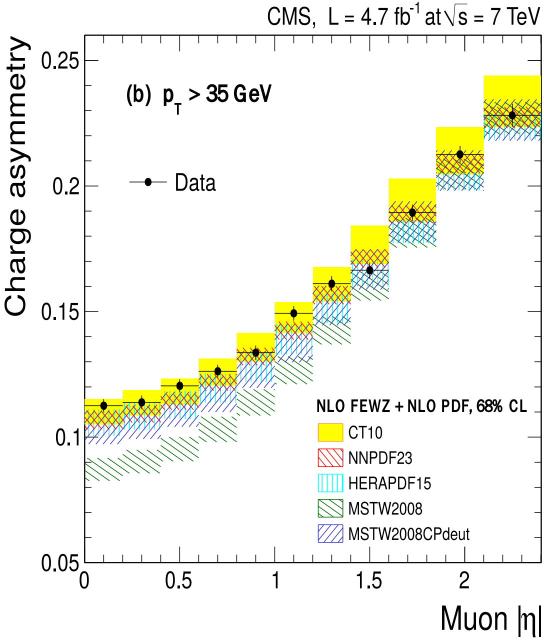
$$\frac{d\sigma^{W^-}}{dy} \propto \bar{u}_A d_B + d_A\bar{u}_B + \dots$$



$$A_{ch}(y) \equiv \frac{\frac{d\sigma^{W^+}}{dy} - \frac{d\sigma^{W^-}}{dy}}{\frac{d\sigma^{W^+}}{dy} + \frac{d\sigma^{W^-}}{dy}}$$

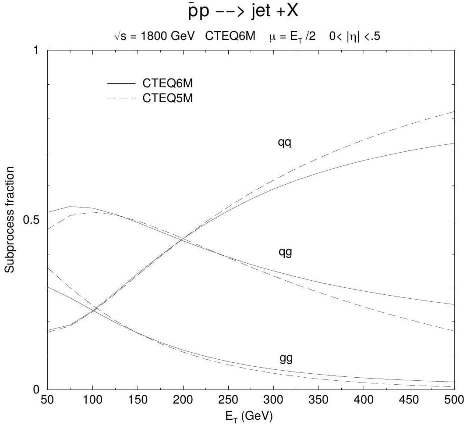
$A_{ch}(y)$ constrains PDF ratio at $Q \approx m_w$:

- d/u at $x \to 1$ at Tevatron 1.96 TeV ($p\bar{p}$)
- d/u at $x > 0.1$ and $\bar{d}/\bar{u}$ at $x \sim 0.01$ at the LHC 7TeV(pp)



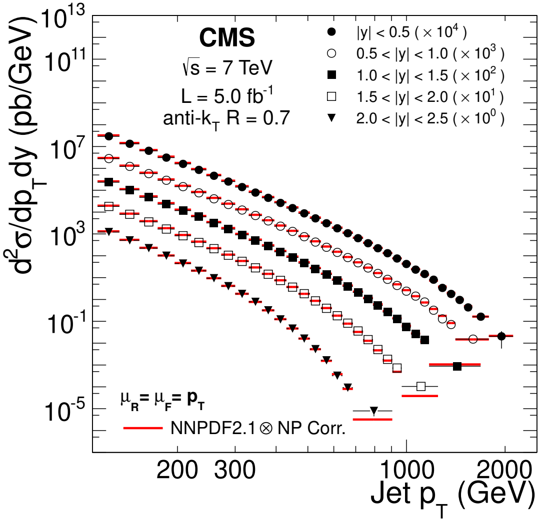
CMS $A_{ch}(\eta)$ data disfavor some d/u parametrizations, motivated an update in MSTW'2008 PDFs

Inclusive jet production



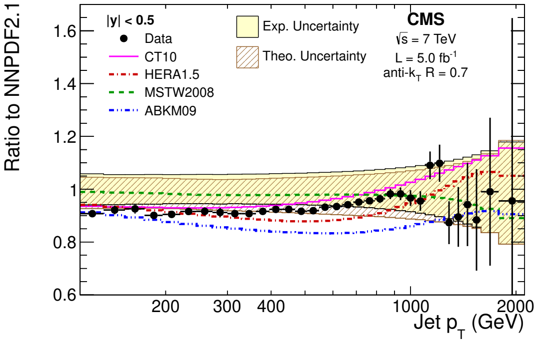
High-ET jets are mostly produced in qq scattering; yet most of the PDF uncertainty arises from qg and gg contributions Here typical x is of order $2E_T/\sqrt{s} \gtrsim 0.1$; e.g. $x \approx 0.2$ for $E_T = 200$ GeV $\sqrt{s} = 1.8$ TeV. At such x, $u(x,Q^2)$ and $d(x,Q^2)$ are known very well, uncertainty arise mostly from $g(x,Q^2)$.

Inclusive jet production in $pp \to$ jet + X



- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error

Inclusive jet production in $pp \to$ jet + X



- Systematic uncertainties dominate, both from the experiment (up to 90 correlated sources of uncertainty) and NLO theoretical cross section (QCD scale dependence)
- The PDF uncertainty would be strongly underestimated if these systematic errors are not included

Criteria for determining PDFs

$$\chi^2_{global} = \sum_i \frac{[D_i - \sum_k \lambda_k \beta_{ki} - T_i(\{a\})]^2}{\sigma_i^2} + \sum_k \lambda_k^2.$$

Where
D_i is the central value of data,
$T_i(\{a\})$ is the theoretical prediction of the data,
σ_i^2 is the qudratic sum of the statistical error and uncorrelated error,
β_{ki} is the matrix for correlated error,
and λ_k are the nuisance parameters.

The PDF is obtained by minimizing the global χ^2 function respect to shape parameters $\{a\}$ and nuisance parameters $\{\lambda\}$.

Source of PDF uncertainty

PDFs receive uncertainty from...

- Experimental uncertainty; statistic and systematic.
- Uncertainties from theoretical input, such like α_s, m_q, α.
- QCD and QED approximation in perturbation (or scale uncertainty).
- Numerical uncertainty from MC integration.
- Choice of the parametrization form of PDF at Q_0 scale.
- Methods for uncertainty estimation.

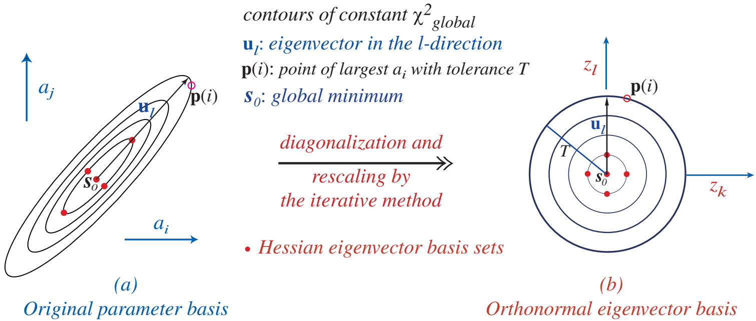
Probing uncertainty of PDFs

The uncertainty of the PDF is estimated by the Hessian method.

$$\chi^2 = \chi_0^2 + \sum_{i,j} H_{ij} y_i y_j \,, \quad H_{ij} = \frac{1}{2} \left(\frac{\partial^2 \chi^2}{\partial y_i \partial y_j} \right)_0 \,,$$

Where $y_i = a_i - a_i^0$ with a_i^0 to be the parameters at minimal χ_0^2.

2-dim (i,j) rendition of d-dim (~16) PDF parameter space



Hessian uncertainty

Let $X = X(\{a_i\})$ to be the observable as a function of fitting parameter. Using the linear approximation of parameter $\{z_i\}$, the symmetry uncertainty of X is,

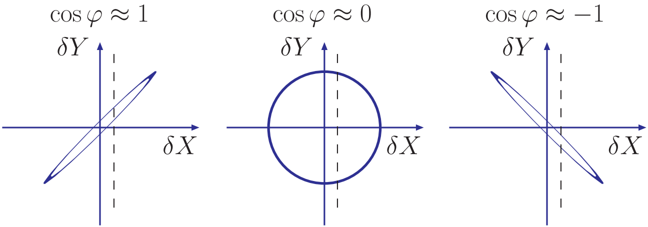
$$\Delta X = \frac{1}{2}\left(\sum_{i=1}^{N_p}\left[X(\{z_i^+\}) - X(\{z_i^-\})\right]^2\right)^{1/2},$$

Where $\{z_1^\pm\} = \{\pm T, 0....\}$, $\{z_2^\pm\} = \{0, \pm T, 0....\}$ and so on.
The asymmetry uncertainty of X is,

$$\delta^+ X = \sqrt{\sum_{i=1}^{N_a}\left[\max\left(X_i^{(+)} - X_0, X_i^{(-)} - X_0, 0\right)\right]^2},$$

$$\delta^- X = \sqrt{\sum_{i=1}^{N_a}\left[\max\left(X_0 - X_i^{(+)}, X_0 - X_i^{(-)}, 0\right)\right]^2},$$

Correlation between observables



In the framework of the Hessian, the correlation between two variables X and Y can be worked out as.

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

where the ΔX and ΔY are their symmetric uncertainties. By this correlation angle φ, the tolerance ellipse is defined by

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

- The correlation cosine between PDF $f(x,\mu)$ and theoretical prediction T_i contains no information of the experimental uncertainty.
- The correlation cosine $C_f(x_i,\mu_i)$ between PDF $f(x,\mu)$ and residual r_i contains no information of the experimental uncertainty in practice

$$C_f(x_i,\mu_i) = \frac{\vec{\nabla} f(x_i,\mu_i) \cdot \vec{\nabla} r_i}{\Delta f(x_i,\mu_i) \Delta r_i}, \quad \text{where}$$

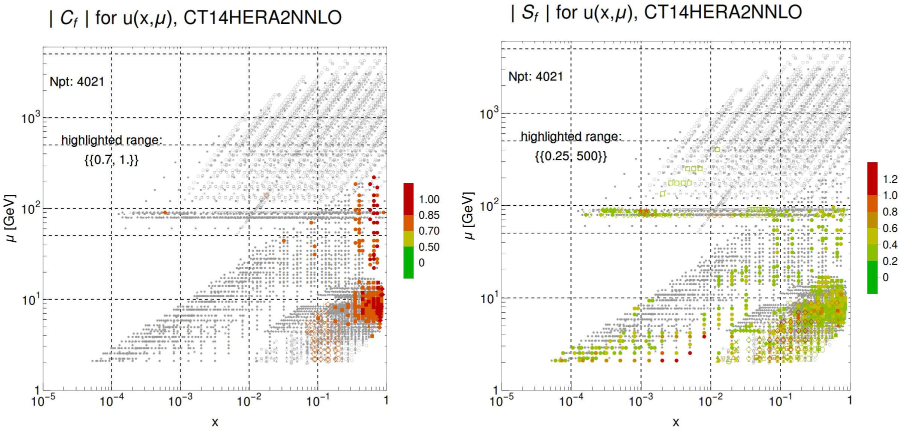
$$\chi^2 = \sum_i^N r_i^2 + \sum_k \lambda_k^2, \quad r_i(\vec{a}) = \frac{D_i - \sum_k \lambda_k \beta_{ki} - T_i(\{a\})}{\sigma_i}$$

- Instead, we concern the "sensitivity" $S_f(x_i,\mu_i)$

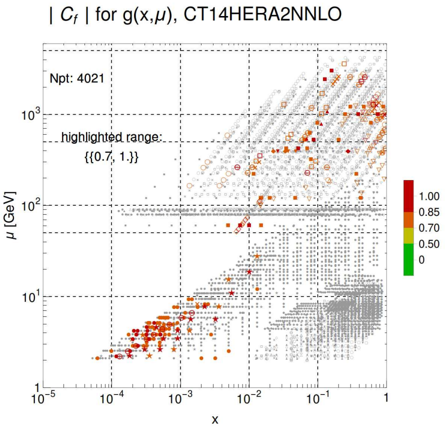
$$S_f(x_i,\mu_i) = C_f(x_i,\mu_i) \frac{\Delta r_i}{\sqrt{\frac{\sum_i^N r_i^2}{N}}}$$

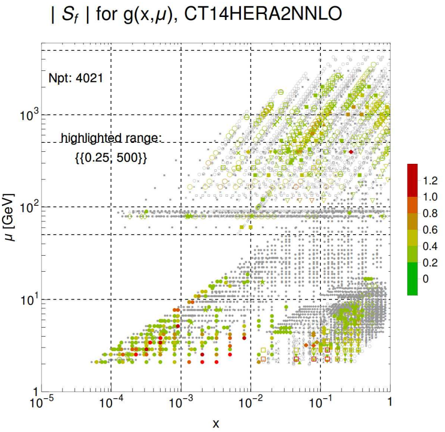
The sensitivity $S_f(x_i, \mu_i)$ help us to visualize the potential impact on PDF in $x - Q$ plane.

$$u(x, \mu)$$



$g(x,\mu)$


| C_f | for g(x,μ), CT14HERA2NNLO
Npt: 4021
highlighted range:
{{0.7, 1.}}
μ [GeV]
x
1.00
0.85
0.70
0.50
0

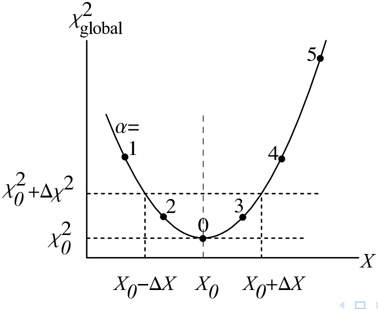

| S_f | for g(x,μ), CT14HERA2NNLO
Npt: 4021
highlighted range:
{{0.25, 500}}
μ [GeV]
x
1.2
1.0
0.8
0.6
0.4
0.2
0

Lagrange Multiplier Method

Conaider a particular physical quantity, say $X(\{a_i\})$, which is a function of PDFs.

$$F(\lambda, \{a_i\}) = \chi^2(\{a_i\}) + \lambda(X(\{a_i\}) - X(\{a_i^{(0)}\}))$$

By minimizing this function with various fixed λ value, say $\lambda_1, \ldots \lambda_j \ldots, \lambda_n$, we will obtain n parameter sets $\{a_i(\lambda_j)\}$ and corresponding $X(\{a_i(\lambda_j)\})$ and $\chi^2(\{a_i(\lambda_j)\})$. With suitable choice of $\Delta\chi^2$, we obtain the uncertainty of the physical quantity $X(\{a_i\})$.



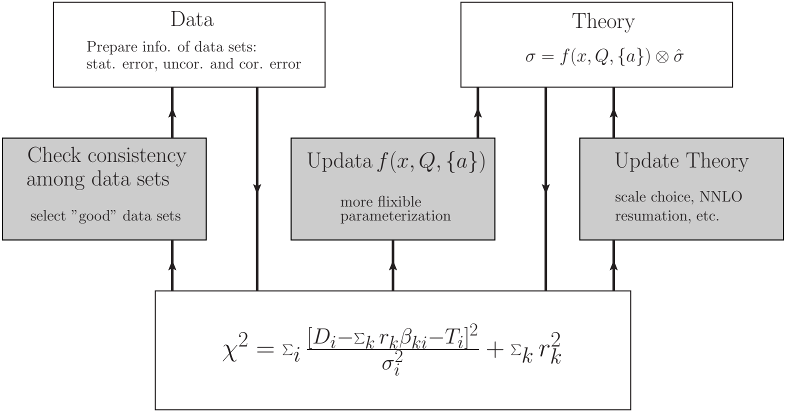
Data

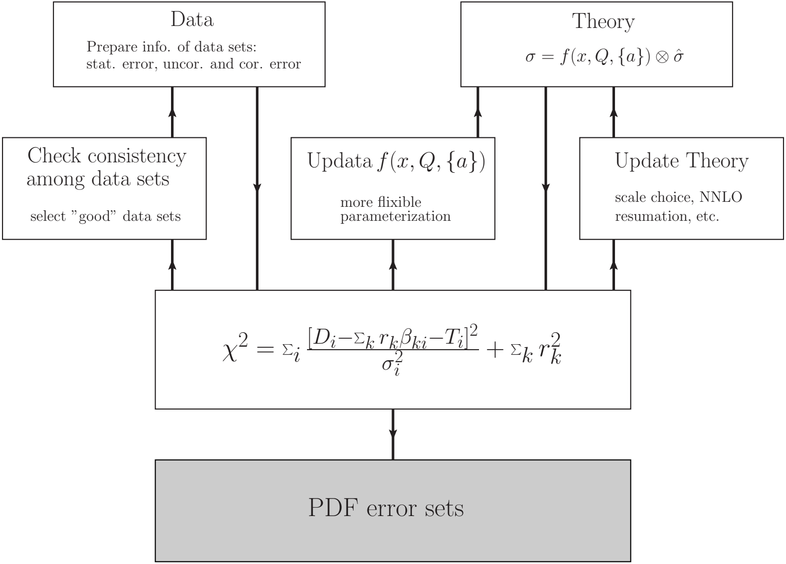
Prepare info. of data sets:
stat. error, uncor. and cor. error

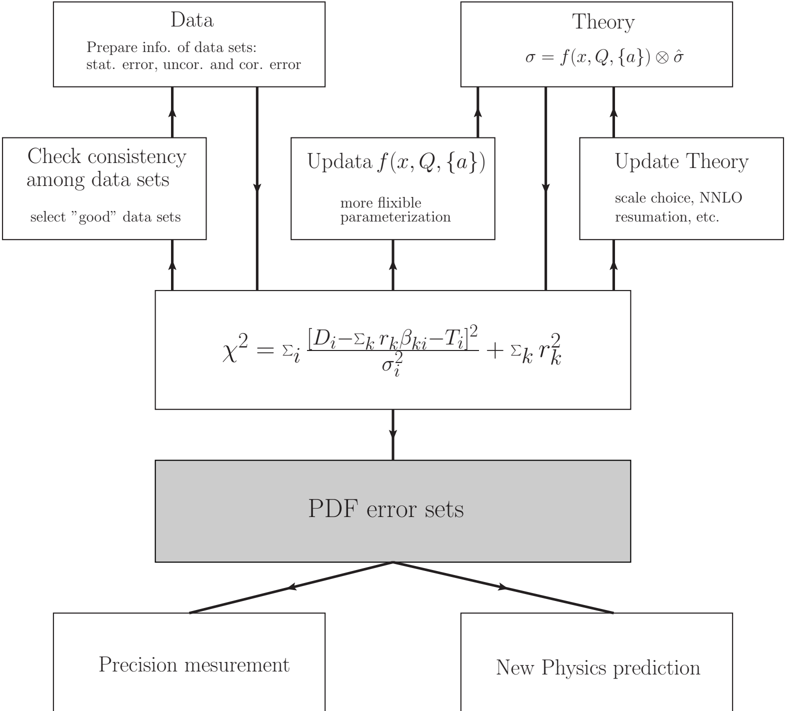
Theory

$\sigma = f(x, Q, \{a\}) \otimes \hat{\sigma}$

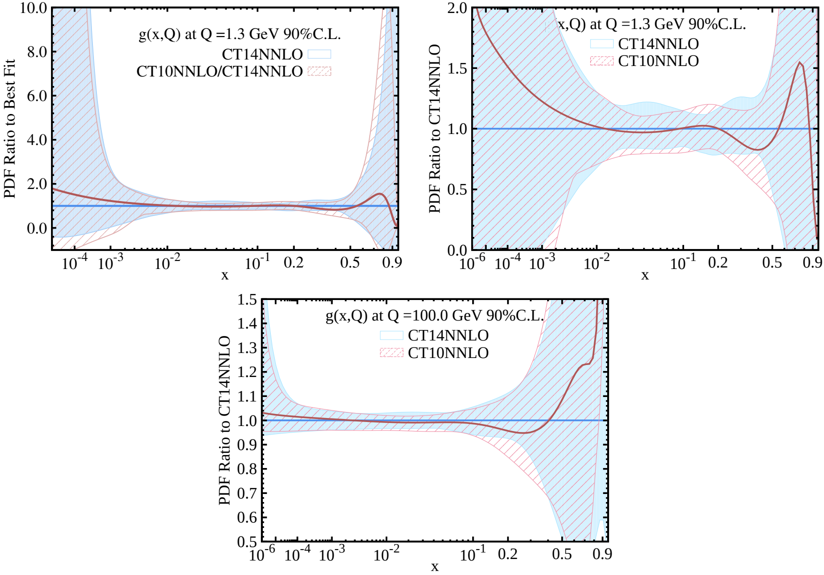
$$\chi^2 = \Sigma_i \frac{[D_i - \Sigma_k r_k \beta_{ki} - T_i]^2}{\sigma_i^2} + \Sigma_k r_k^2$$


Data
Prepare info. of data sets:
stat. error, uncor. and cor. error
Theory
$\sigma = f(x, Q, \{a\}) \otimes \hat{\sigma}$
Check consistency among data sets
select "good" data sets
Updata $f(x, Q, \{a\})$
more flixible parameterization
Update Theory
scale choice, NNLO resumation, etc.
$\chi^2 = \Sigma_i \frac{[D_i - \Sigma_k r_k \beta_{ki} - T_i]^2}{\sigma_i^2} + \Sigma_k r_k^2$


Data
Prepare info. of data sets:
stat. error, uncor. and cor. error
Theory
$\sigma = f(x, Q, \{a\}) \otimes \hat{\sigma}$
Check consistency among data sets
select "good" data sets
Updata $f(x, Q, \{a\})$
more flixible parameterization
Update Theory
scale choice, NNLO resumation, etc.
$\chi^2 = \Sigma_i \frac{[D_i - \Sigma_k r_k \beta_{ki} - T_i]^2}{\sigma_i^2} + \Sigma_k r_k^2$
PDF error sets

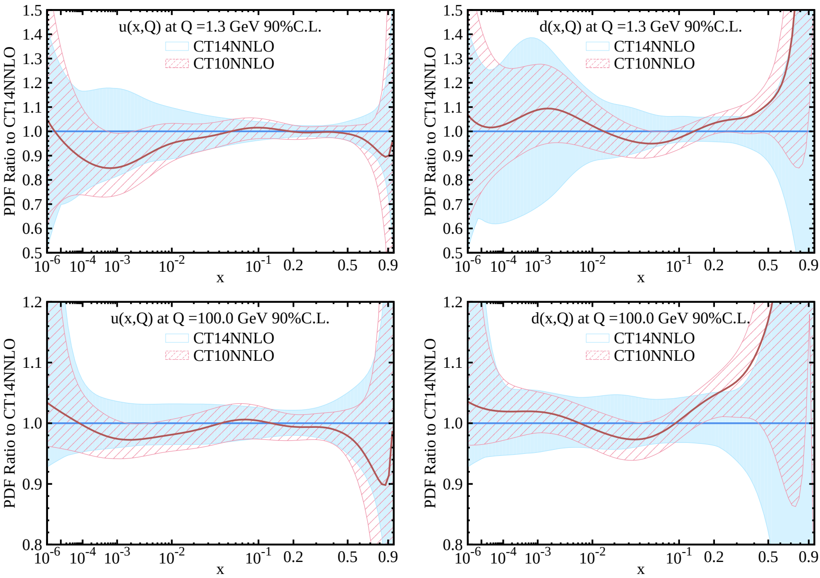

Data
Prepare info. of data sets:
stat. error, uncor. and cor. error
Theory
$\sigma = f(x, Q, \{a\}) \otimes \hat{\sigma}$
Check consistency among data sets
select "good" data sets
Updata $f(x, Q, \{a\})$
more flixible parameterization
Update Theory
scale choice, NNLO resumation, etc.
$\chi^2 = \Sigma_i \frac{[D_i - \Sigma_k r_k \beta_{ki} - T_i]^2}{\sigma_i^2} + \Sigma_k r_k^2$
PDF error sets
Precision mesurement
New Physics prediction

$g(x,Q)$



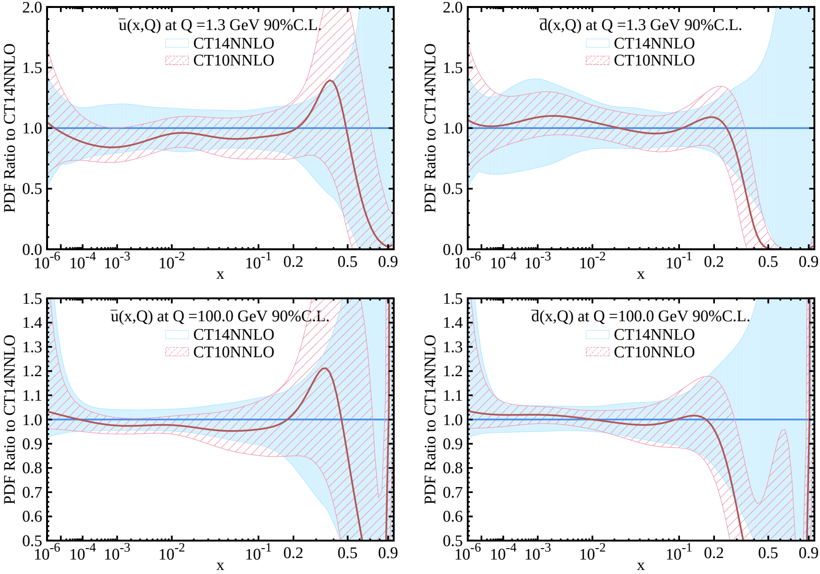
g(x,Q) at Q =1.3 GeV 90%C.L.
CT14NNLO
CT10NNLO/CT14NNLO
PDF Ratio to Best Fit
x
g(x,Q) at Q =1.3 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x
g(x,Q) at Q =100.0 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x

$u(x, Q)$ $d(x, Q)$

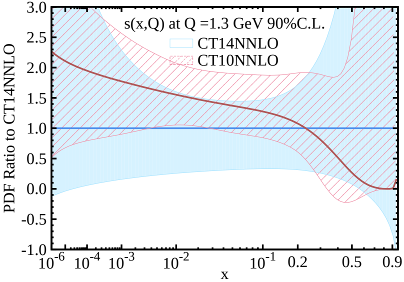


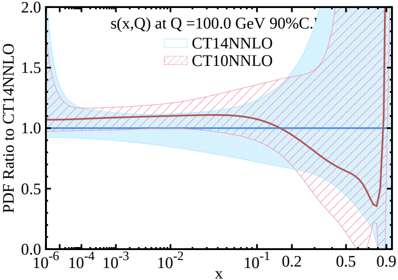
u(x,Q) at Q =1.3 GeV 90%C.L.
CT14NNLO
CT10NNLO
d(x,Q) at Q =1.3 GeV 90%C.L.
CT14NNLO
CT10NNLO
u(x,Q) at Q =100.0 GeV 90%C.L.
CT14NNLO
CT10NNLO
d(x,Q) at Q =100.0 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x

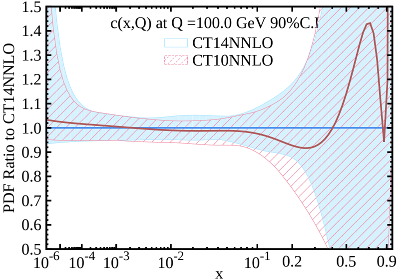
$\bar{u}(x,Q)$ $\bar{d}(x,Q)$



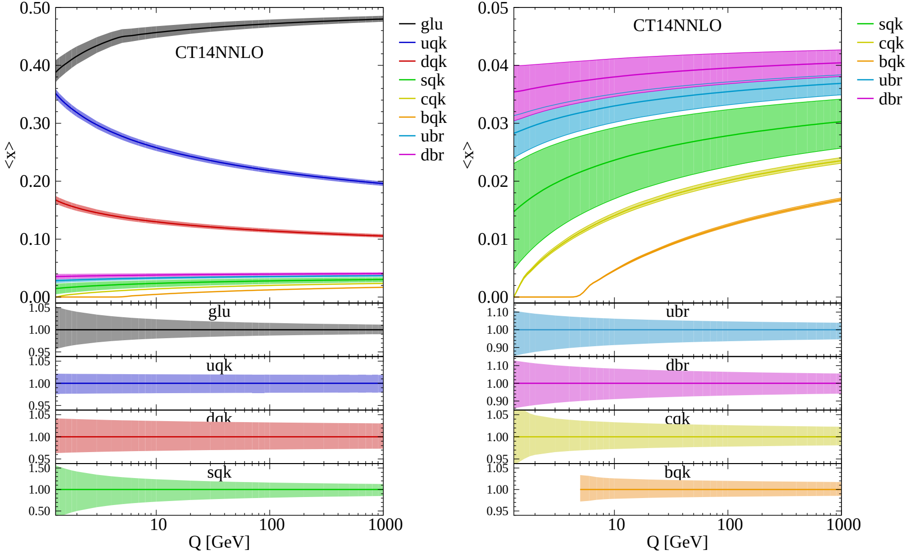
$S(x,Q)$ $c(x,Q)$


s(x,Q) at Q =1.3 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x


s(x,Q) at Q =100.0 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x

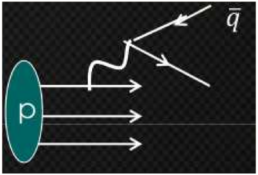

c(x,Q) at Q =100.0 GeV 90%C.L.
CT14NNLO
CT10NNLO
PDF Ratio to CT14NNLO
x

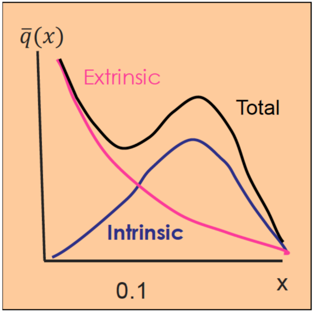
Momentum fraction



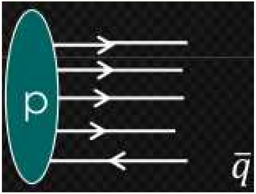
Is there a non-perturbative contribution to charm PDF?

”Extrinsic” charm

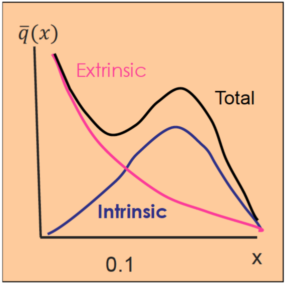




”Intrinsic” charm



1.The Sea-like(extrinsic) component:

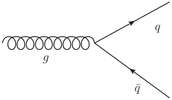


- Monotonic in x, satisfies

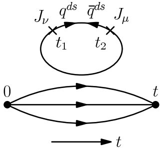
$$q(x) \propto x^{-1}, \text{ for } x \to 0$$

- May be generate in several ways, e.g.

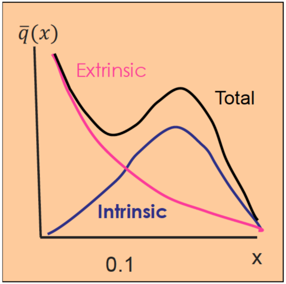
  In PQCD, from gluon splittings



  In Lattice QCD, from disconnected diagrams



2.Valence-like (intrinsic) component:

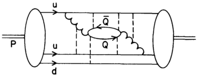


- peaks in x, satisfies

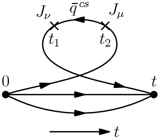
$$q(x) \propto x^{-1/2}, \text{ for } x \to 0$$

- May be generate in several ways, e.g.

For all flaovers, nonperturbatively from a $|uudQ\bar{Q}>$ Fock state: (Brodsky, Peterson, Sakai, PRD 1981)



In Lattice QCD, from connected diagrams



Parametrizations for BHPS and SEA models

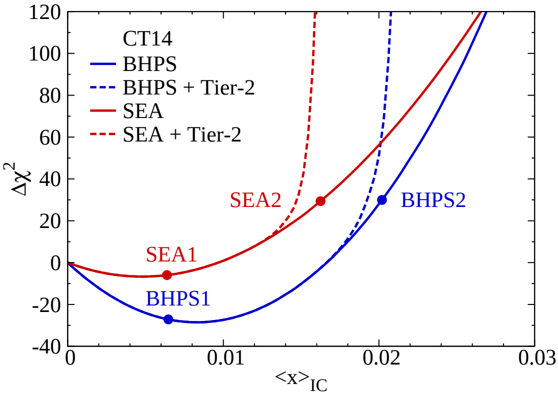
- ”Valence-like” charm quark PDF according to the BHPS model (scale is unknown in this model):

$$c(x) = \bar{c}(x) = \frac{1}{2} A\, x^2 \left[\frac{1}{3}(1-x)(1+10x+x^2) - 2x(1+x)\ln(1/x) \right].$$

- ”Sea-like” charm quark distribution, similar to that of the light flavor sea quarks:

$$c(x) = \bar{c}(x) = A \left[\overline{d}(x, Q_0) + \overline{u}(x, Q_0) \right]$$

Intrinsic charm

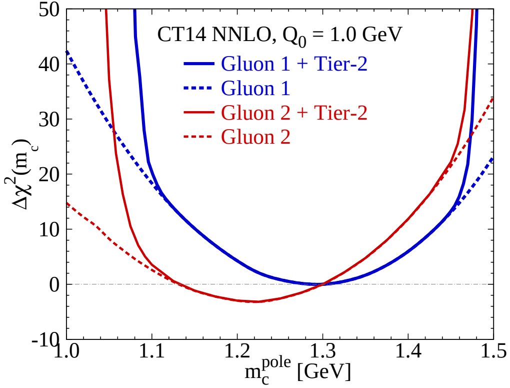


Intrinsic charm (IC) can carry up to 1% of the proton momentum

$$\langle x\rangle_{\mathrm{IC}} \lesssim 0.021 \text{ for CT14 BHPS,}$$
$$\langle x\rangle_{\mathrm{IC}} \lesssim 0.016 \text{ for CT14 SEA.}$$

arXiv: 1707.00657

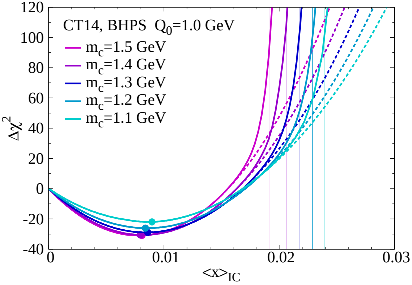
m_c^{pole} scan

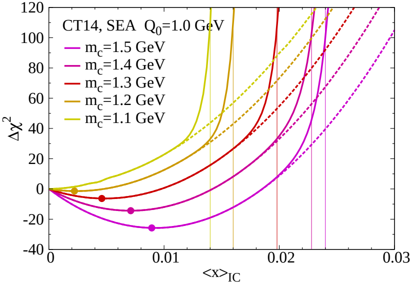


- The location of minimal χ^2 on m_c depend on the choice of Q_0 and parametrization.
- The charm quark mass is allowed to be 1.1 GeV $\lesssim m_c \lesssim$ 1.5 GeV.

DEPENDENCE OF FIT ON THE CHARM-QUARK MASS

The combined HERA charm production and inclusive DIS data play an important role in the description of the goodness of fit. m_c is a key input scale.

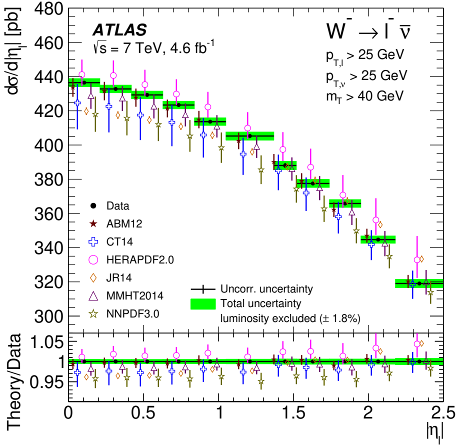




BHPS model: the position of the minimum is relatively stable as m_c is varied, while the upper limit on the amount of IC decreases to 1.7%. BHPS model is not dramatically affected by variations of m_c

SEA model: limits on the amount of IC allowable are shifted towards higher values. ubar and dbar are well constrained by data (vector boson production in pp and $p\bar{p}$) in the intermediate/small x region, and cannot change too much

The precision of the experimental measurement at Large hardron Collider(LHC) reach 1% level already. This require the precision of the PDF has to be precise to the same level.



W production in LHC

Green: error of measument, about 1%

error bar: error of various PDFs, about 5% 7%

arXiv: 1612.03016

Is there a way to see the potential impact to PDF before real global analysis?

- How to know the potential impact among plenty of very precise data of LHC?
- How to know if the data help on the determination of particular observable?

Hessian Updating

- Updated Chi-square function :

$$
\begin{aligned}
\Delta\chi^2(Z) &= \Delta\chi^2_{old} + (X^E_\alpha - X_\alpha) C^{-1}_{\alpha\beta} (X^E_\beta - X_\beta) \\
&= T^2 [c_i - \bar{c}_i]^2 - \bar{c}_i \bar{c}_i + (X^E_\alpha - X^0_\alpha) C^{-1}_{\alpha\beta} (X^E_\beta - X^0_\beta)
\end{aligned}
$$

Where $c_j = \sqrt{1+\lambda_j} V^T_{jk} z_k$, $\bar{c}_i = \frac{1}{\sqrt{\lambda_i}} V_{ji} A_i$

$$
A^i = \frac{-2}{T^2} (X^E_\alpha - X^0_\alpha) C^{-1}_{\alpha\beta} \Delta X^i_\beta, \quad M^{ij} = \frac{1}{T^2} \Delta X^i_\alpha C^{-1}_{\alpha\beta} \Delta X^j_\beta
$$

- Minimize to find new best fit happen when

$$
Z^2_{new} = (1+M)^{-1} A \quad \text{or} \quad c = \bar{c}
$$

- New best-fit PDF and error PDFs :

$$
\begin{aligned}
f^0_{new} &= f^0 + \Delta f \cdot Z \\
f^{\pm(r)} &= f^0_{new} \pm \frac{1}{\sqrt{1+\lambda^{(r)}}} \Delta f \cdot U^{(r)}
\end{aligned}
$$

arXiv: 1806.07950

ePump
Error PDF Updating Method Package

FullCT14HERA2.in — Edited

```
+++ N(EV pairs)                        N(Data Sets)    PDFtype(C/L/N)  Dyn_Tol?(Y/N)  Tol_squared
         27                                 33                C               Y            100.0
+++ ObservableFile                                N(Observables)   Data?(Y/N)      Error_type        Weight
  CT14HERA2ex/tabs/E160.If1363                        1120            Y               1              1.0
  CT14HERA2ex/tabs/E101.If1363                         337            Y               1              1.0
  CT14HERA2ex/tabs/E102.If1363                         250            Y               1              1.0
  CT14HERA2ex/tabs/E104.If1363                         123            Y               1              1.0
  CT14HERA2ex/tabs/E108.If1363                          85            Y               1              1.0
  CT14HERA2ex/tabs/E109.If1363                          96            Y               1              1.0
  CT14HERA2ex/tabs/E110.If1363                          69            Y               1              1.0
  CT14HERA2ex/tabs/E111.If1363                          86            Y               1              1.0
  CT14HERA2ex/tabs/E124.If1363                          38            Y               1              1.0
  CT14HERA2ex/tabs/E125.If1363                          33            Y               1              1.0
  CT14HERA2ex/tabs/E126.If1363                          40            Y               1              1.0
  CT14HERA2ex/tabs/E127.If1363                          38            Y               1              1.0
  CT14HERA2ex/tabs/E147.If1363                          47            Y               1              1.0
  CT14HERA2ex/tabs/E145.If1363                          10            Y               1              1.0
  CT14HERA2ex/tabs/E169.If1363                           9            Y               1              1.0
  CT14HERA2ex/tabs/E201.If1363                         119            Y               1              1.0
  CT14HERA2ex/tabs/E203.If1363                          15            Y               1              1.0
  CT14HERA2ex/tabs/E204.If1363                         184            Y               1              1.0
  CT14HERA2ex/tabs/E225.If1363                          11            Y               1              1.0
  CT14HERA2ex/tabs/E227.If1363                          11            Y               1              1.0
  CT14HERA2ex/tabs/E234.If1363                           9            Y               1              1.0
  CT14HERA2ex/tabs/E260.If1363                          28            Y               1              1.0
  CT14HERA2ex/tabs/E261.If1363                          29            Y               1              1.0
  CT14HERA2ex/tabs/E267.If1363                          11            Y               1              1.0
  CT14HERA2ex/tabs/E268.If1363                          41            Y               1              1.0
  CT14HERA2ex/tabs/E240.If1363                          14            Y               1              1.0
  CT14HERA2ex/tabs/E241.If1363                           5            Y               1              1.0
  CT14HERA2ex/tabs/E281.If1363                          13            Y               1              1.0
  CT14HERA2ex/tabs/E266.If1363                          11            Y               1              1.0
  CT14HERA2ex/tabs/E504.If1363                          72            Y               1              1.0
  CT14HERA2ex/tabs/E514.If1363                         110            Y               1              1.0
  CT14HERA2ex/tabs/E535.If1363                          90            Y               1              1.0
  CT14HERA2ex/tabs/E538.If1363                         133            Y               1              1.0
+++      PDFin                 PDFout
  PDFs/CT14HERA2ex/If1363    CT14HERA2ex/PDFtmp/If1363
```

".in" file

How to use ePump

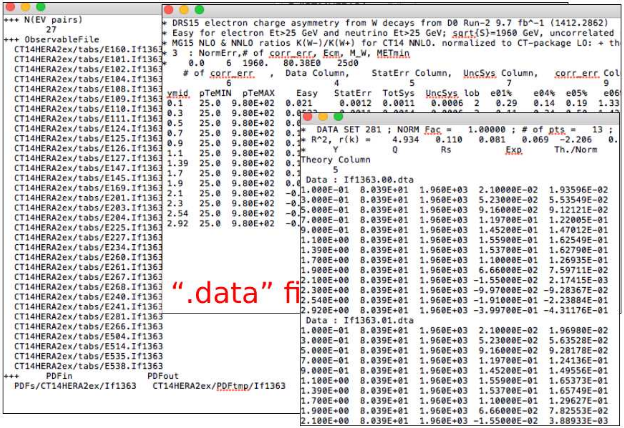
```
+++ N(EV pairs)
        27
+++ ObservableFile
    CT14HERA2ex/tabs/E160.If1363
    CT14HERA2ex/tabs/E101.If1363
    CT14HERA2ex/tabs/E102.If1363
    CT14HERA2ex/tabs/E104.If1363
    CT14HERA2ex/tabs/E108.If1363
    CT14HERA2ex/tabs/E109.If1363
    CT14HERA2ex/tabs/E110.If1363
    CT14HERA2ex/tabs/E111.If1363
    CT14HERA2ex/tabs/E124.If1363
    CT14HERA2ex/tabs/E125.If1363
    CT14HERA2ex/tabs/E126.If1363
    CT14HERA2ex/tabs/E127.If1363
    CT14HERA2ex/tabs/E147.If1363
    CT14HERA2ex/tabs/E145.If1363
    CT14HERA2ex/tabs/E169.If1363
    CT14HERA2ex/tabs/E201.If1363
    CT14HERA2ex/tabs/E203.If1363
    CT14HERA2ex/tabs/E204.If1363
    CT14HERA2ex/tabs/E225.If1363
    CT14HERA2ex/tabs/E227.If1363
    CT14HERA2ex/tabs/E234.If1363
    CT14HERA2ex/tabs/E260.If1363
    CT14HERA2ex/tabs/E261.If1363
    CT14HERA2ex/tabs/E267.If1363
    CT14HERA2ex/tabs/E268.If1363
    CT14HERA2ex/tabs/E240.If1363
    CT14HERA2ex/tabs/E241.If1363
    CT14HERA2ex/tabs/E281.If1363                13           Y          1         1.0
    CT14HERA2ex/tabs/E266.If1363                11           Y          1         1.0
    CT14HERA2ex/tabs/E504.If1363                72           Y          1         1.0
    CT14HERA2ex/tabs/E514.If1363               110           Y          1         1.0
    CT14HERA2ex/tabs/E535.If1363                90           Y          1         1.0
    CT14HERA2ex/tabs/E538.If1363               133           Y          1         1.0
+++      PDFin            PDFout
    PDFs/CT14HERA2ex/If1363   CT14HERA2ex/PDFtmp/If1363
```

```
* DRS15 electron charge asymmetry from W decays from D0 Run-2 9.7 fb^-1 (1412.2862)
* Easy for electron Et>25 GeV and neutrino Et>25 GeV; sqrt{S}=1960 GeV, uncorrelated
* MG15 NLO & NNLO ratios K(W-)/K(W+) for CT14 NNLO. normalized to CT-package LO: + th
* 3  : NormErr,# of corr_err, Ecm, M_W, METmin
*    0.0    6  1960.   80.38E0   25d0
   # of corr_err   ,  Data Column,     StatErr Column,  UncSys Column,    corr_err Col
           6                 4                   5                7                 9
ymid  pTeMIN  pTeMAX     Easy     StatErr  TotSys   UncSys lob  e01%    e04%  e05%   e06
0.1    25.0  9.80E+02  0.021      0.0012  0.0011   0.0006  2   0.29    0.14  0.19  1.33
0.3    25.0  9.80E+02  0.0523     0.0011  0.0014   0.0006  2   0.11    0.34  0.50  1.43
0.5    25.0  9.80E+02  0.0916     0.0011  0.0018   0.0007  2   0.13    0.27  0.38  1.15
0.7    25.0  9.80E+02  0.1197     0.0011  0.0025   0.0007  2   0.06    0.28  0.41  1.1
0.9    25.0  9.80E+02  0.1452     0.0012  0.0032   0.0008  2   0.08    0.25  0.37  1.18
1.1    25.0  9.80E+02  0.1559     0.0018  0.0041   0.0013  2   0.06    0.24  0.35  1.55
1.39   25.0  9.80E+02  0.1537     0.0067  0.0061   0.0018  2   0.02    0.25  0.38  2.67
1.7    25.0  9.80E+02  0.1100     0.0031  0.0049   0.0016  2   0.10    0.2   0.31  3.86
1.9    25.0  9.80E+02  0.0666     0.0120  0.0053   0.0027  2   0.51    0.08  0.15  6.32
2.1    25.0  9.80E+02  -0.0155    0.0053  0.0061   0.0046  2   2.39    0.26  0.45 21.48
2.3    25.0  9.80E+02  -0.0997    0.0071  0.0088   0.0077  2   0.19    0.02  0.05  3.41
2.54   25.0  9.80E+02  -0.1910    0.0041  0.0116   0.0103  2   0.11    0.15  0.25  2.22
2.92   25.0  9.80E+02  -0.3997    0.0090  0.0223   0.0210  2   0.01    0.2   0.33  0.87
```

“.data” file

“.in” file

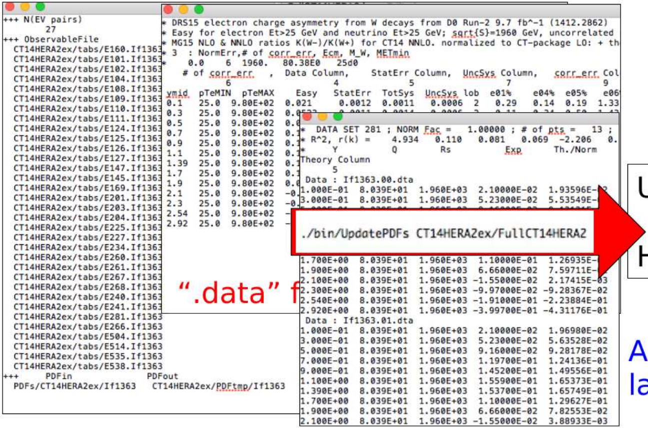
How to use ePump



“.in” file

“.theory” file

How to use ePump



".data" f

Updated best-fit
and
Hessian Error PDFs

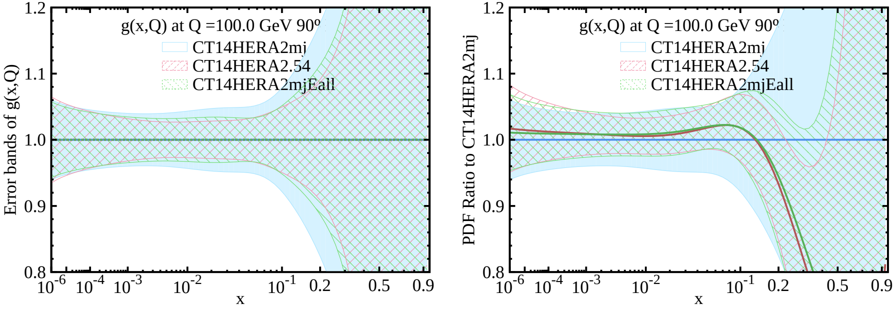
A few seconds
later…

".in" file

".theory" file

Test : CT14HERA2 minus Jets

- Remove all CDF, D0, ATLAS 7TeV, CMS TeV jet data from CT14HERA2 and refit → CT14HERA2mj.
- Add back the 4 data sets to CT14HERA2mj by ePump and compare with CT14HERA2.



gluon - PDF

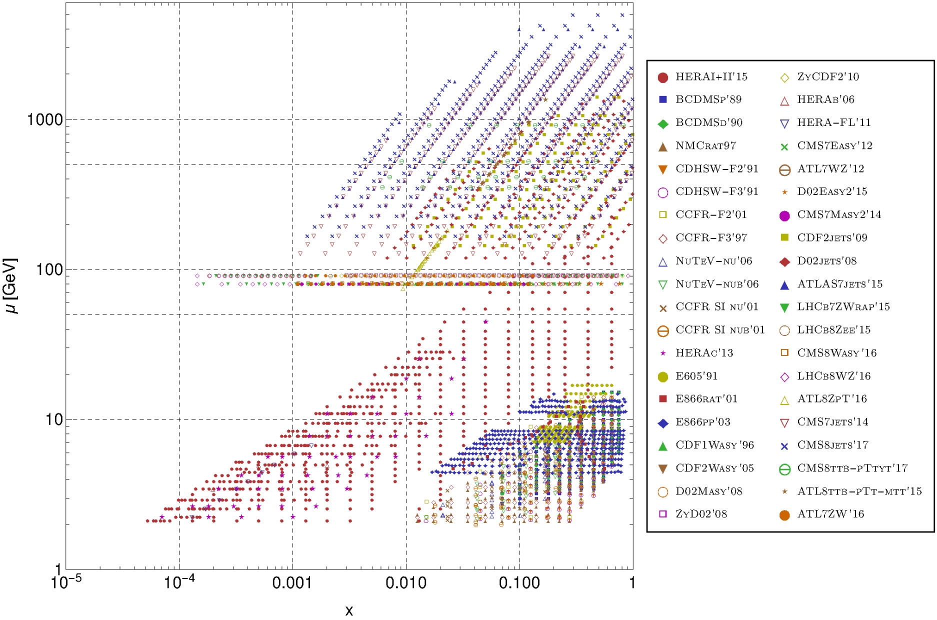
CT18 in a nutshell

- Start with CT14-HERAII (HERAII combined data released after publication of CT14)
- Use as much relevant LHC data as possible
- Using applgrid/fastNLO interfaces to data sets, with NNLO/NLO K-factors, or fastNNLO tables in the case of top pair (single and double differential) data
- Examine a wide range of non-perturbative PDF parameterizations
- Implement a parallelization of the global PDF fitting to allow for faster turn-around time
- Lagrange Multiplier studies to examine constraints of specific data sets on PDF distributions, or on $\alpha_s(m_z)$ and (in some case) the tensions (useful information)

LHC data sets included in CT18

245	1505.07024	LHCb Z (W) muon rapidity at 7 TeV(applgrid)
246	1503.00963	LHCb 8 TeV Z rapidity (applgrid);
249	1603.01803	CMS W lepton asymmetry at 8 TeV (applgrid)
250	1511.08039	LHCb Z (W) muon rapidity at 8 TeV(applgrid)
253	1512.02192	ATLAS 7 TeV Z p_T (applgrid)
542	1406.0324	CMS incl. jet at 7 TeV with R=0.7 (fastNLO)
544	1410.8857	ATLAS incl. jet at 7 TeV with R=0.6 (applgrid)
545	1609.05331	CMS incl. jet at 8 TeV with R=0.7 (fastNLO)
573	1703.01630	CMS 8 TeV $t\bar{t}$ (p_T, y_t) double diff. distributions (fastNNLO)
580	1511.04716	ATLAS 8 TeV $t\bar{t}$ p_T and $m_{t\bar{t}}$ diff. distributions (fastNNLO)
248	1612.03016	ATLAS 7 TeV Z and W rapidity (applgrid) $\to$ CT18Z PDFs

Experimental data in CT18 PDF analysis

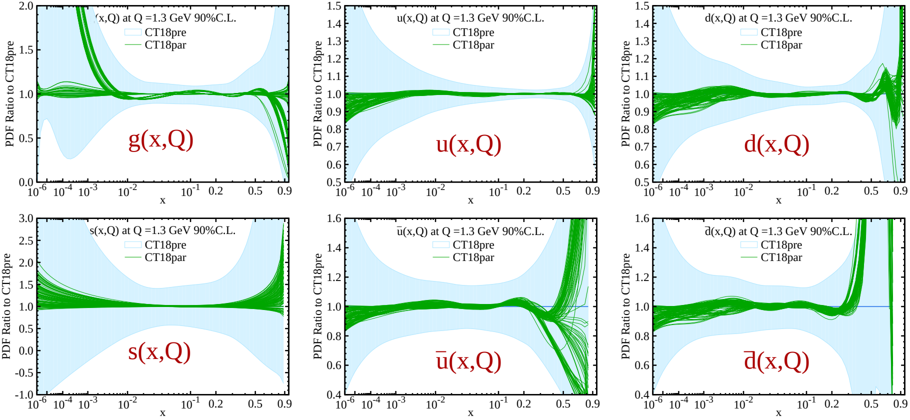


Theory calculations @NNLO

Obs.	Expt.	fast table	NLO code	K-factors	R,F scales
Inclusive jet	ATL 7 CMS 7/8	APPLgrid fastNLO	NLOJet++	NNLOJet	$\mathrm{p_T}, \mathrm{p_T^1}$
$\mathrm{p_T^Z}$	ATL 8	APPLgrid	MCFM	NNLOJet	$\sqrt{\mathrm{Q}^2 + \mathrm{p}^2_{\mathrm{T,Z}}}$
W/Z rapidity W asymmetry	LHCb 7/8 ATL 7 CMS 8	APPLgrid	MCFM/aMCfast	FEWZ/MCFM	$\mathrm{M_{W,Z}}$
DY (low,high mass)	ATL 7/8 CMS 8	APPLgrid	MCFM/aMCfast	FEWZ/MCFM	$\mathrm{Q_{ll}}$
$\mathrm{t\bar{t}}$	ATL 8 CMS 8	fastNNLO			$\frac{\mathrm{H_T}}{4}, \frac{\mathrm{m_T}}{2}$

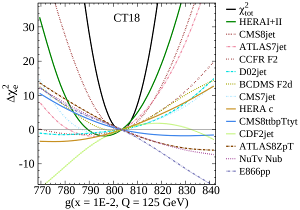
- Studies of QCD scale dependence and other theory uncertainties for DIS, high-p_T Z, jet production
- An uncorrelated error of 0.5% is included for:
  ATLAS 7 TeV and CMS 7/8 TeV jet production, and
  ATLAS 8 TeV high-p_T Z production to account for numerical uncertainties in the MC integration of NNLO cross sections.
- Alternative renormalization/factorization scale choices were examined in high-p_T Z production, do not significantly alter the conclusions.

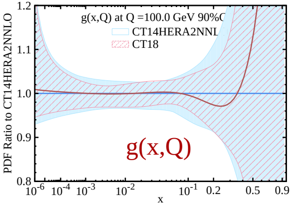
Explore various non-perturbative parametrization forms of PDFs

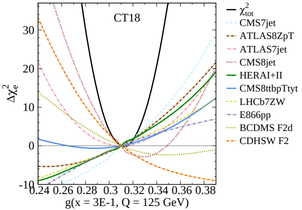


- CT18 – sample result of exploring various non-perturbative parametrization forms.
- There is no data to constrain very large or very small x region.

Preview of CT18 PDFs (g-PDF)



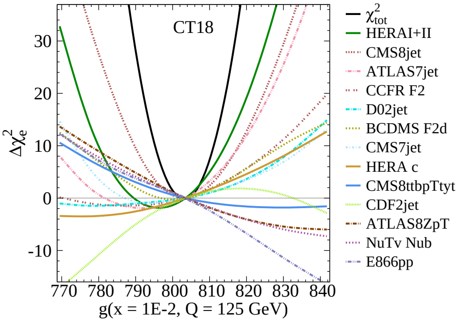


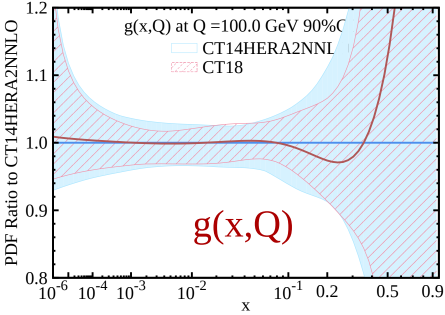


Lagrange Multiplier Scans

- The gluon PDF as $x \to 1$ is parametrization form dependent.

Preview of CT18 PDFs
(g-PDF)

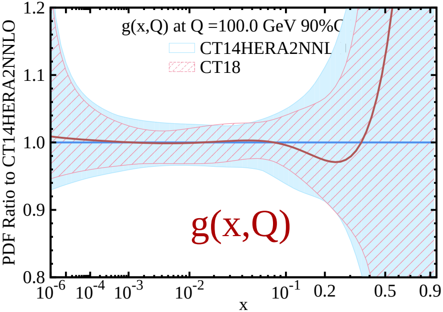


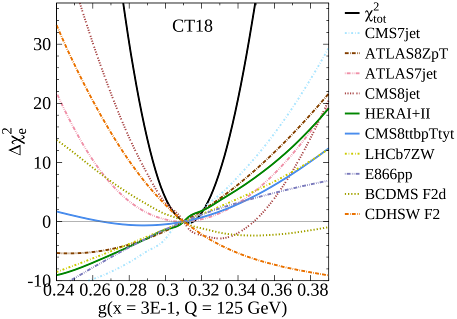


Lagrange Multiplier Scans

- At x around 0.01, ATLAS8 Z p_T data prefer a slightly larger gluon PDF.

Preview of CT18 PDFs
(g-PDF)

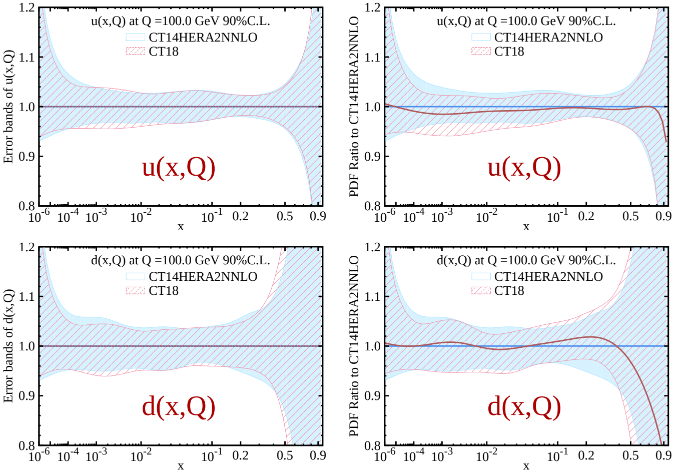




Lagrange Multiplier Scans

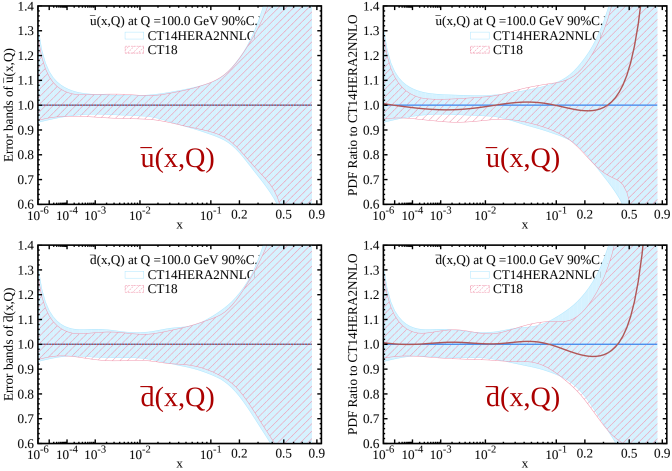
- At x around 0.3, competing with the CDHSW F_2 and Tevatron jet data, which prefer larger gluon, the ATLAS7 jet, CMS7 jet and ATLAS8 Z p_T data prefer a smaller gluon; some tension found in CMS7 and CMS8 jet data.

Preview of CT18 (u-PDF and d-PDF)



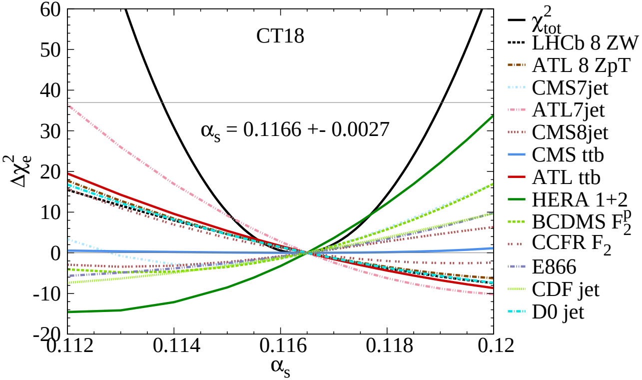
- Some changes on u and d at small x, and d around 0.2; mainly come from LHCb W and Z rapidity data, at 7 and 8 TeV.

Preview of CT18 ($\bar{u}$-PDF and $\bar{d}$-PDF)



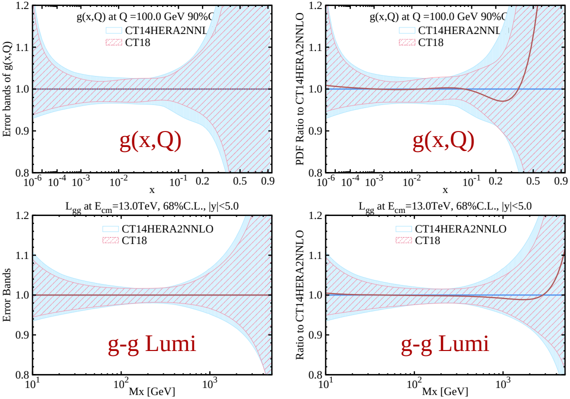
- Minor changes on ubar and dbar PDFs at small x region mainly come from LHCb W and Z rapidity data, at 7 and 8 TeV.
- The behavior of ubar and dbar PDFs, as $x \to 1$, is parametrization form dependent.

$\alpha_s(M_z)$



- The fixed target F_2 data and HERA DIS data prefer smaller α_s value.
- The ATLAS 8TeV Z p_T, ATLAS 7 TeV incl. jet data, bring the central value of $\alpha_s(M_z)$ from $0.115^{+0.006}_{-0.004}$ (CT14) to 0.1166 ± 0.0027 (CT18).

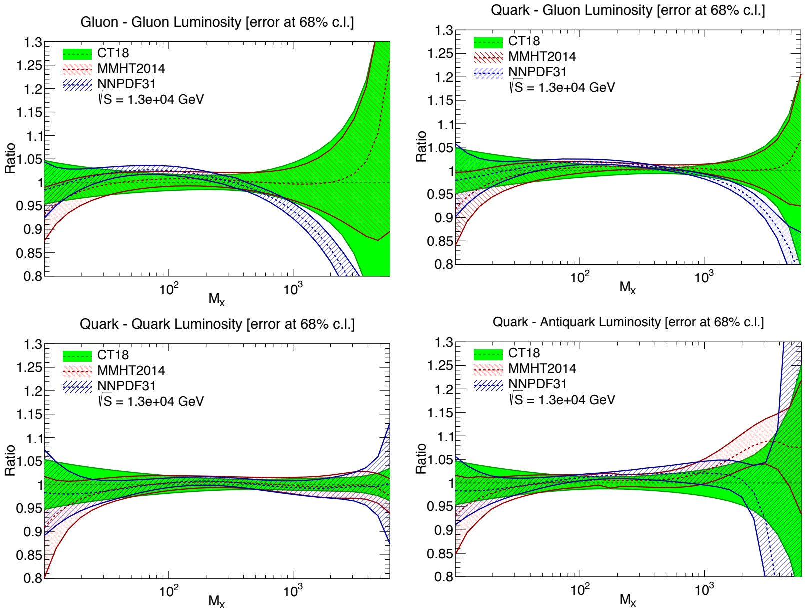
$\sigma(gg \to H)$ CT18 v.s. CT14



	$\sigma(gg \to h)$ 7 TeV
CT14	14.67 $\pm$0.46
CT18	14.57 $\pm$0.44
	$\sigma(gg \to h)$ 8 TeV
CT14	18.70 $\pm$0.57
CT18	18.45 $\pm$0.55
	$\sigma(gg \to h)$ 13 TeV
CT14	42.78 $\pm$1.32
CT18	42.43 $\pm$1.26
	$\sigma(gg \to h)$ 14 TeV
CT14	48.23 $\pm$1.50
CT18	47.91 $\pm$1.42

- PDF induced errors (at $90\% C.L.$) are reduced by about 5% as compared to CT14 predictions.

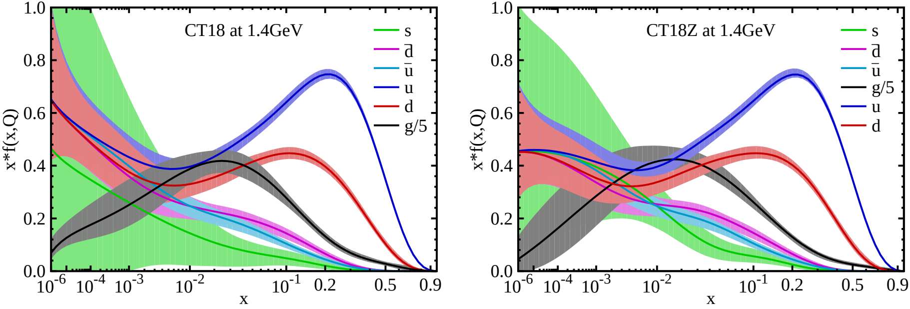
PDF Luminosities at 13 TeV LHC CT18, MMHT14 and NNPDF3.1



CT18Z
LHC data treatment

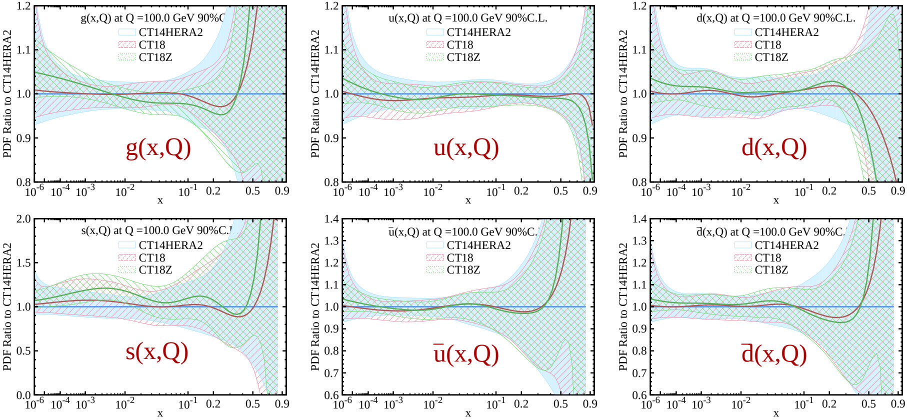
- Start with CT18 data set
- Add in ATLAS 7 TeV W and Z rapidity data (arXiv:1612.03016; $4.6 fb^{-1}$); large $\chi^2/d.o.f \sim 2.1$
- Remove CDHSW data
- Use a special x-dependent factorization scale mDIS,x at NNLO calculation
- CT18Z uses a combination of mDIS,x (preferred by DIS) and an increased $m_c^{pole} = 1.4$ GeV (preferred by LHC vector boson production, disfavored by DIS)

PDF uncertainty bands
CT18 vs. CT18Z



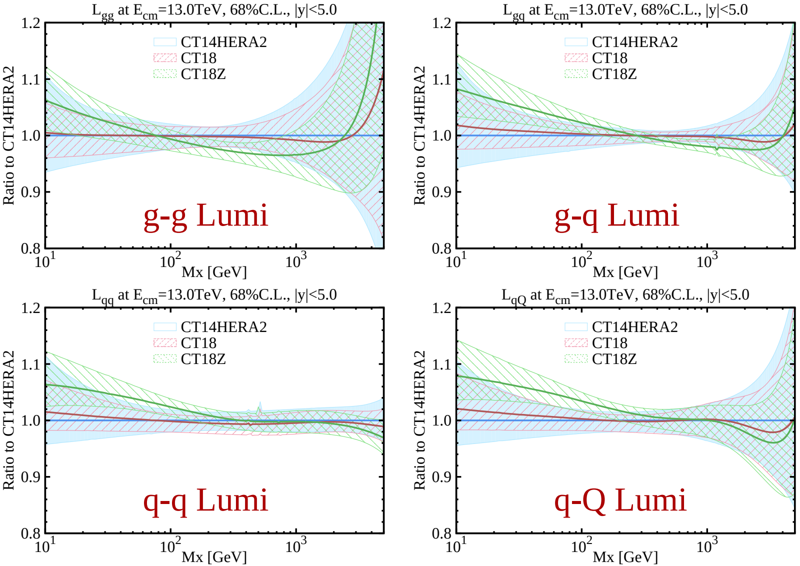
CT18Z has enhanced gluon, u-, d- and s-PDFs at $x \sim 10^{-4}$, and reduced g-PDFs at $x > 10^{-2}$. The CT18Z fit is performed so as to maximize the differences from CT18 PDFs, while preserving about the same goodness-of-fit as for CT18 analysis.

CT18Z vs.CT18 PDFs

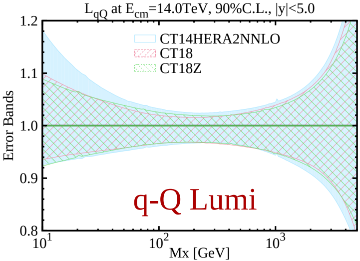


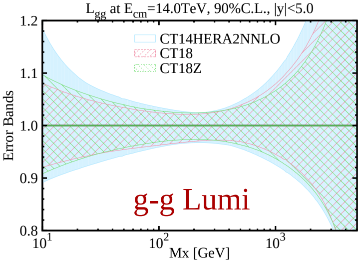
- G increases at small-x, and decreases at $x \sim 0.010.3$
- u and d increase at small-x
- d increases at $x \sim 0.2 - 0.3$
- s increases at small-x

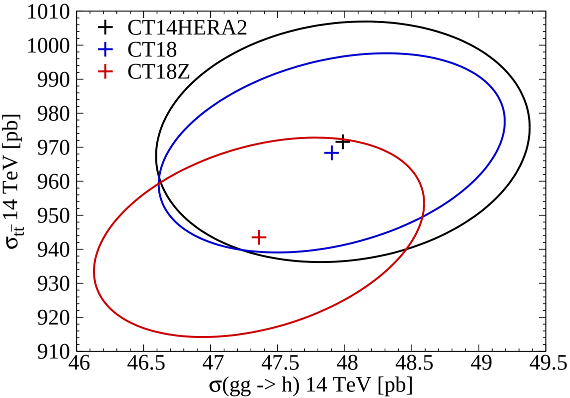
PDF Luminosities at 13 TeV LHC
CT14HERA2, CT18 and CT18Z



Mild reduction in nominal PDF error bands and cross section uncertainties







Summary

- Parton distribution function(PDF) is a necessary ingredient for a hadron collision process.
- Currently, PDF is determined by global analysis.
- Study on PDF is not just a window to let us know about the inner stucture of hadron, but also the fundation to probe new physics.
- New generation PDF CT18 and CT18Z is very close to publish.