Bread & Butter Physics at the LHC

C.-P. Yuan
Michigan State University
Wu-Ki Tung Endowed Professor

May 8, 2023 @ U of Pittsburgh

Pheno 23 Symposium

In collaboration with CTEQ-TEA members

CTEQ – Tung et al. (TEA)
in memory of Prof. Wu-Ki Tung

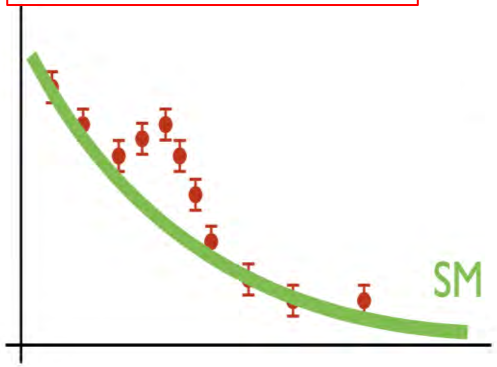
C T E Q

What is the bread-and-butter physics at the LHC?

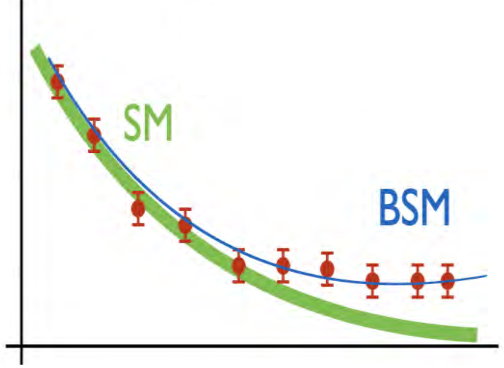
The bread and butter of a situation or activity is its most *basic* or *important* aspects. --- Dictionary

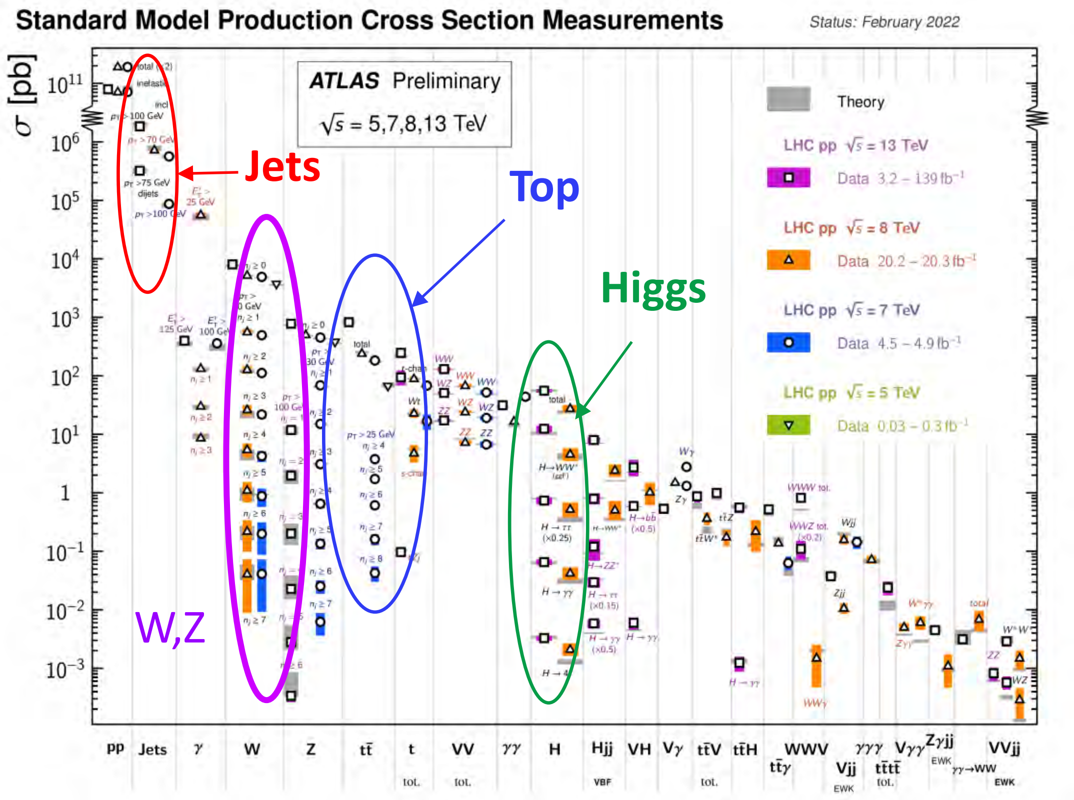
- **Goals:** 1. Test Standard Model (SM)
  2. Find New Physics (NP)

Resonance search



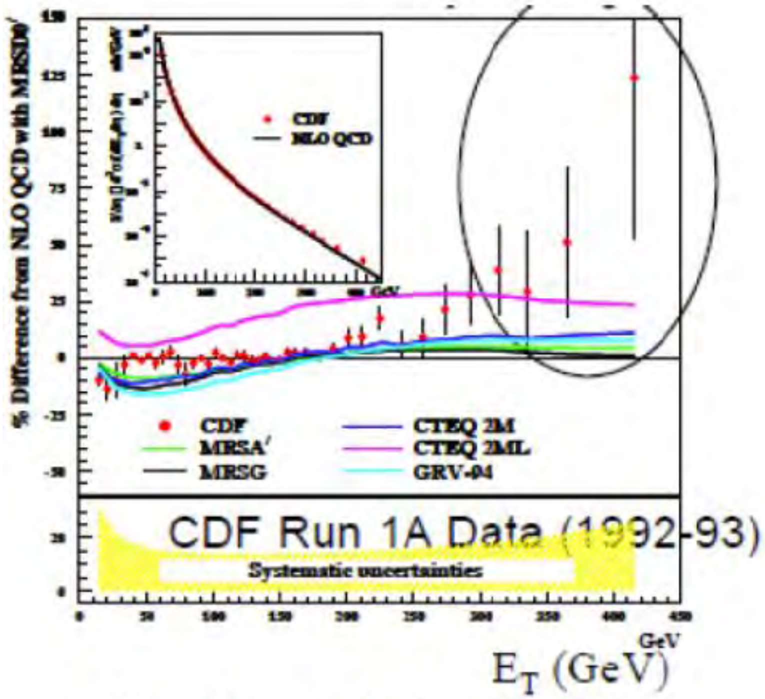
Deviation search





New Physics Found (in 1996) ?

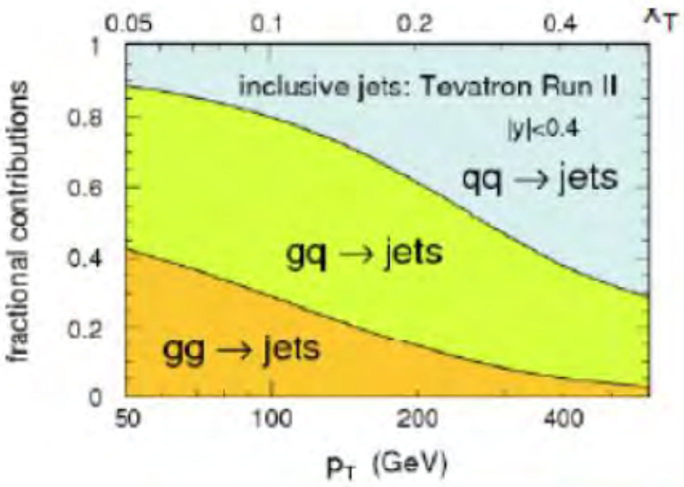
C T E Q

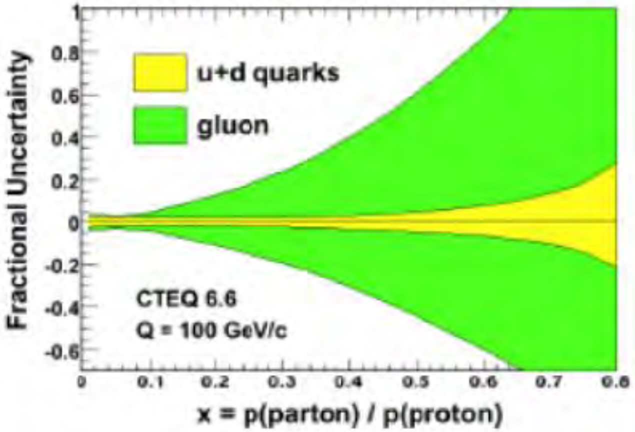


Phys. Rev. Lett. 77, 438 (1996)

High-x gluon not well known

...can be accommodated in the Standard Model



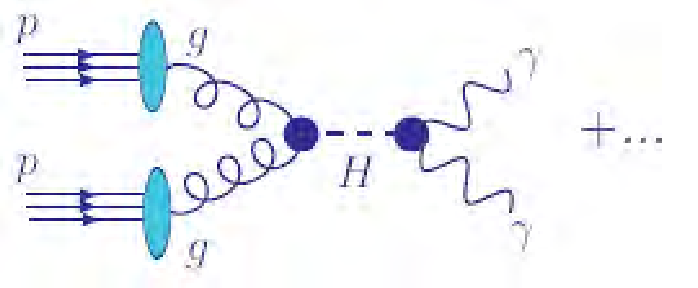


Explained by having better determined PDFs from global analysis;
no need for NP scenario yet.

J. Huston, E. Kovacs, S. Kuhlmann, J.L. Lai, J.F. Owens, D. Soper, W.K. Tung , Phys. Rev. Lett. 77 (1996) 444.

QCD Factorization Theorem and PDFs

$$\sigma_{pp\to H\to\gamma\gamma X}(Q) = \sum_{a,b=g,q,\bar{q}} \int_0^1 d\xi_a \int_0^1 d\xi_b \hat{\sigma}_{ab\to H\to\gamma\gamma}\left(\frac{x_a}{\xi_a}, \frac{x_b}{\xi_b}, \frac{Q}{\mu_R}, \frac{Q}{\mu_F}; \alpha_s(\mu_R)\right)$$
$$\times f_a(\xi_a,\mu_F) f_b(\xi_b,\mu_F) + O\left(\frac{\Lambda_{QCD}^2}{Q^2}\right)$$



$\hat{\sigma}$ is the hard cross section; computed order-by-order in $\alpha_s(\mu_R)$
$f_a(x,\mu_F)$ is the distribution for parton a with momentum fraction x, at scale μ_F

$$f_{a/h}(x,Q)$$

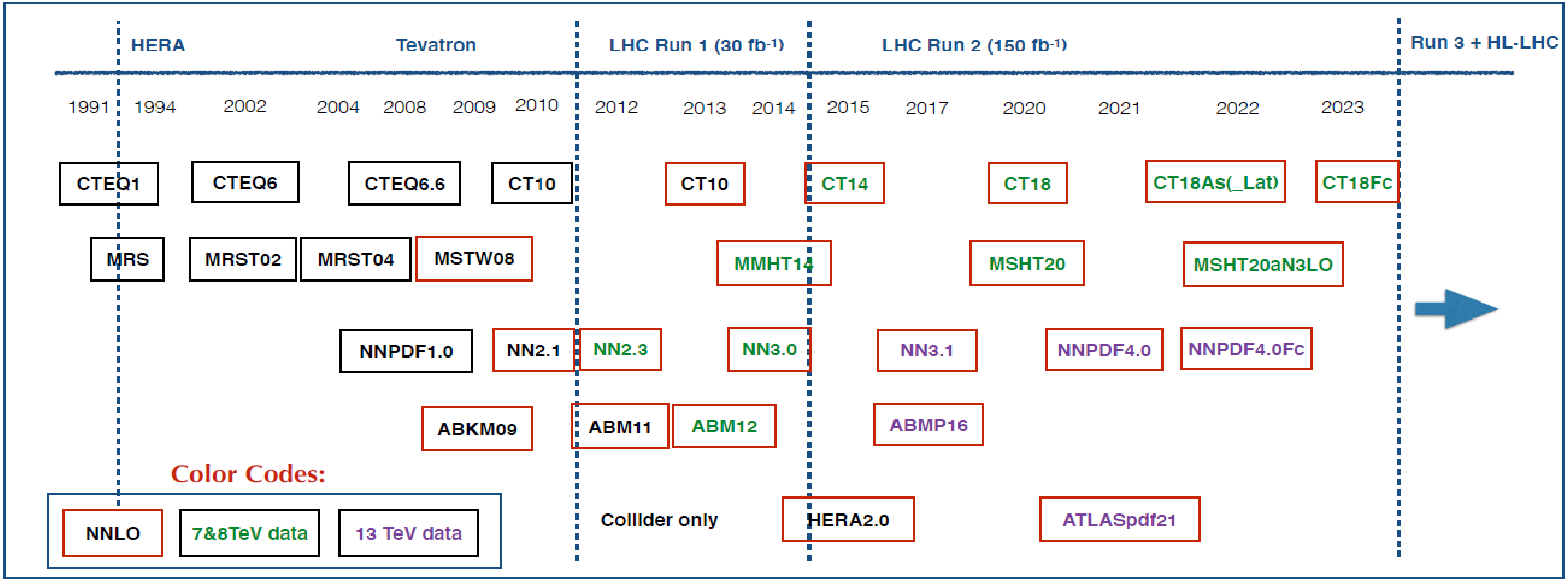
Unpolarized collinear parton distribution functions (PDFs) $f_{a/h}(x,Q)$ are associated with probabilities for finding a parton a with the "+" momentum xp^+ in a hadron h with the "+" momentum p^+ for $p^+ \to \infty$, at a resolution scale $Q >$ 1 GeV .

The (unpolarized) collinear PDFs describe long-distance dynamics of (single parton scattering) in high-energy collisions.

Some PDF sets of QCD global analysis

Jun Gao, DIS 2022

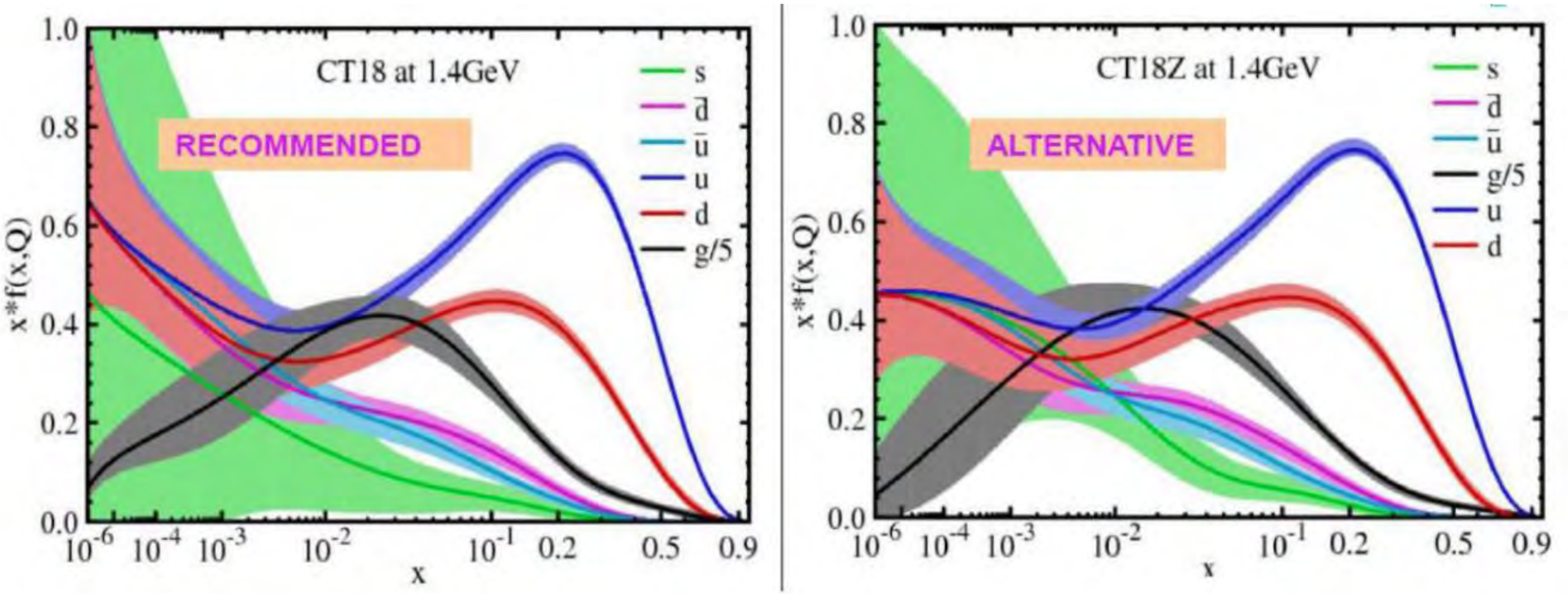
✦ PDFs provided by several major analysis groups (CT, MSHT, NNPDF, ABM, HERAPDF, ATLASpdf, CJ, JAM…) using slightly different heavy-quark schemes, selections of data, and methodologies



CT18 family PDFs

https://ct.hepforge.org/PDFs/ct18/

arXiv:1912.10053



*** CT18 (N)NLO PDF set is recommended for the majority of LHC applications**

- CT18Z has enhanced gluon and strange PDFs at $x \sim 10^{-4}$, and reduced light-quark PDFs at $x < 10^{-2}$. The CT18Z maximizes the differences from CT18 PDFs, while preserving about the same goodness-of-fit as for CT18.
- CT18A and CT18X include some features of CT18Z, lie between CT18 and CT18Z.

CT18
CT18A
CT18X
CT18Z
CT18_as

NNLO
NLO

Extensions

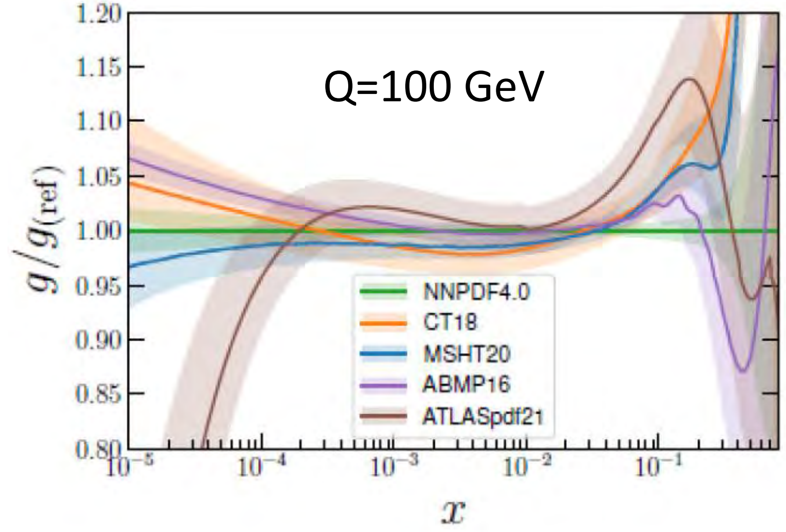
CT18qed
CT18As
CT18As_Lat
CT18FC
CT18LO

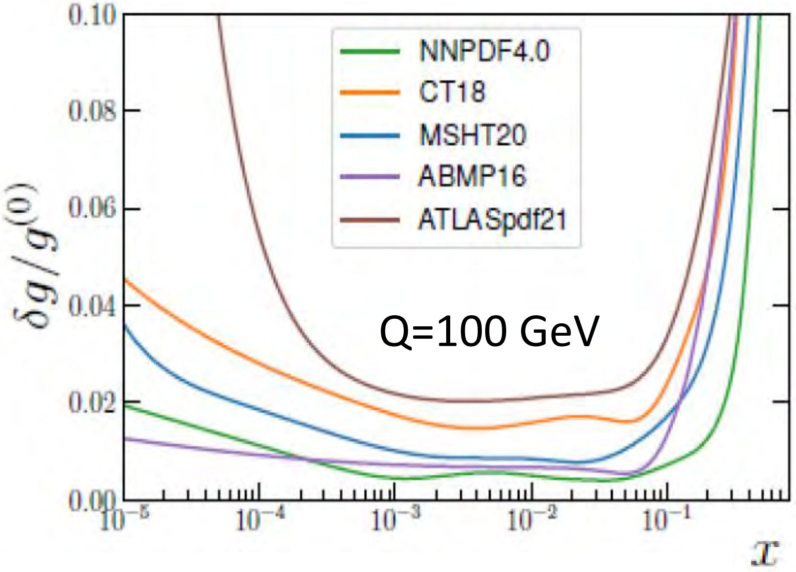
Comparing predictions from various QCD global analysis groups

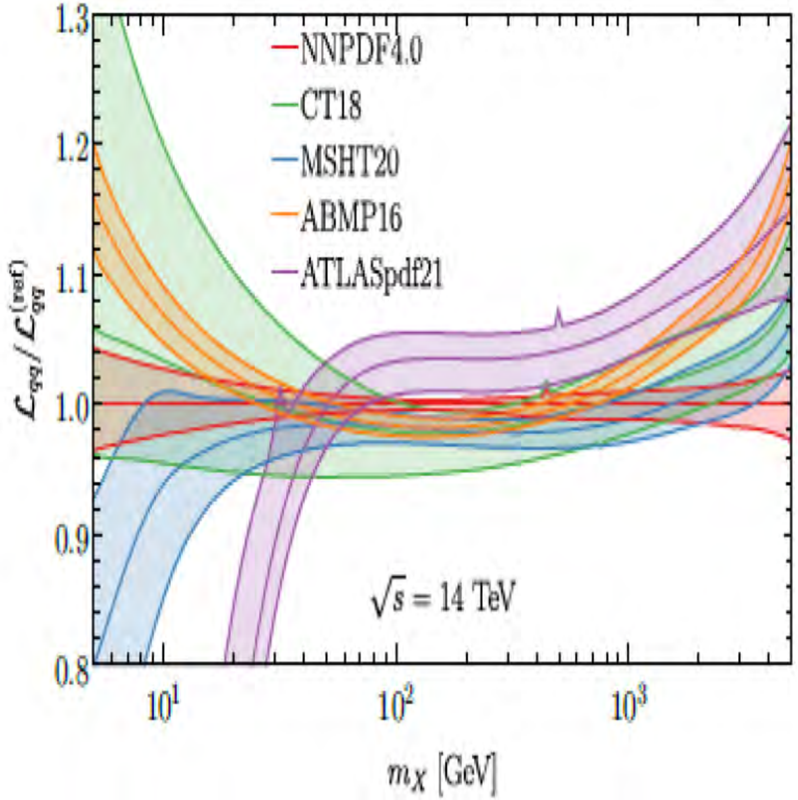
Snowmass 2021, 2203.13923

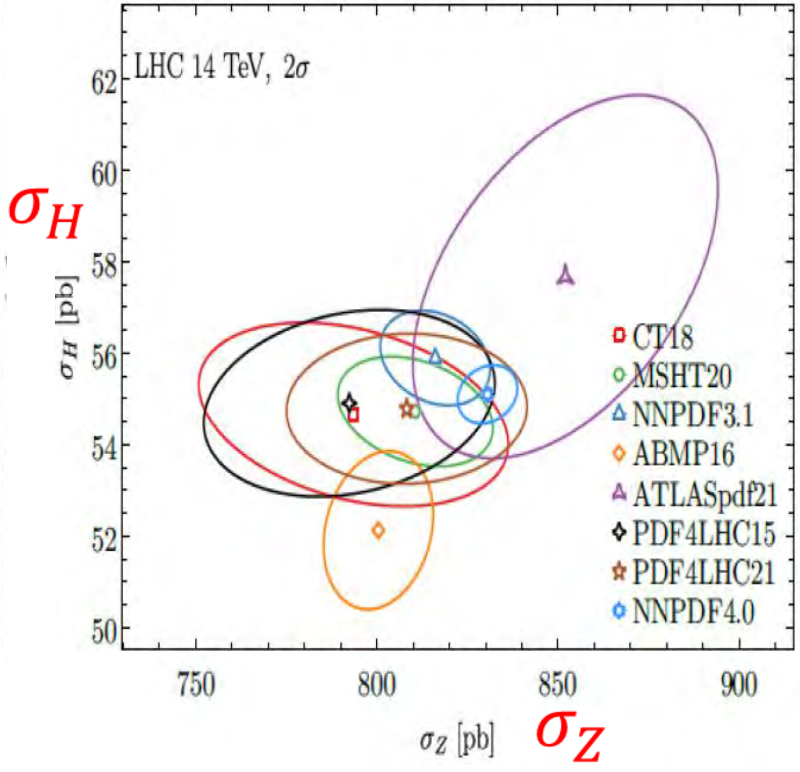
Smaller PDF errors lead to smaller

PDF luminosity errors, then smaller PDF-induced errors in cross sections.







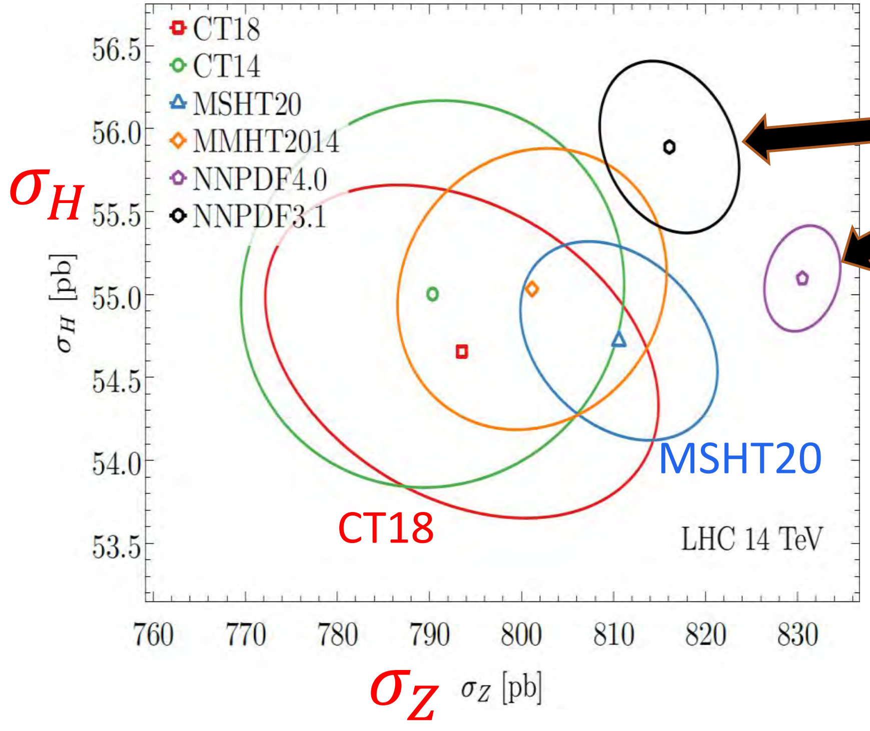


Comparing predictions from various QCD global analysis groups

Snowmass 2021, 2203.13923

The PDF-induced errors @ 68% CL in $gg \to h$ and $q\,\bar{q} \to Z$ NNLO cross sections

Due to different choices of



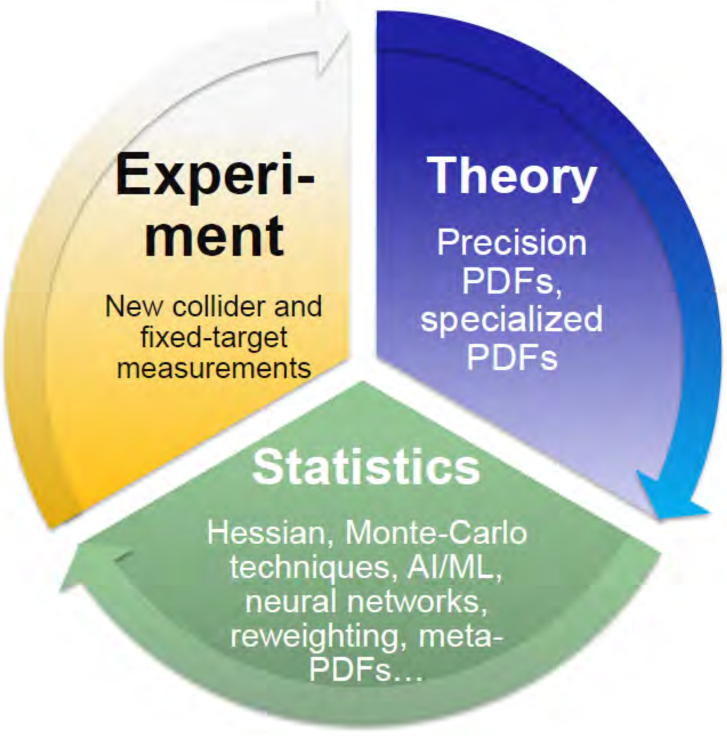
NNPDF3.1

NNPDF4.0

Their predictions do not overlap at 1σ level.

Different (though mostly consistent) predictions on

- central values and error estimates of PDFs,
- parton luminosities,
- physical cross sections, and
- various correlations among PDFs and data …



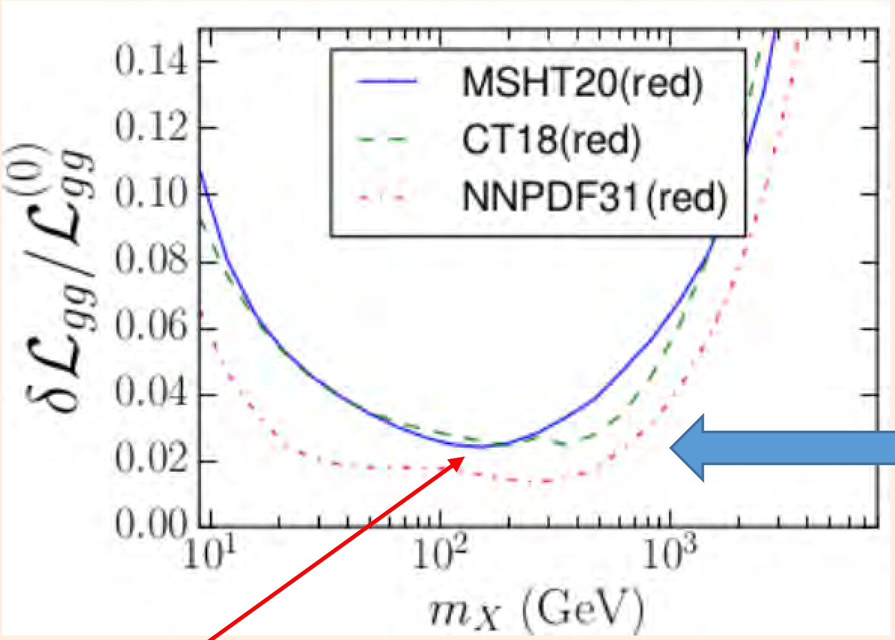
Components of a global QCD fit

Benchmark Study: PDF4LHC21

arXiv:2203.05506

Relative PDF uncertainties on the gg luminosity at 14 TeV in three PDF4LHC21 fits to the **identical** reduced global data set

arXiv:2203.05506



× 1.5 – 2 difference

- Each analysis group (CT, MSHT, NNPDF) used the same (reduced) data sets and same theory predictions in the analysis

Smaller error size found by NNPDF

- NNPDF3.1' and especially 4.0 (based on the NN's+ MC technique) tend to give smaller uncertainties in data-constrained regions

The size of PDF error estimates depends on the methodology of global analysis adopted by the PDF fitting group.

Sources of PDF errors

Factorization Theorem:

Data = PDFs ⊗ Hard part cross sections (Wilson coeff.)

Experimental errors:

- Statistical
- Systematic
  - uncorrelated
  - correlated
- χ^2 definition (experimental or t_0)
- Possible tensions among data sets

Extracted with errors, dependent of methodology of analysis

- Non-perturbative parametrization forms of PDFs
- Additional theory prior
- Choice of Tolerance (T^2) value

Theoretical errors:

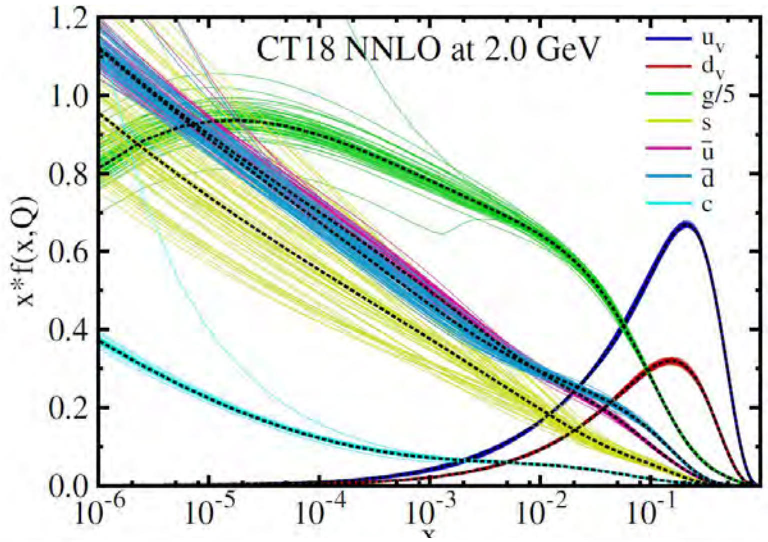
- Which order: (NLO, NNLO, …, resummation – BFKL, qT, threshold)
- Which scale: (μ_F, μ_R)
- Which code: (antenna subtraction, sector decomposition,…, qT, N-jettiness,,…)
- Monte Carlo error: (most efficient implementation,…)

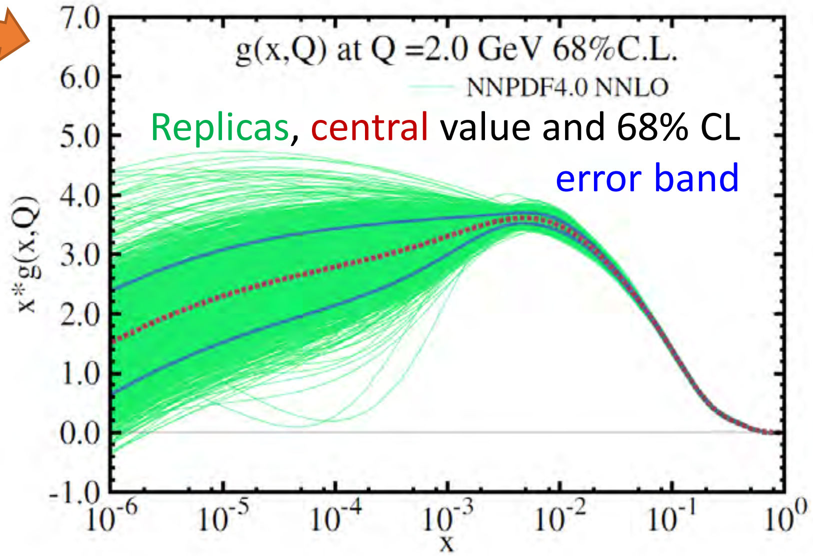
How to estimate PDF errors in QCD global analysis

- Error estimate is important.
- Two different methodology in global analysis
- Hessian PDF eigenvector (EV) sets, from analytic parametrizations of PDFs
  → (ABM, CTEQ, HERA, MSHT, …)
- Monte Carlo (MC) PDF replicas, from Neural Network (NN) parametrizations
  → (NNPDF)

- Both methods assume some non-perturbative input of PDFs at the initial Q_0 scale, around 1 GeV. (analytical parametrization vs. NN architecture)

- They are two powerful and complementary representations.
- Hessian PDFs can be converted into MC ones, and vice versa.





How to quantify PDF uncertainties

was first introduced in 2001 by

Jon Pumplin, Dan Stump and Wu-Ki Tung

@ Michigan State University

hep-ph/0101032

Uncertainties of predictions from PDFs:

The Hessian method

hep-ph/0101051

Uncertainties of predictions from PDFs:

The Lagrange multiplier method

$$\chi^2 = \chi_0^2 + \sum_{i,j} H_{ij} (a_i - a_i^0)(a_j - a_j^0)$$

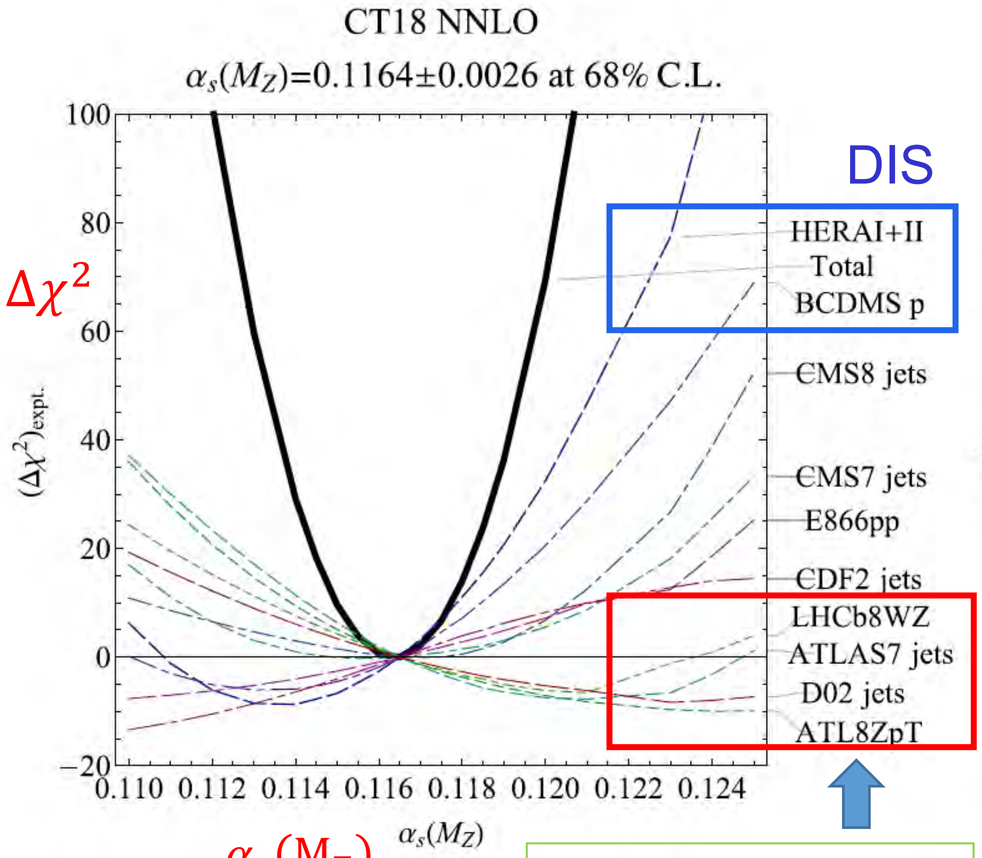
It was first implemented in CTQE6 PDFs.

They were used to determine uncertainty of PDFs, physical cross sections, α_s and m_t as well as exploring tensions among data sets in the CTEQ-TEA analysis.

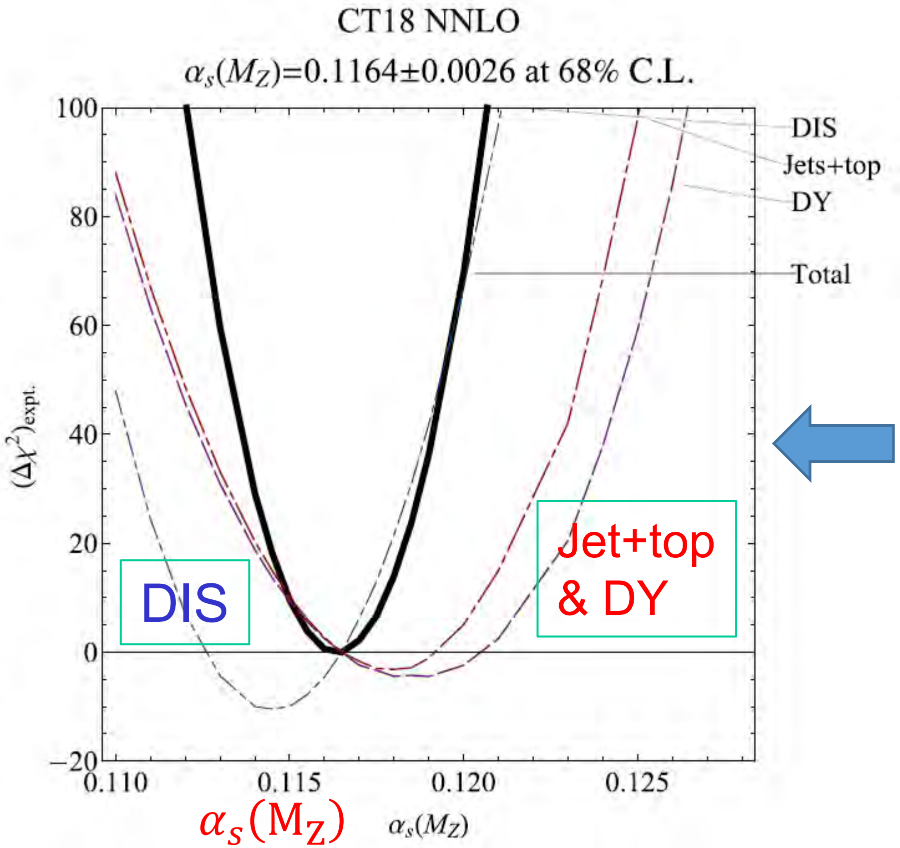
Lagrange Multiplier scan

To explore PDF-induced errors in the determination of $\alpha_s(\mathrm{M_Z})$ and tensions among data sets included in the fit

arXiv:1912.10053



Prefer larger $\alpha_s(M_Z)$



- The opposing pulls (i.e., tensions) of DIS and jet+top&DY experiments significantly exceed $\Delta\chi^2 = 1$ variation, as implied by the simplest statistical framework.
- Require a large value of Tolerance T^2, the maximum allowed total $\Delta\chi^2$, with $\Delta\chi^2 > 1$

The scan of $\alpha_s(\mathrm{M_Z})$ values in CT18 NNLO PDF analysis.

Tolerance (T^2) values in various PDF analysis groups

CTEQ

- Tolerance T^2, the maximum allowed total $\Delta\chi^2$ value away from the best (or central) fit, was introduced to account for the sampling of
  - non-perturbative parametrization of PDFs (or NN architecture, smoothness, positivity) and
  - the allowed PDF variation due to various choices of data sets and theory calculations, etc.

- Roughly speaking, at the 68% CL,
  - CTEQ-TEA (CT) Tier-1 $T^2 \sim 30$
  - MSHT dynamical $T^2 \sim 10$
  - NNPDF effective $T^2 \sim 2$ (for MC replicas and their Hessian representation)

- A smaller T^2 value typically yields a smaller PDF error estimate.

CT tolerance includes both Tier-1 and Tier-2 contributions.

To reduce PDF uncertainty, one must maximize both

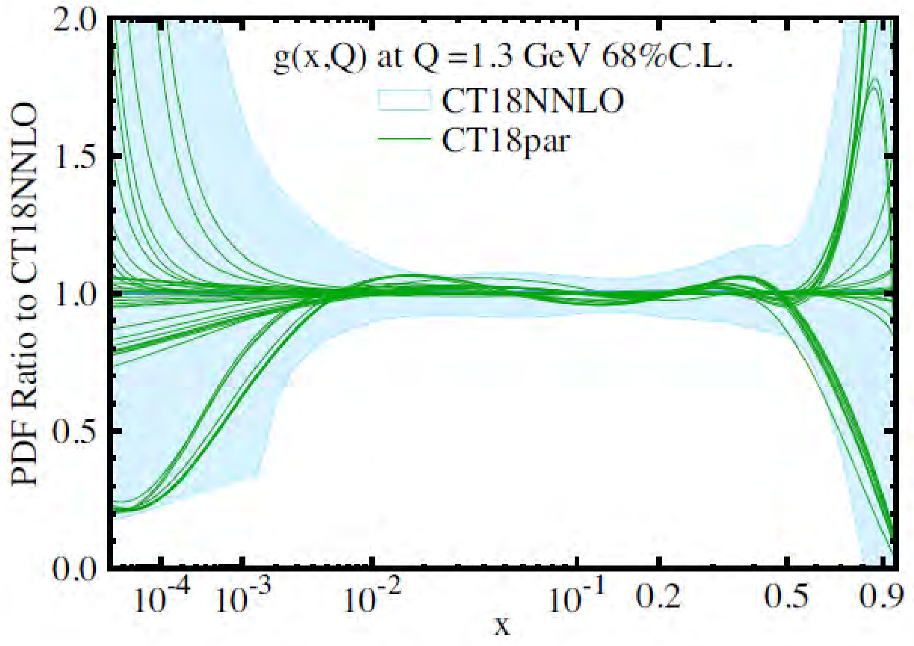
PDF fitting accuracy
(accuracy of experimental, theoretical and other inputs)

and

PDF sampling accuracy
(adequacy of sampling in space of possible solutions)

Large Tolerance value T^2 in the CT fits

$\Delta\chi^2 \gg 1$



A large part of the CT18 PDF uncertainty accounts for the sampling over

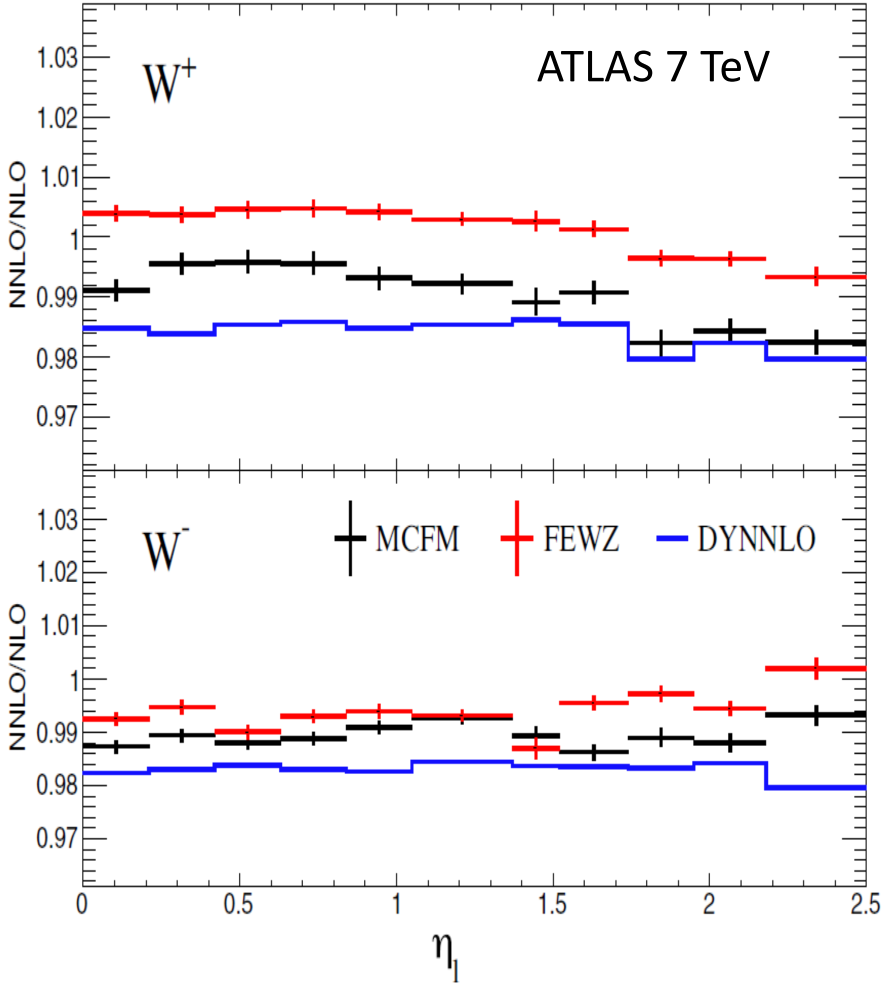
- 250-350 parametrization forms of PDFs,
- possible choices of fitted experiments, definitions of χ^2,
- theory predictions, and
- analysis method (Hessian, LM, MC)

Total $\Delta\chi^2 \gg 1$ @ 68% CL

- In CT fits, we estimate that varying non-perturbative inputs at the initial scale of the PDFs (at 1.3 GeV) could contribute to T^2 around 10 to 15 units @ 68% CL.
- Hence, CT PDF errors are defined with T^2=100 @ 90% CL, or equivalently, T^2=100/$(1.645)^2 = 37$ @ 68% CL
- A more advanced CT tolerance prescription is under development.

Different (NNLO) theory predictions from various codes; require $\Delta\chi^2 > 1$

arXiv:1912.10053



- Compare predictions of three different codes:
  - FEWZ (sector decomposition)
  - MCFM (N-jettiness)
  - DYNNLO (qT)
- Their predictions agree well at NLO.
- Their NNLO predictions agree well for inclusive cross sections (without imposing kinematic cuts).
- Their NNLO predictions for fiducial cross sections (with kinematic cuts) can differ at percent level, while the statistical error of the data is at the sub-percent level.

✓ The resulting PDFs from various theory predictions only differ slightly, when including this data in the CT18A fit.
✓ The kind of theory uncertainty is accounted for by choosing a larger Tolerance value than 1 (i.e., $\Delta\chi^2 > 1$) at the 68% CL.

Some data requires all-order (resummation) calculations

- When applying a symmetric p_T cut (with same magnitude) on the decay leptons of inclusive W or Z boson production, the two leptons are almost back-to-back, decaying from a low p_T gauge boson.
- Fixed order predictions cannot correctly predict the low p_T distribution of W or Z.
- It requires a resummation calculation, such as ResBos, to resum all the large logs arising from multiple soft-gluon radiation.

What's QCD Resummation?

- Perturbative expansion

$$\frac{d\hat{\sigma}}{dq_T^2} \sim \alpha_s \left\{ 1 + \alpha_s + \alpha_s^2 + \cdots \right\}$$

- The singular pieces, as $\frac{1}{q_T^2}$ (1 or log's)

$$\frac{d\hat{\sigma}}{dq_T^2} \sim \frac{1}{q_T^2} \sum_{n=1}^{\infty} \sum_{m=0}^{2n-1} \alpha_s^{(n)} \ln^{(m)} \left(\frac{Q^2}{q_T^2} \right)$$

$$\sim \frac{1}{q_T^2} \{ \alpha_s (\underline{L} + \underline{1}) + \alpha_s^2 (\underline{L^3} + \underline{L^2} + \underline{L+1}) + \alpha_s^3 (\underline{L^5} + \underline{L^4} + \underline{L^3} + \underline{L^2} + \underline{L+1}) + \cdots \}$$

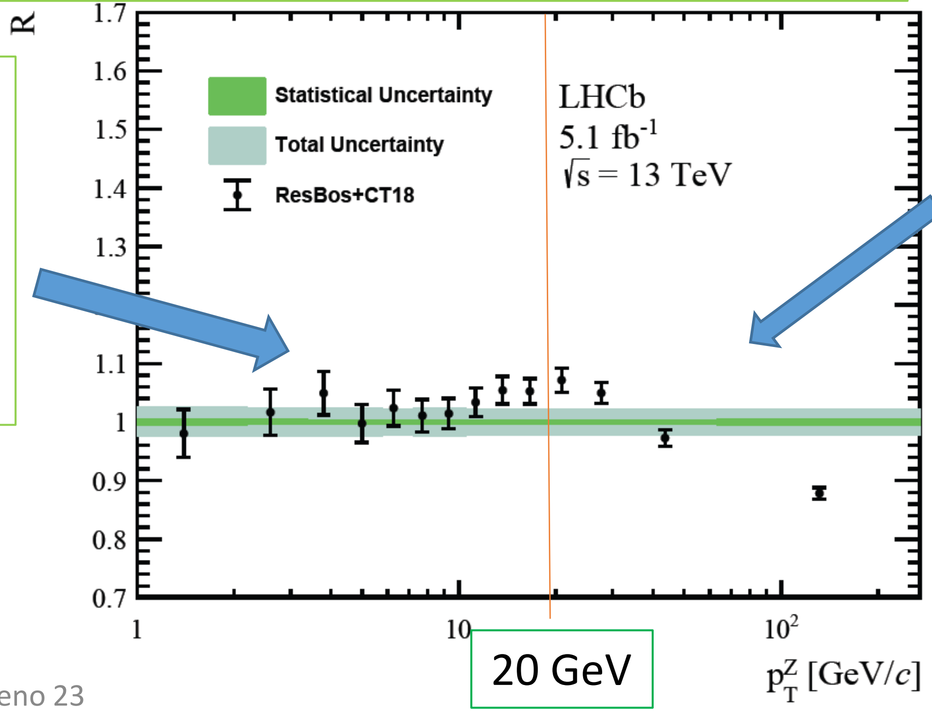
$$L \equiv \ln\left(\frac{Q^2}{q_T^2}\right)$$

Resummation is to reorganize the results in terms of the large Log's.

ResBos + CT18 can describe well low $p_T(Z)$ region, with $p_T(Z) < 20$ GeV

$$\alpha_s \ln\left(\frac{Q^2}{q_T^2}\right) \sim 1$$

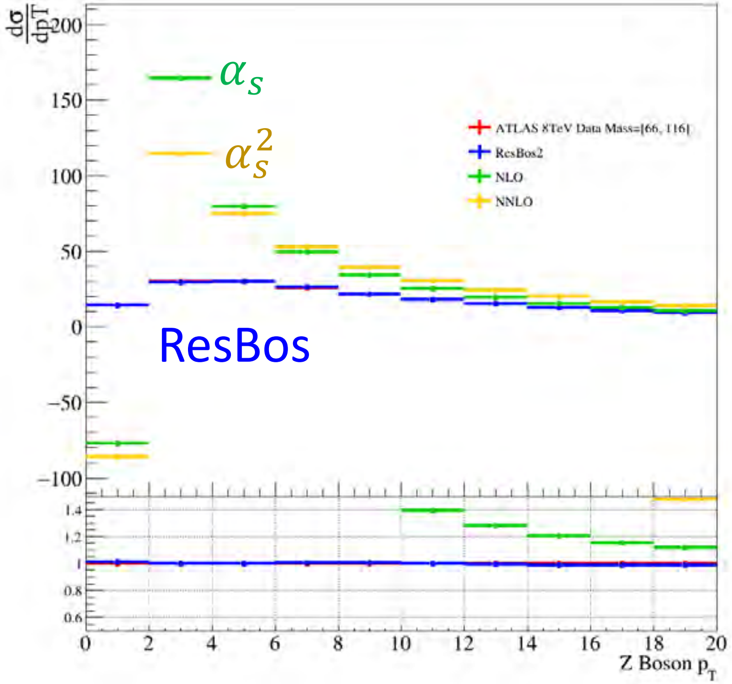
Compare to LHCb 13 TeV Z data; arXiv:2112.07458

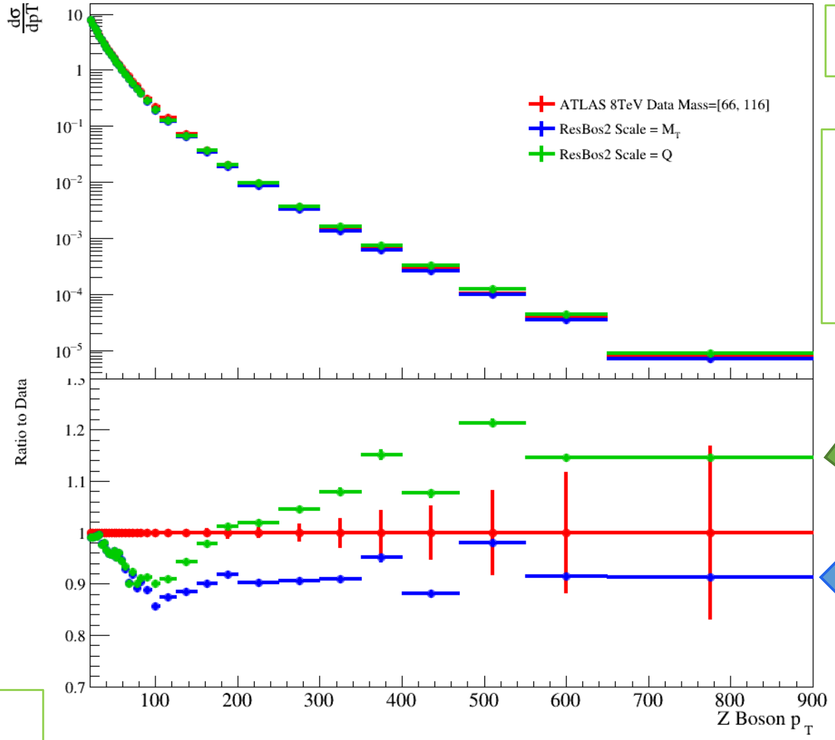


High $p_T(Z)$ region needs α_s^3 contrinution

Some data requires all-order (resummation) calculations: ResBos

Compare to ATLAS 8 TeV Z data; arXiv:1606.00689





https://gitlab.com/resbos2

arXiv:2205.02788

- Sensitive to scale choices at α_s^2
- High $p_T(Z)$ region requires yet higher order (α_s^3) contribution.

Use $\mu_F = \mu_R = \mathrm{Q}$
Invariant mass, at α_s^2

Use $\mu_F = \mu_R = \mathrm{m_T}$
where $m_T = \sqrt{Q^2 + p_T^2}$
Transverse mass, at α_s^2

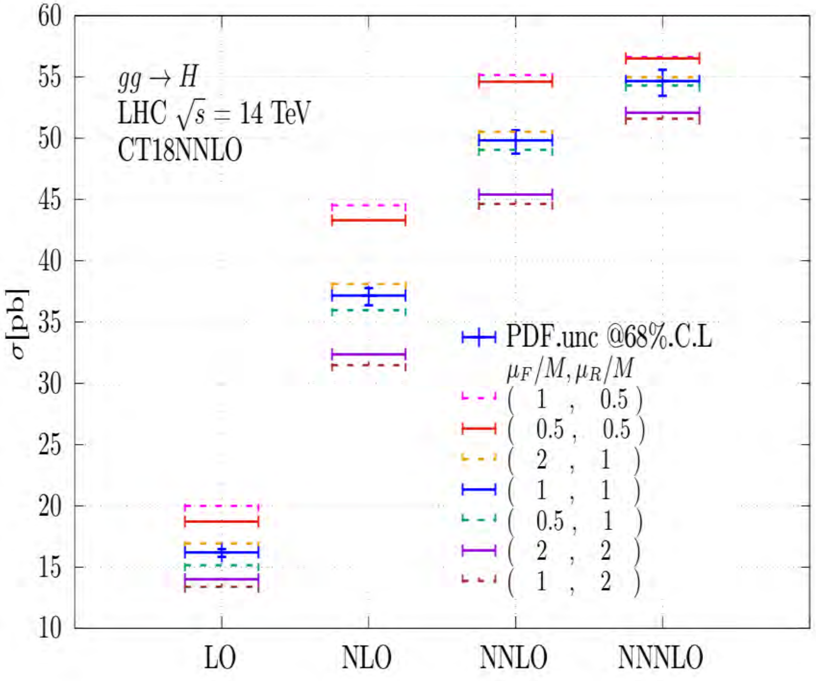
The low p_T Z data, with $p_T(Z) < 20$ GeV, can be described well by ResBos, but not fixed order (NLO, NNLO,…) calculations which yield singular result as $p_T(Z) \to 0$.

Require higher (fixed) order calculations for $p_T(Z) > 20$ GeV; α_s^3 correction increases the rate by about 10% when using the scale m_T and renders a good agreement with data.

Missing higher order (MHO) uncertainty estimated by scale variation

- General wisdom: Varying a "typical scale" by a factor of 2 (or 7-point scales) to estimate missing higher order (MHO) contribution.
- This wisdom does not always work. Namely, varying the factorization and normalization scales by a factor of 2 cannot accurately estimate MHO contribution.

The complete higher order calculations in QCD, EW, and the mixed QCD+EW are all very important for making precision theory prediction to compare to precision experimental data in order to extract precision PDFs.



Tools : ggHiggs(Marco Bonvini)

$\sigma(gg \to H)$ at 14 TeV LHC

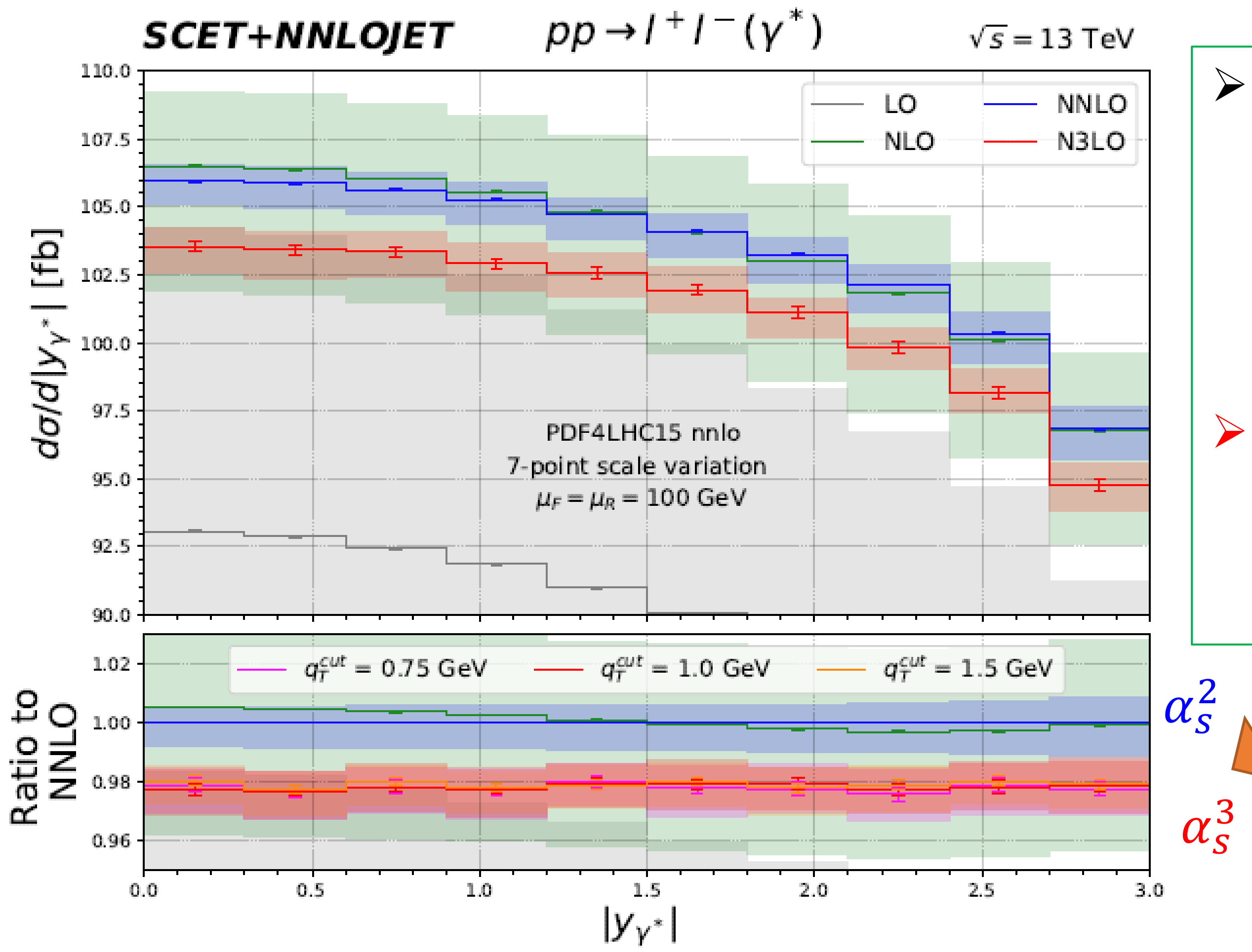
7-point scale variation at N3LO in QCD for $m_t = 172.5$ GeV and $M = m_H = 125$ GeV

μ_R/M \ μ_F/M	0.5	1	2
0.5	3.4%	3.6%	-
1	-0.6%	-	0.6%
2	-	-5.6%	-4.7%

- The K-factor of electroweak (EW) correction is about 1.05
- The PDF uncertainty is about 2.8%

Estimating missing higher order contribution via varying μ_f and μ_R scales

arXiv:2107.09085



- Varying the factorization μ_f and renormalization μ_R scales by a factor of 2 around their nominal values (with 7-point scale variation) does not always lead to a good estimate of missing higher order (MHO) effect in the perturbative calculation.
- The N3LO correction is outside the scale variation band predicted at NNLO, due to accidental cancellation among various partonic subprocess contributions.

This comparison does not include PDF and α_s induced errors.

Finding the faithful PDF uncertainty on QCD cross sections: Hopscotch scans

arXiv: 2205.10444

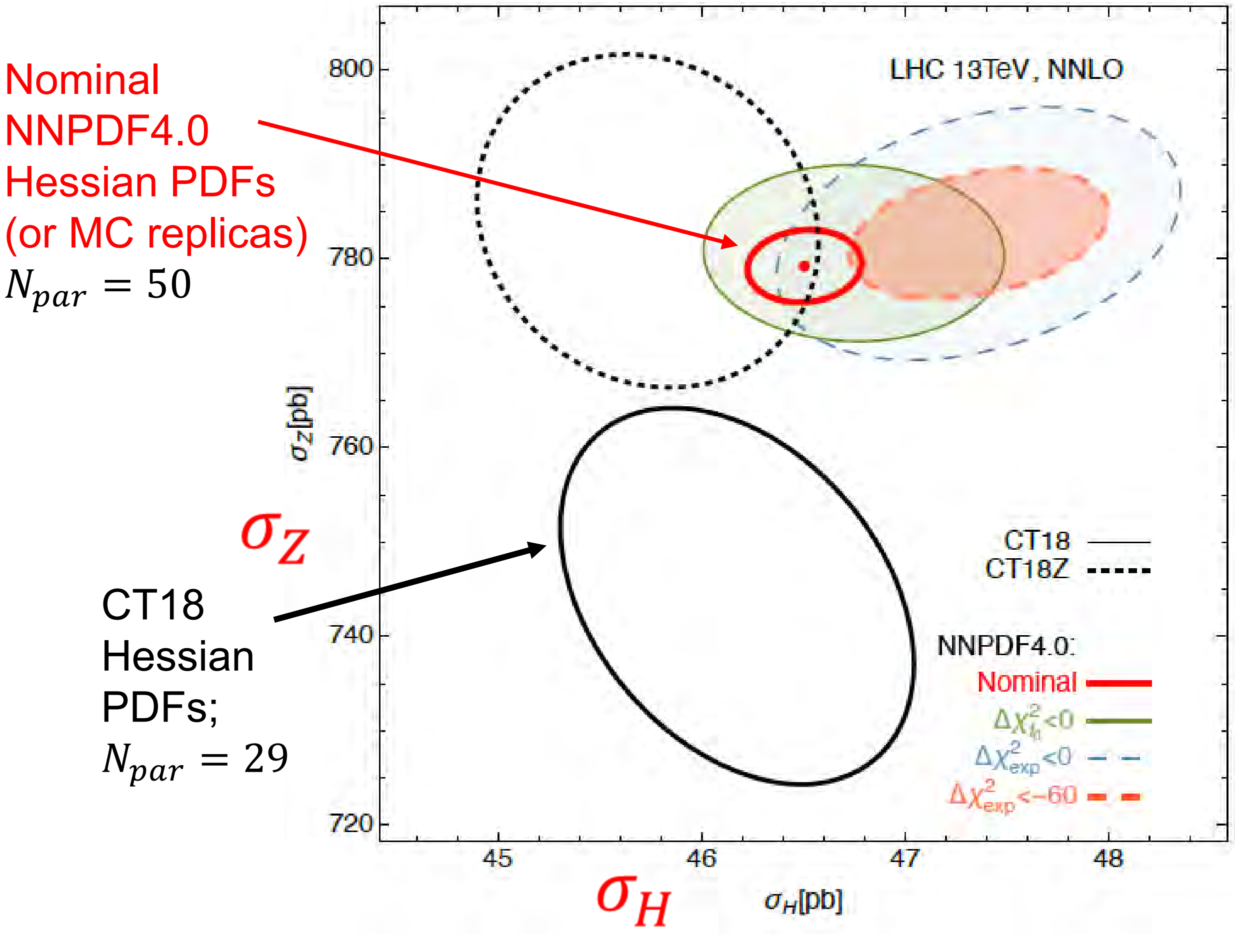
CTEQ

P. Nadolsky, MWDays 2023 @ CERN

- The *prior uncertainty due to methodology* (parametrization/NN architecture, smoothness, data tensions, model for syst. errors, ...) is comparable to the impact of most recent data sets.
- An undetected sampling bias may result in a wrong prediction with a low nominal uncertainty.

Investigate using the hopscotch scans

Hopscotch scans find regions containing well-behaving PDF solutions with $\Delta\chi^2 < 0$, suggesting enlarged NNPDF4.0 PDF uncertainties.

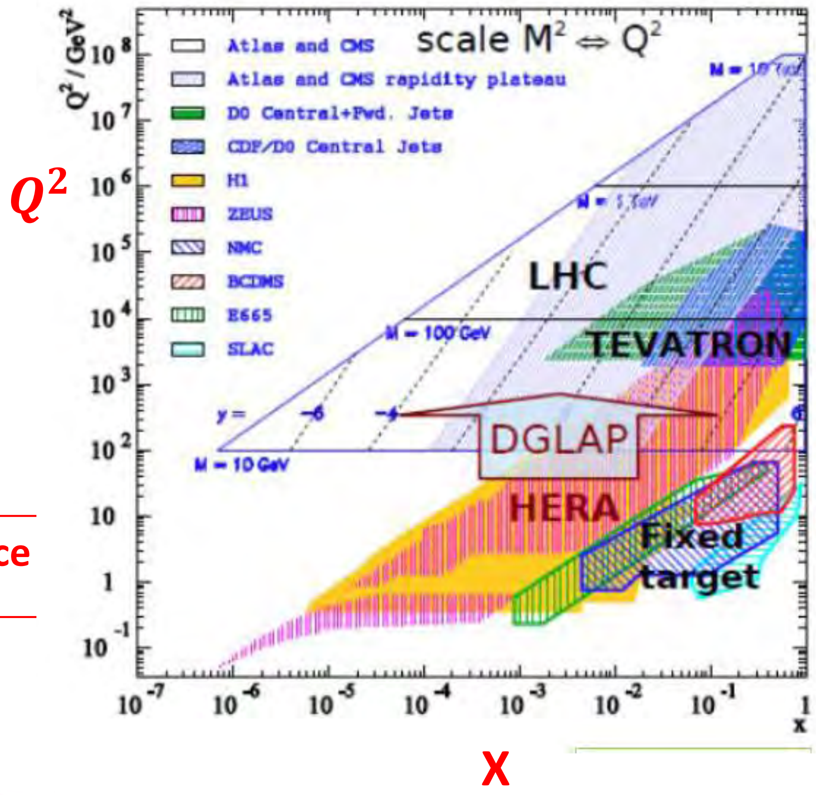


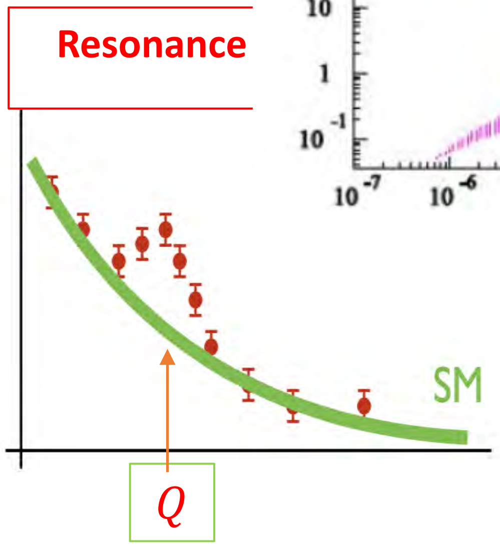
The ellipses are projections of 68% CL ellipsoids in N_{par}-dim PDF shape parameter spaces.

How to use PDFs and their tools

from a user’s point of view

Some basics about PDFs: relevant kinematics in (x, Q^2)





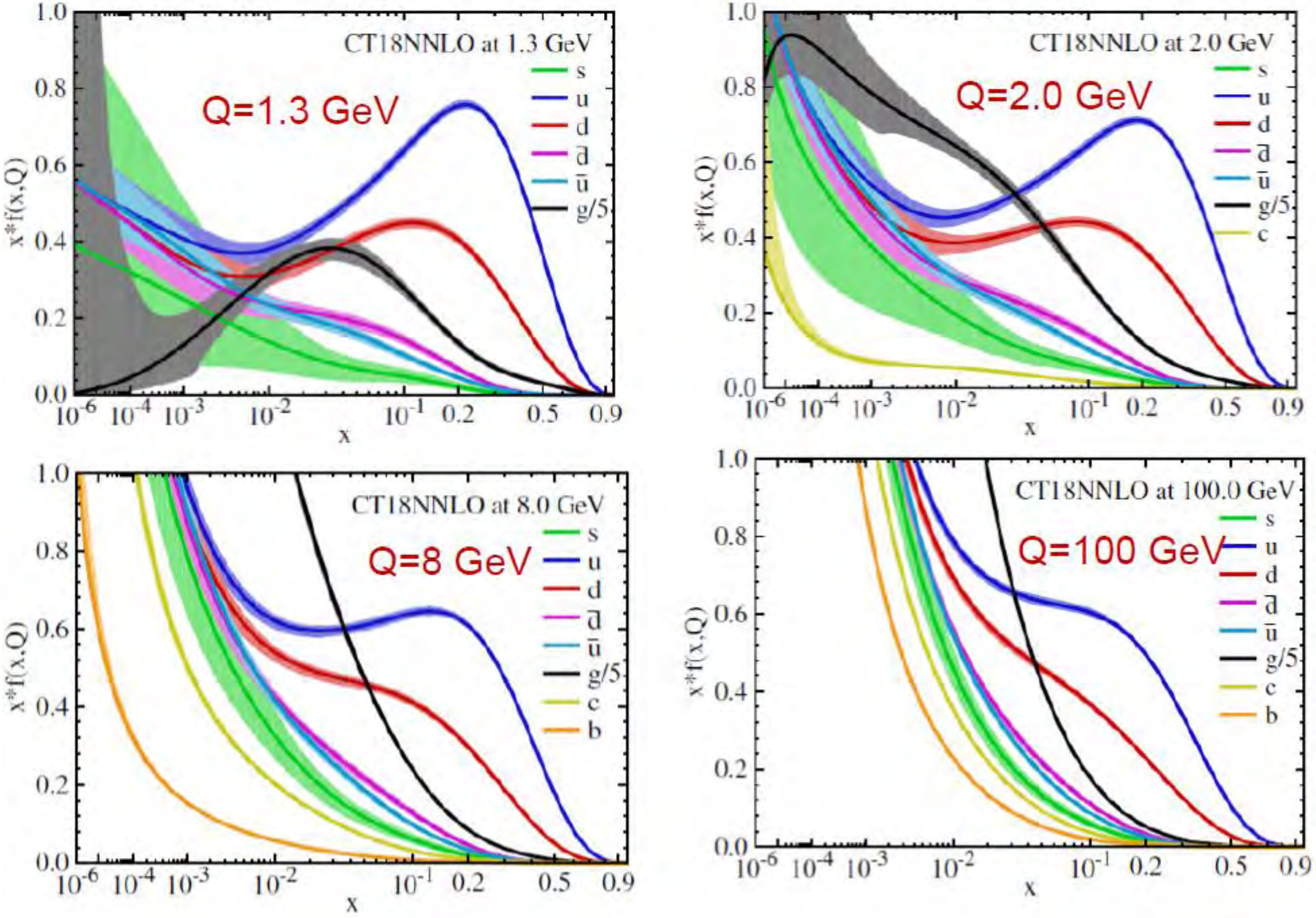
- Parton Distribution Function $f(x, Q)$
- Given a heavy resonance with mass Q produced at hadron collider with c.m. energy $\sqrt{S}$
- What's the typical x value?

$$\langle x \rangle = \frac{Q}{\sqrt{S}}$$ at central rapidity (y=0)

- Generally, $x_1 = \frac{Q}{\sqrt{S}} e^{y}$ *and* $x_2 = \frac{Q}{\sqrt{S}} e^{-y}$

$$x_1 + x_2 = 2\frac{Q}{\sqrt{S}} \cosh(y) \Rightarrow y_{\max} : x_1 + x_2 = 1$$

PDF uncertainties vary as Q via DGLAP evolution

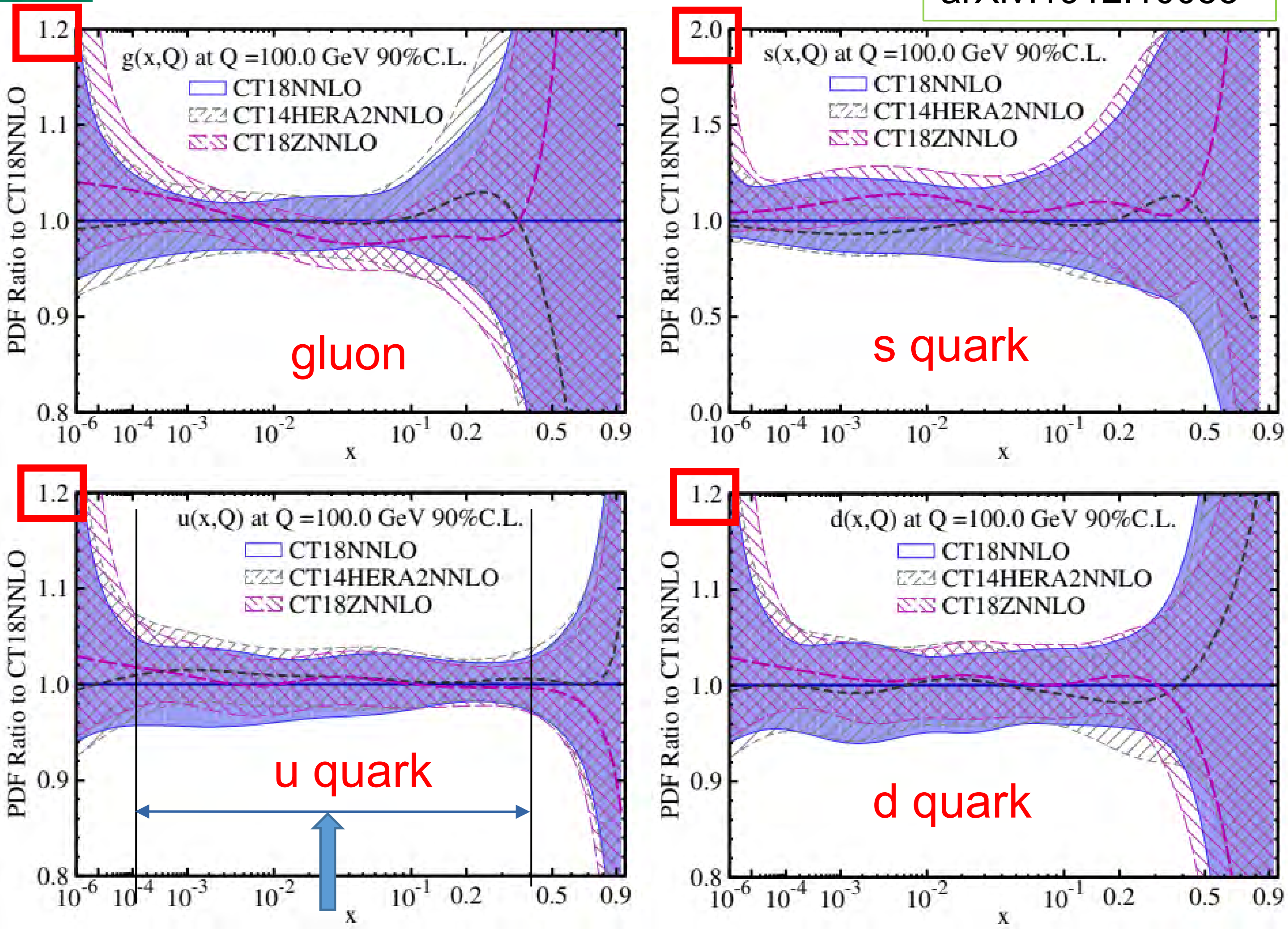


CT18 NNLO PDFs

- Faster DGLAP evolution at low Q values.
- Smaller PDF error bands at higher Q values.
- At high Q, perturbaive contribution becomes more important than the non-perturbative part of PDF.

CT18 PDFs and their uncertainties

arXiv:1912.10053



Better constrained by precision experimental data

- PDFs are better determined at $10^{-4} < x < 0.4$
- Regions of x→1 and x→0 are not experimentally accessible; could use lattice QCD predictions at large x
- Large uncertainty for strangeness PDF, especially in large x region.

Using Hessian method:

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

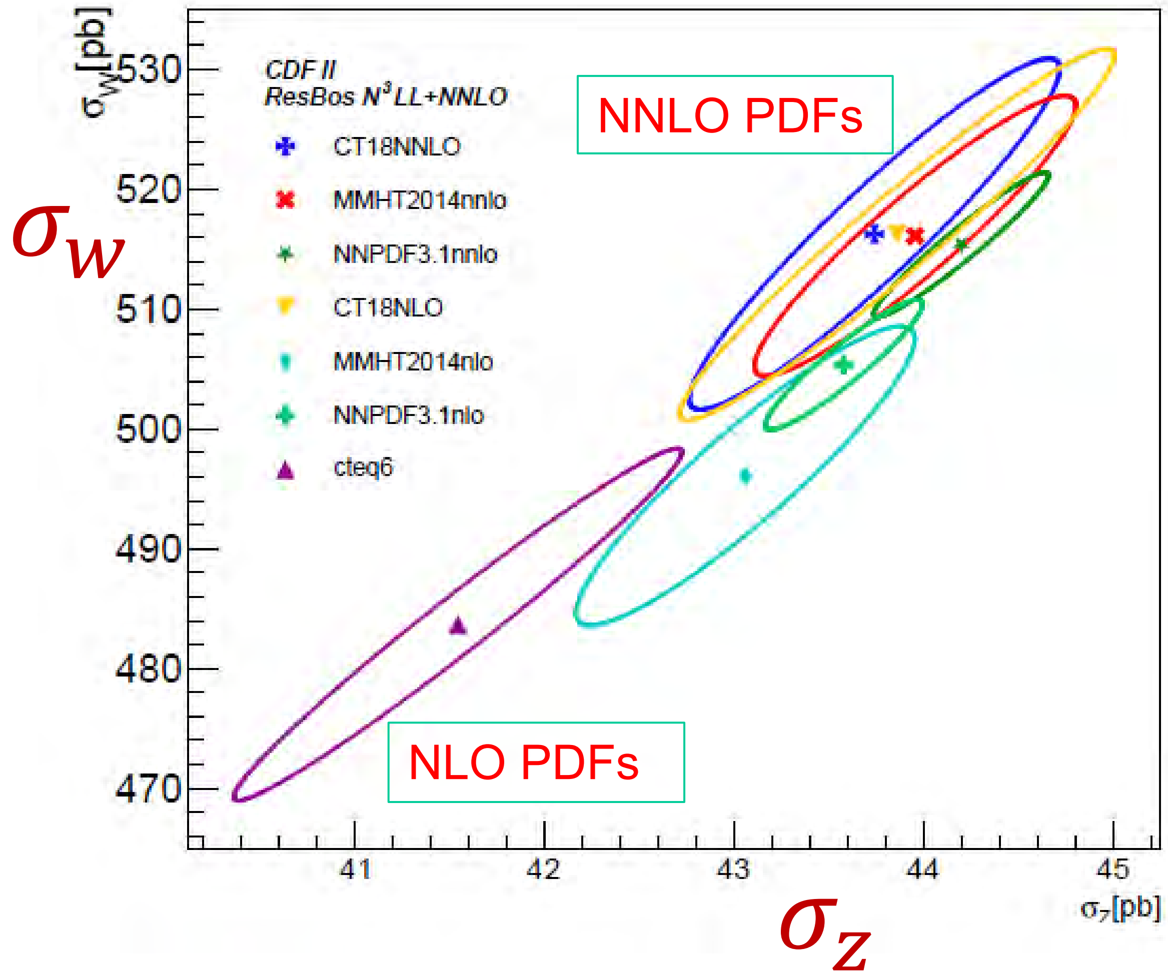
$$\delta^- X = \sqrt{\sum_{i=1}^{N_a}\left[\max\left(X_0 - X_i^{(+)}, X_0 - X_i^{(-)}, 0\right)\right]^2},$$

For CT18, $N_a = 29$

PDF-induced correlation ellipses for CDF II W-mass measurement

ResBos2 predictions

arXiv:2205.02788



- Gluon PDF can contribute to NLO and NNLO predictions.
- Slightly larger PDF-induced errors by NLO PDF sets than NNLO PDF sets.
- Correlation of W and Z (fiducial) cross sections, varies with different PDF sets.
- σ_Z/σ_W is strongly correlated with s PDF.
- $\sigma_Z + \sigma_W$ is strongly correlated with g PDF.

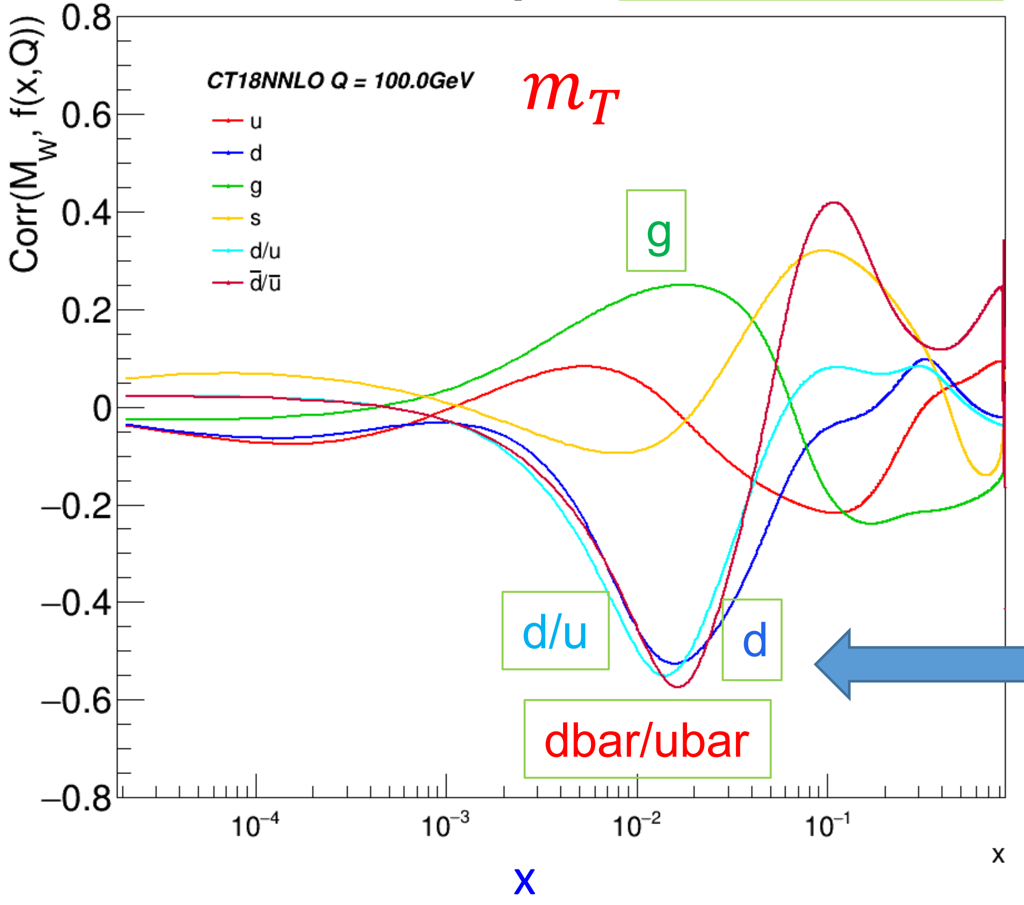
FIG. 10. PDF-induced correlation ellipses, at the 68% confidence level (C.L.), between the fiducial cross sections of W and Z boson production at the Tevatron Run II.

Correlation cosine between the extracted M_W (at CDF II) and CT18 PDFs

Which flavor PDF's error affects most the M_W measurement?

Extracted from $\frac{1}{\sigma}\frac{d\sigma}{dm_T}$

arXiv:2205.02788



Using Hessian method, the correlation between two observables X and Y, which are functions of PDFs, can be described by the correlation cosine

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

with symmetric error $$\Delta X = \frac{1}{2}\left(\sum_{i=1}^{N_a}\left[X(\{z_i^+\}) - X(\{z_i^-\})\right]^2\right)^{1/2}$$

It shows that the CDF-II W boson mass extracted from the normalized m_T distribution is most sensitive to dbar/ubar, d/u and d PDFs at x around 0.01 to 0.1

L_2 Sensitivity

Quantify the degree of tensions among data sets in a fit

https://ct.hepforge.org/PDFs/ct18/figures/L2Sensitivity/

The L_2 Sensitivity for each experiment, E, is defined as

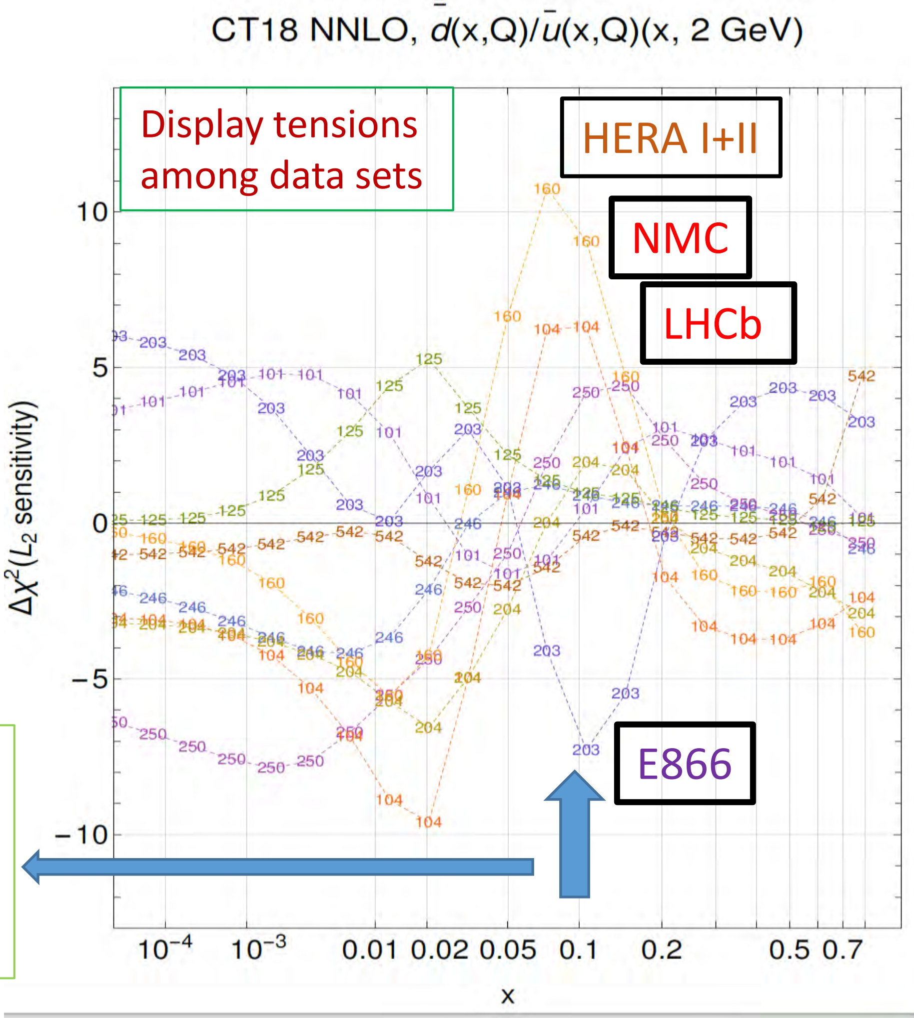
$$S_{f,L2}(E) = \vec{\nabla}\chi_E^2 \cdot \frac{\vec{\nabla} f}{|\vec{\nabla} f|} = \Delta\chi_E^2 \cos\varphi(f,\chi_E^2)$$

where the correlation angle between PDFs and χ_E^2 is

$$\varphi(f,\chi_E^2) = \cos^{-1}\left(\frac{\vec{\nabla} f}{|\vec{\nabla} f|} \cdot \frac{\vec{\nabla}\chi_E^2}{|\vec{\nabla}\chi_E^2|}\right)$$

The L_2 Sensitivity is a fast approximation to the Lagrange Multiplier (LM) scan

When increasing $\bar{d}(x)/\bar{u}(x)$ at x=0.1 and $Q = 2$ GeV, the $\Delta\chi^2$ of E866 pd/pp (203) decreases and that of HERA I+II (160), NMC pd/pp (104), and LHCb 8 W/Z (250) increases. Hence, they have tensions.



When increasing $\bar{d}(x)/\bar{u}(x)$ at 2 GeV by 1 σ error, the change in $\Delta\chi^2$.

$T^2 = 37$ @68% CL

Perform a fast fit to explore the impact of new data on updating the existing PDFs: Hessian profiling

arXiv: 1806.07950
arXiv: 1907.12177

ePump
(error PDF Updating Method Package)

- A tool to examine the impact of a new data set to further constrain the existing PDFs without using a global analysis code.
- A tool to reduce the total number of error PDF sets relevant to specific experimental observables.
- Getting the numerical results in minutes, not hours or days.

ePump-updating

ePump-optimization

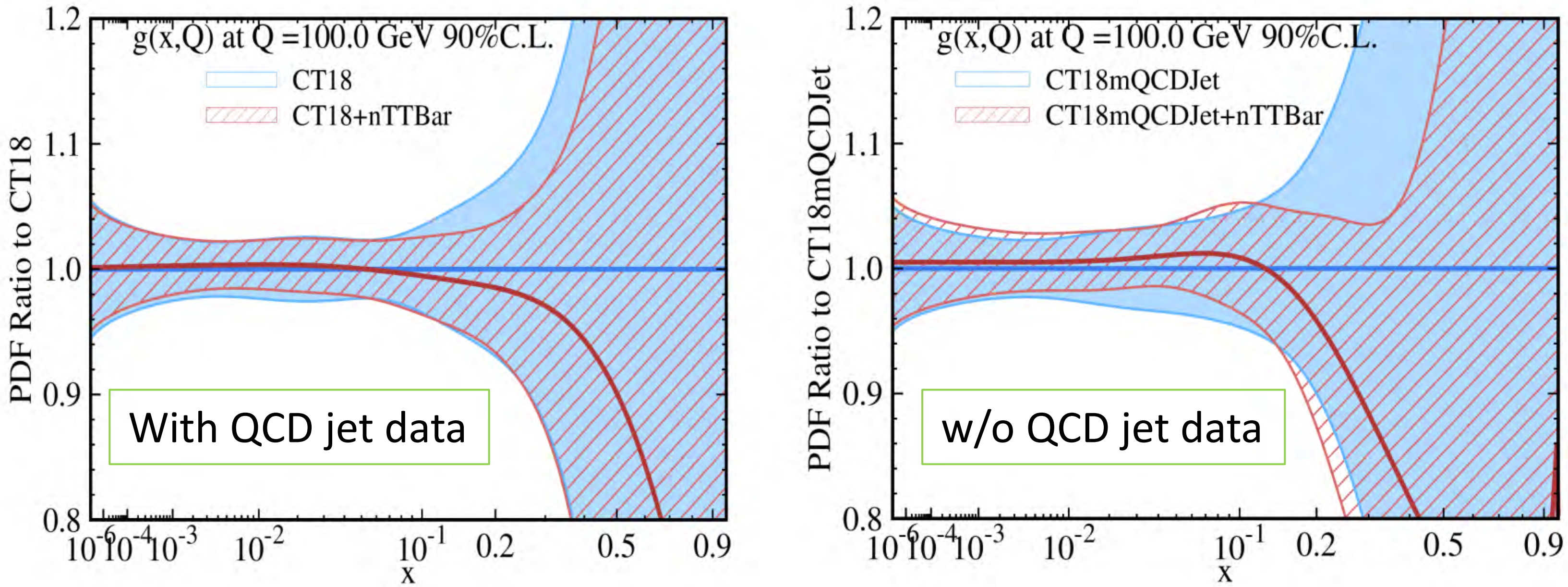
https://epump.hepforge.org/

Impact of $\bar{t}t$ data to large-x gluon PDF
Interplay between top-quark and jet data in CT2X fit

Do $\bar{t}t$ data strongly modify and constrain the PDFs, especially g-PDF at large x?

Use ePump to update the CT18 PDFs by including the post-CT18 $\bar{t}t$ data (nTTBar) at the 13 TeV LHC.

It depends:



The "base" PDF set CT18mQCDJet was obtained without including QCD jet data in the fit.

- The 13 TeV LHC $\bar{t}t$ data prefer a softer gluon-PDF, than CT18, at large x.
- If precision jet data (typically with more data points) are not included in the "base" PDFs, the impact of $\bar{t}t$ data to large-x gluon PDF would become stronger.
- Different scale choice yields different g-PDF

Hessian profiling of CT and MSHT PDFs cannot use $\Delta\chi^2 = 1$

MICHIGAN STATE

CTEQ

ATLAS-CONF-2023-015

The statistical analysis for the determination of $\alpha_s(m_Z)$ is performed with the xFitter framework [60]. The value of $\alpha_s(m_Z)$ is determined by minimising a χ^2 function which includes both the experimental uncertainties and the theoretical uncertainties arising from PDF variations:

$$\chi^2(\beta_{\text{exp}}, \beta_{\text{th}}) = \sum_{i=1}^{N_{\text{data}}} \frac{\left(\sigma_i^{\text{exp}} + \sum_j \Gamma_{ij}^{\text{exp}} \beta_{j,\text{exp}} - \sigma_i^{\text{th}} - \sum_k \Gamma_{ik}^{\text{th}} \beta_{k,\text{th}}\right)^2}{\Delta_i^2} + \sum_j \beta_{j,\text{exp}}^2 + \sum_k \beta_{k,\text{th}}^2 .$$

arXiv: 1907.12177

arXiv:1912.10053

- xFitter profiling uses $\Delta\chi^2 = 1$, by default.
- For CT (or MSHT) PDFs, using $\Delta\chi^2 = 1$ in profiling is equivalent to assigning a weight of about 30 (or 10) to the new data included in the fit. Hence, it will overestimate the impact of new data.
- CT: $T^2 \sim 30$; MSHT: $T^2 \sim 10$

profiling of CT and MSHT PDFs requires to include a tolerance factor $T^2 > 10$ as in the ePump code

When profiling a new experiment with the prior imposed on PDF nuisance parameters $\lambda_{\alpha,th}$:

$$\chi^2(\vec{\lambda}_{\text{exp}}, \vec{\lambda}_{\text{th}}) = \sum_{i=1}^{N_{pt}} \frac{\left[D_i + \sum_\alpha \beta_{i,\alpha}^{\text{exp}} \lambda_{\alpha,\text{exp}} - T_i - \sum_\alpha \beta_{i,\alpha}^{\text{th}} \lambda_{\alpha,\text{th}}\right]^2}{s_i^2} + \sum_\alpha \lambda_{\alpha,\text{exp}}^2 + \sum_\alpha T^2 \lambda_{\alpha,\text{th}}^2 .$$

$$\beta_{i,\alpha}^{\text{th}} = \frac{T_i(f_\alpha^+) - T_i(f_\alpha^-)}{2},$$

new experiment

priors on expt. systematics and PDF params

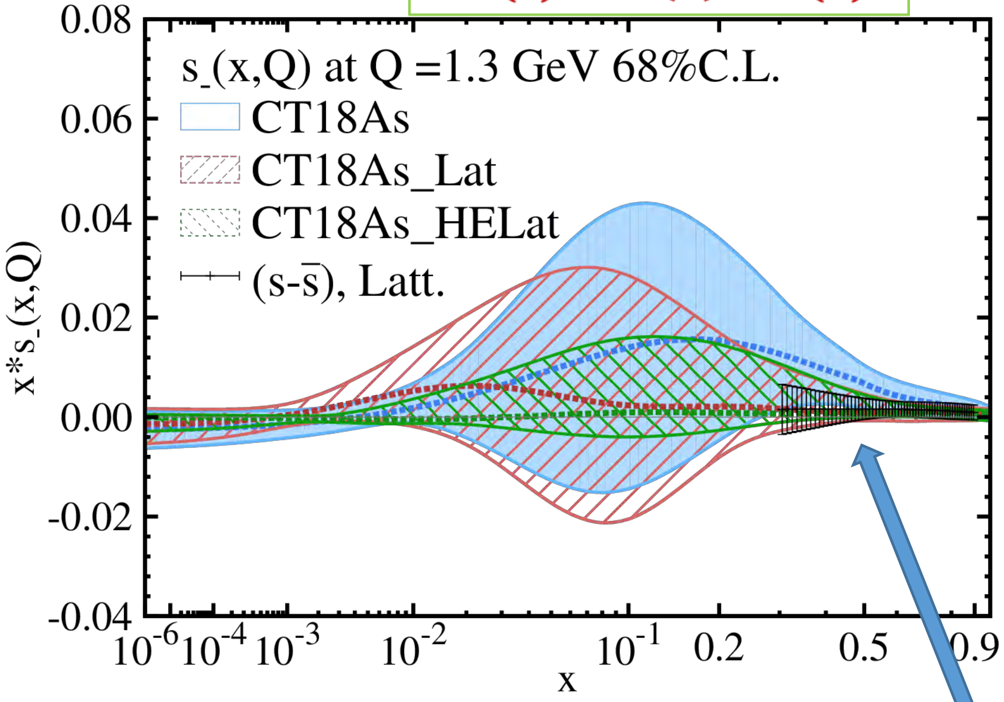
Extensions of CT18 family PDFs: post-CT18

- CT18As: CT18A (a CT18 fit with the inclusion of ATLAS 7 TeV W, Z data), but with non-zero strangeness asymmetry $s_{-}(x, Q_0) = s(x, Q_0) - \bar{s}(x, Q_0)$ at Q_0 = 1.3 GeV.
- CT18As_Lat: CT18As, but including Lattice QCD data on strangeness asymmetry $s_{-}(x, Q_0) = s(x, Q_0) - \bar{s}(x, Q_0)$
- CT18FC: fitted charm PDF $c(x, Q_0) \neq 0$; for c$(x, Q_0) =$ or $\neq \bar{c}(x, Q_0)$
- CT18qed: take photon as a parton of proton; $\gamma(x, Q_0) \neq 0$
- Machine Learning approach: A fast version of Lagrange Multiplier scan (for simultaneous fit to PDFs and SMEFT)
- CT18LO: LO PDF for event generators, e.g., PYTHIA

Lattice QCD data as an input to PDF global analysis

arXiv: 2211.11064

$s_{-}(x) = s(x) - \bar{s}(x)$



- The uncertainties of PDFs can be further reduced by including Lattice QCD predictions in global analysis
- Complementarity of collider experimental data and lattice QCD data

CT18As: CT18A with non-zero strangeness asymmetry $s_{-}(x)$ at Q_0 = 1.3 GeV.
CT18As_Lat: CT18As PDFs with lattice input on $s_{-}(x)$
CT18As_HELat: CT18As_Lat with the lattice errors reduced by half.

CT18A = CT18 + ATLAS W,Z data

- Lattice QCD calculation provides prediction at 0.3 < x < 0.8, while NuTeV and CCFR SIDIS di-muon data constraint strangeness PDFs at 0.015 < x < 0.336.
- Lattice QCD data are consistent with $s(x) = \bar{s}(x)$ at large x.
- CT18 assumes $s(x, Q_0) = \bar{s}(x, Q_0)$; NNLO DGLAP evolution generates $s(x, Q) \neq \bar{s}(x, Q)$ at $Q > Q_0$

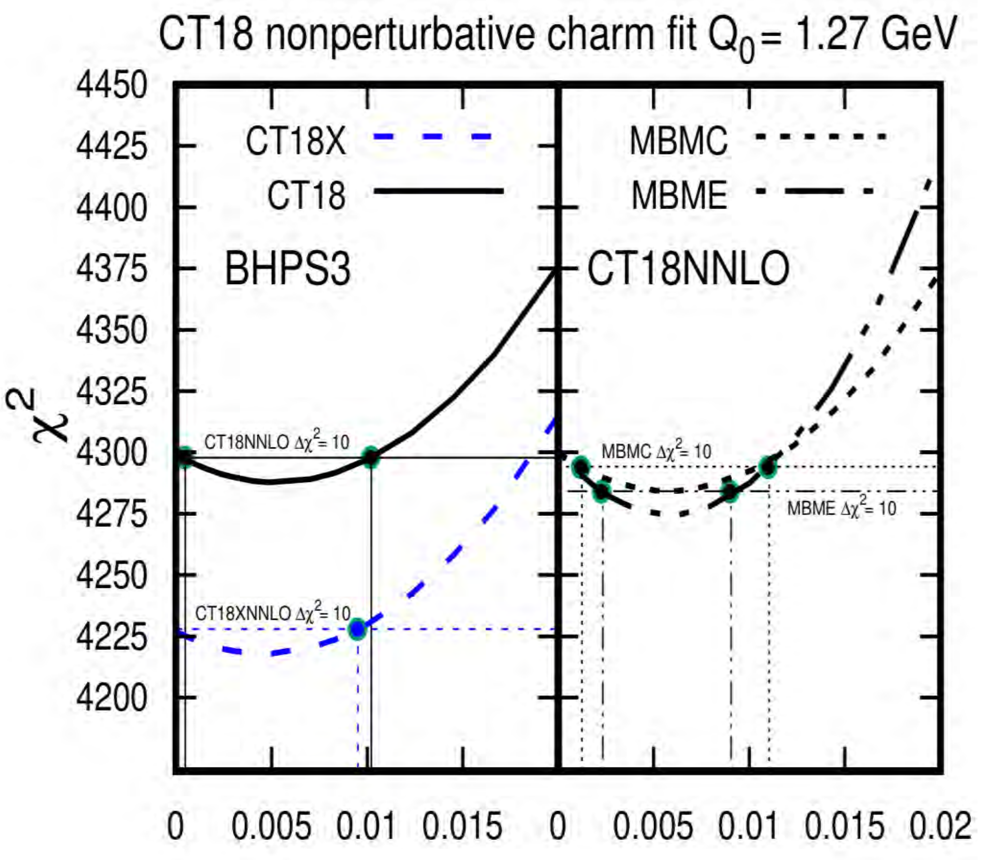
arXiv: 2005.12015

Nonperturbative (intrinsic) charm of proton CT18FC

arXiv:2211.01387

- Proton's intrinsic charm, a non-vanishing charm PDF at Q_0 (around 1 GeV) scale, remains indeterminate.
- Challenging to formulate a rigorous definition of intrinsic charm (IC) and its relation to fitted charm (FC).
- Need more NNLO and better showering calculations.
  - Z+c theory predictions have sizable uncertainties, e.g., flavor-tag jet definition, multi-parton interaction (MPI), showering effect.
- Need more sensitive data

arXiv: 2302.12844



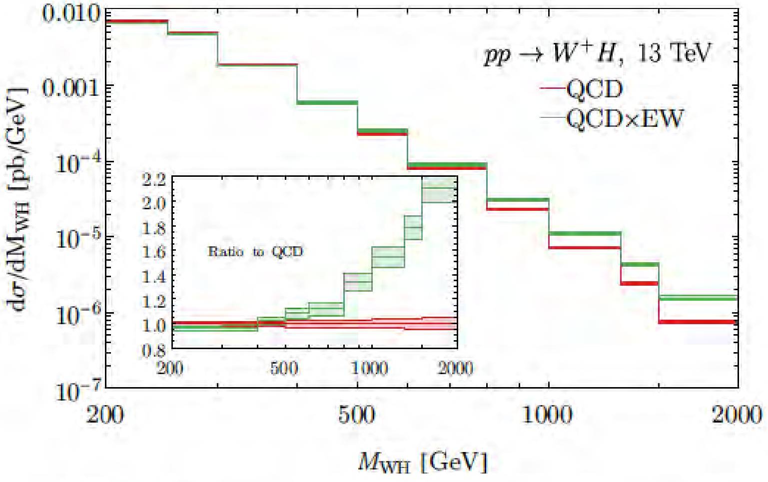
- CT18FC study found no significant evidence for non-zero IC, as NNPDF4.0 IC, Nature 608 (2022) 7923, 483.
- FC in CT18FC study is currently consistent with zero, and with shallower $\Delta\chi^2$ than CT14IC.

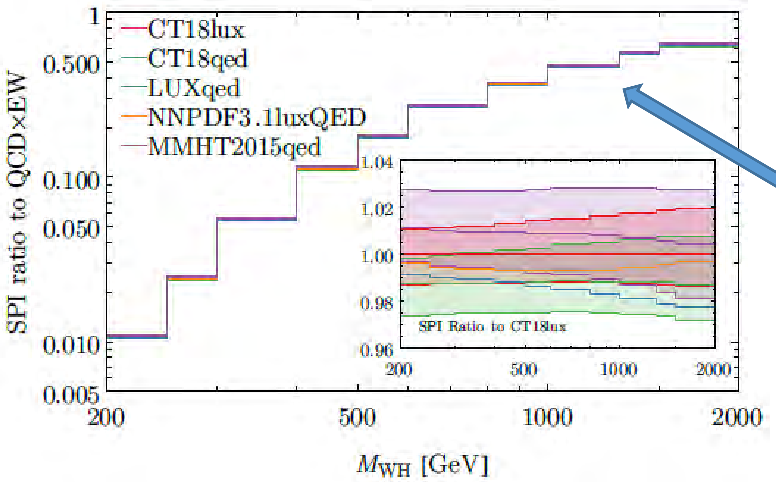
$$\langle x\rangle_{\rm FC} \approx 0.5\% \ (\Delta\chi^2 \gtrsim -25) \text{ vs. } \langle x\rangle_{\rm FC} \approx 0.8-1\% \ (\Delta\chi^2 \gtrsim -40)$$ in CT14 IC

Photon PDF of proton: CT18qed

arXiv:2106.10299

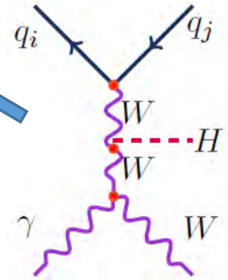
$\sigma(pp \to W^{+}H)$

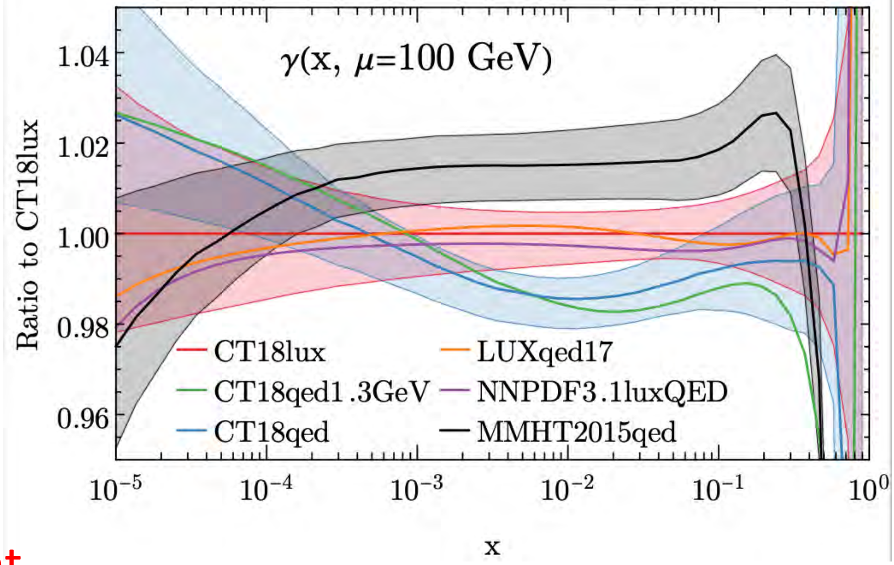




At α_s^2 accuracy, EW corrections and explicit photon PDF $\gamma(x,Q^2)$ are needed.

Single-photon-initiated (SPI) process; important at TeV scale



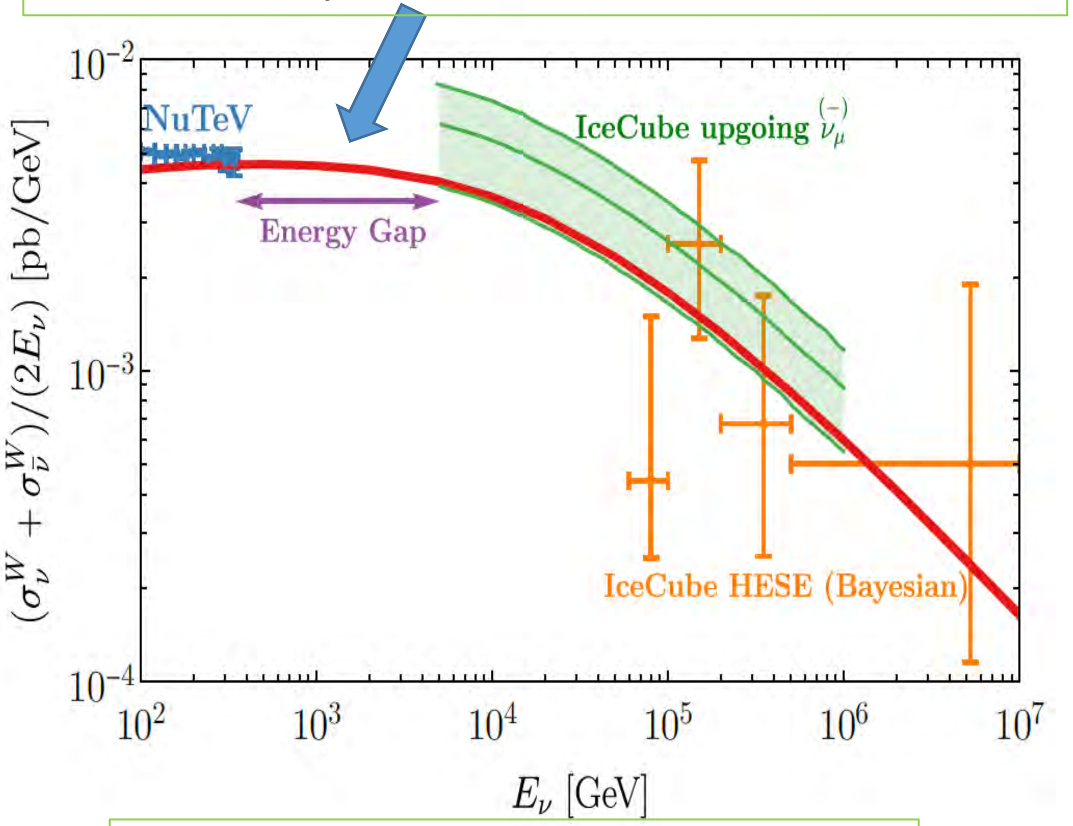


- CT18lux provides the photon PDF at all scales, μ.
- CT18qed initializes photon PDF at μ_0, and evolves to high scales.
- CT18lux gives the photon in between LUXqed(17) and MMHT2015qed, while CT18qed gives smaller photon.

CT18 NNLO high-energy neutrino-nucleon DIS cross sections from 10^2 to 10^7 GeV

arXiv:2303.13607

To be filled by FASER measurement at the LHC



Future data can further constrain PDFs.

We published the first GM NNLO calculation for charged current DIS processes in arXiv:2107.00460, which is needed for

- DUNE (Deep underground neutrino exp)
- EIC (Electron-Ion Collider)
- IceCube Neutrino Observatory
- FASER (Forward search exp at the LHC; the first observation of collider neutrino events)

arXiv:2303.14185

At low E_ν the contributions from quasi-elastic (QE) scattering and resonance (RES) production are important, and not included in this comparison.

(See talk by Keping Xie, Pheno23)

Machine Learning in CTEQ-TEA analysis: SMEFT

C T E Q

It is a simultaneous fit of PDFs and SMEFT couplings.

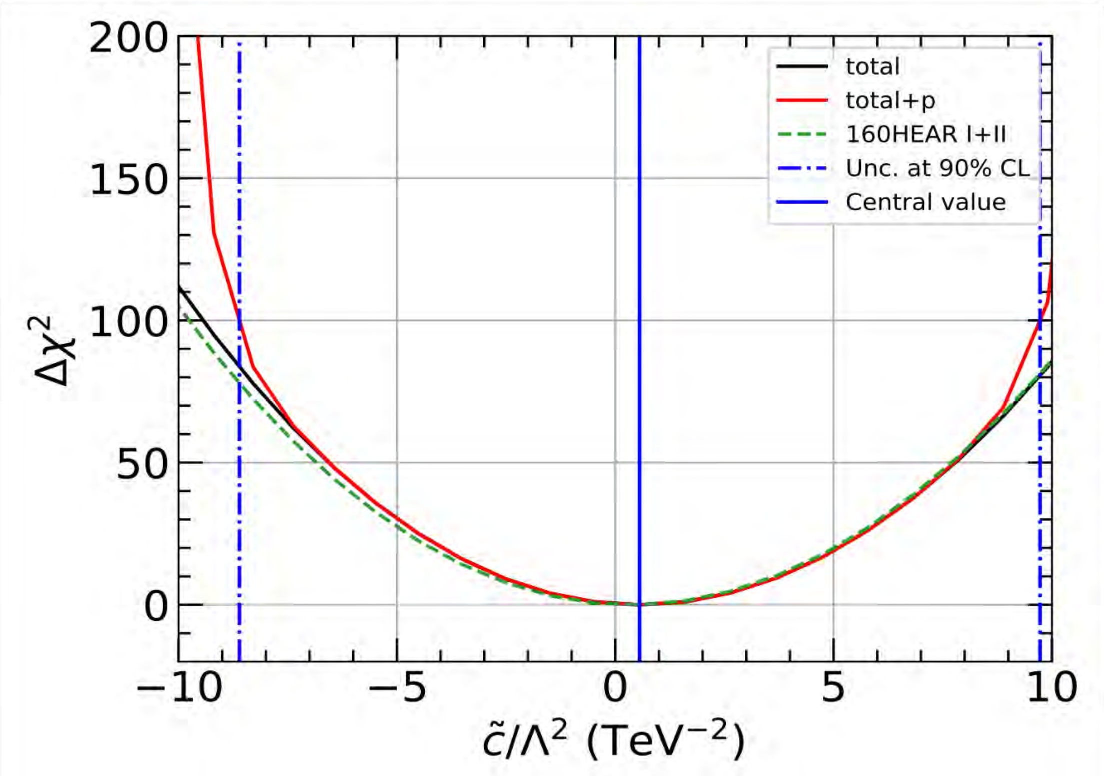
The machine-learning (ML) approach ensures efficient scans over the full PDF parameter space, especially the Lagrange Multiplier scans of χ^2, as demonstrated for a study on the constraint of SMEFT couplings.

arXiv:2201.06586

Lepton-quark contact interactions of SMEFT

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_{i,j} \frac{c_{ij}}{\Lambda^2} (\bar{q}_i \gamma_\mu q_i)(\bar{l}_j \gamma^\mu l_j)$$
$$= \mathcal{L}_{\mathrm{SM}} + \frac{\tilde{c}}{\Lambda^2} \sum_{i,j} e_{q_i} e_{l_j} (\bar{q}_i \gamma_\mu q_i)(\bar{l}_j \gamma^\mu l_j)$$

LM scans on SMEFT couplings



More…

- Up to now, we only discuss PDFs at NNLO accuracy, but some progress has been made toward aN3LO.
- We only focus on higher order QCD corrections, but precision data also require knowing EW and QCD+EW corrections, including resummation (with heavy-flavor mass effect…)
- …

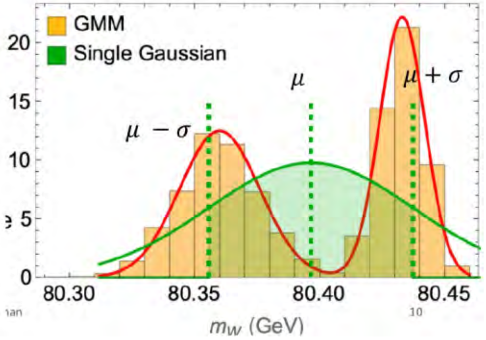
MSHTaN3LO: [arxiv:2207.04739]

NNPDF@aN3LO:
G. Magni, DIS2023

- NNPDF and MSHT aN3LO do not agree on xP_{qg}.
  - NNPDF group makes an ansatz in N-space
  - MSHT group makes an ansatz in x-space.
  - The current information about splitting functions is still not complete.

Tevatron/LHC Combination

- An essentially completed project, waiting to be published
- Addressed QCD and PDF corrections needed to “match” the available measurements; not on the scale of the presently observed discrepancy
- Final presentation of results still under discussion (difficult!)



Kirtimaan Mohan, WG1

The W-boson mass
M. Boonekamp
DIS2023

- How to perform a combined fit to data sets with obvious tension?

(See talk by Kirtimaan Mohan at Pheno23)

Toward a new generation of CT202X PDFs

1. Identify sensitive, mutually consistent new experimental data sets using preliminary fits and fast techniques (L_2 sensitivities and $ePump$)
2. Implement N3LO QCD and NLO EW contributions as they become available. N3LO accuracy is reached only when N3LO terms are **fully** implemented.
   – Meanwhile, "**NNLO+" PDFs**: e.g., include theoretical uncertainty due to QCD scale dependence for key processes as has been done in CT18/CT18X NNLO PDFs
3. Explore quark sea flavor dependence: $s - \bar{s}$ (CT18As), fitted charm (CT18FC),…
4. Include lattice QCD constraints (CT18As_Lat)
5. Next-generation PDF uncertainty quantification: META PDFs, Bézier curves, MC sampling, multi-Gaussian combination, …
6. Lattice QCD: Provides constraints on hadron structures not accessible experimentally

CTEQ-TEA group

C T E Q

- CTEQ – Tung Et Al. (TEA)

in memory of Prof. Wu-Ki Tung, who co-established CTEQ Collaboration in early 90's

- Current members:

China: Sayipjamal Dulat, Ibrahim Sitiwaldi, Alim Albet (Xinjiang U.), Jun Gao (Shanghai Jiaotong U.), Mengshi Yan (Peking U.) , Tie-Jiun Hou (U. of South China), Yao Fu (USTC)

Mexico: Aurore Courtoy (Unam, Mexico)

USA: Marco Guzzi (Kennesaw State U.), Tim Hobbs (Argonne Lab), Pavel Nadolsky, Xiaoxian Jing (Southern Methodist U.), Keping Xie (Pittsburgh U.) Joey Huston, Huey-Wen Lin, Dan Stump, Carl Schmidt, CPY (Michigan State U.)

Some useful websites:

- CT18 PDFs — https://ct.hepforge.org/PDFs/ct18/
- L2 Sensitivity — https://ct.hepforge.org/PDFs/ct18/figures/L2Sensitivity/
- ePump — https://epump.hepforge.org/
- ResBos2 — https://gitlab.com/resbos2

Lesson learned from W mass measurement

C T E Q

Experimentalists ➕ Theorists

2017 Featured Story #1: Million-dollar gift establishes endowed professorship in honor of the late Dr. Wu-Ki Tung

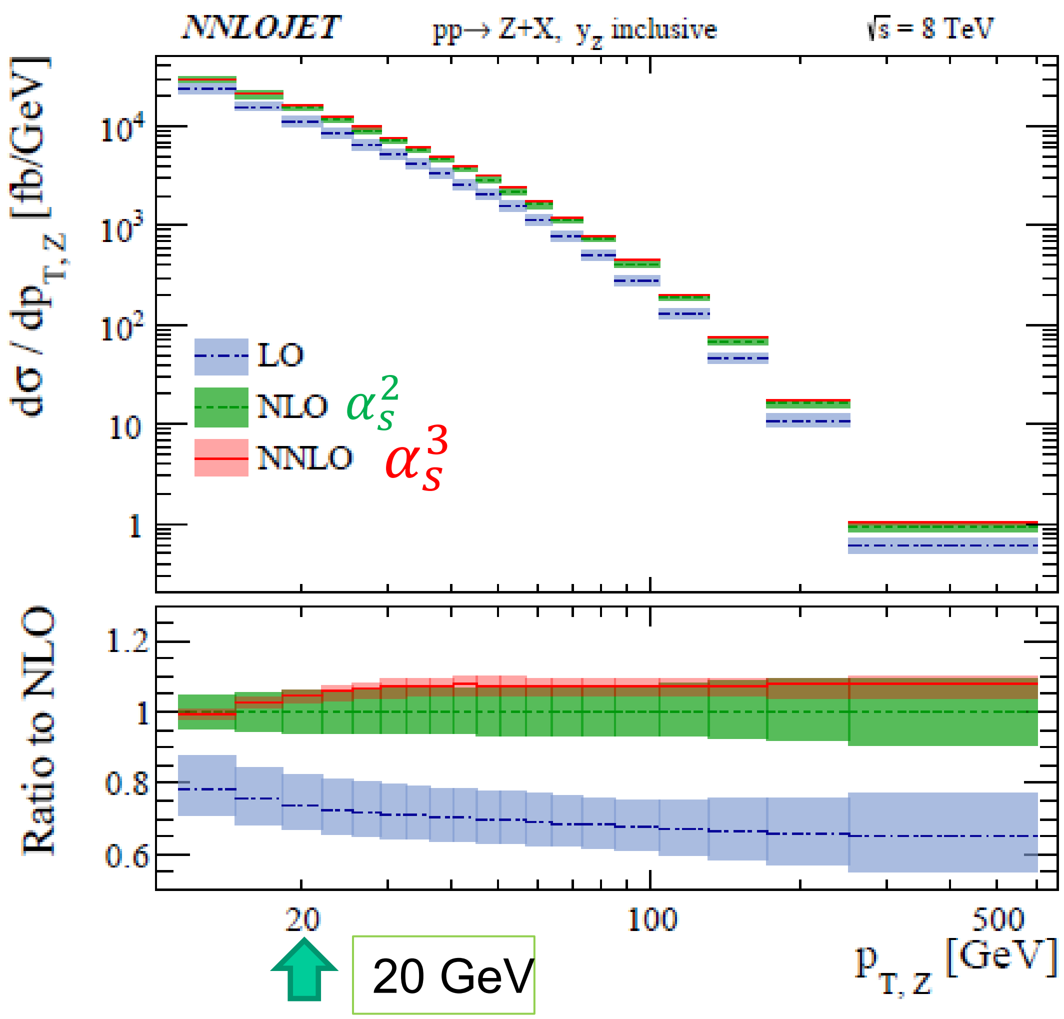
Michigan State University
(1992-2009)

http://www.pa.msu.edu/node/5921

- Co-founder of CTEQ (**The Coordinated Theoretical-Experimental Project on QCD)** in 1989 – present
- Nowadays, many other collaborations are doing precisely that.

Backup slides

Higher order contributions are important



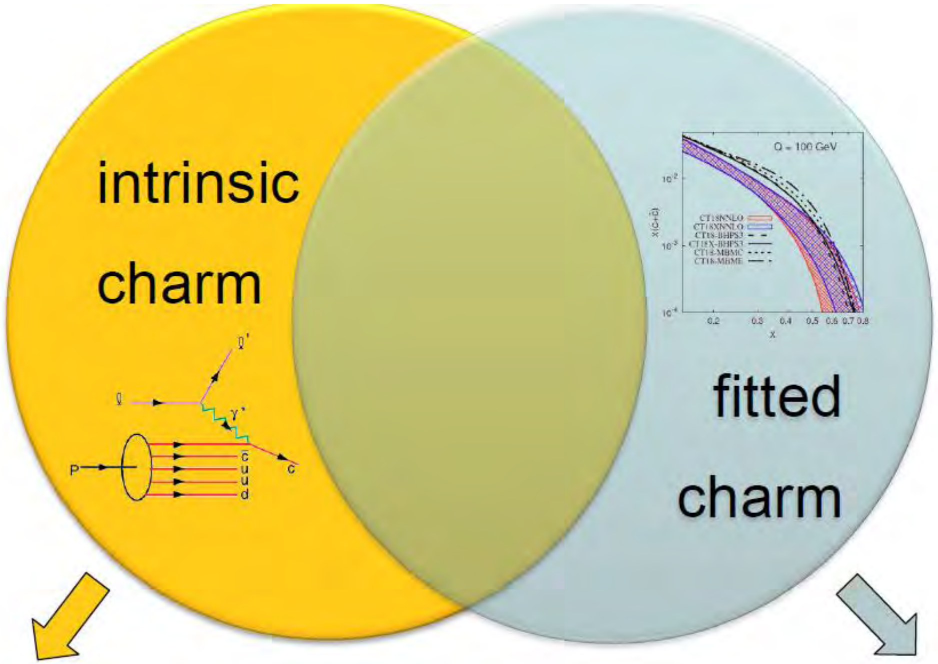
arXiv:1708.00008

- The α_s^3 prediction has much smaller scale variation as compared to α_s^2 calculation.
- For $p_{T(Z)} > 20$ GeV, the K-factor of α_s^3/α_s^2 is roughly a constant, about 1.1

Challenging to formulate a rigorous definition of Intrinsic Charm and its relation to Fitted Charm

C T E Q

arXiv:2211.01387



- The concept of nonperturbative methods
- Can refer to a component of the hadronic Fock state or the type of the hard process
- Predicts a typical enhancement of the charm PDF at $x \gtrsim 0.2$

Connection?

- A charm PDF parametrization at scale $Q_0 \approx 1$ GeV found by global fits [CT, NNPDF, …]
- Arises in perturbative QCD expansions over α_s and operator products
- May absorb process-dependent or unrelated radiative contributions

ePump-optimization

arXiv: 1806.07950; 1907.12177

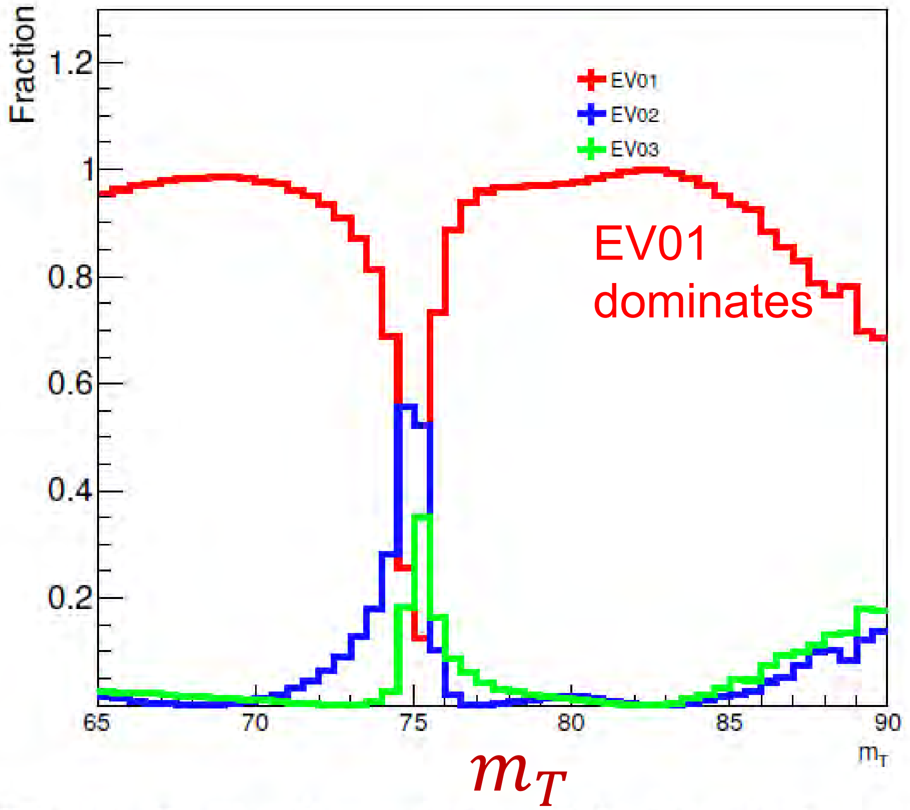


FIG. 12. Fractional contribution of the three leading optimized eigenvector PDFs (EV01, EV02 and EV03) to the variance of the m_T distribution, normalized to each bin, obtained from the ePump-optimization analysis.

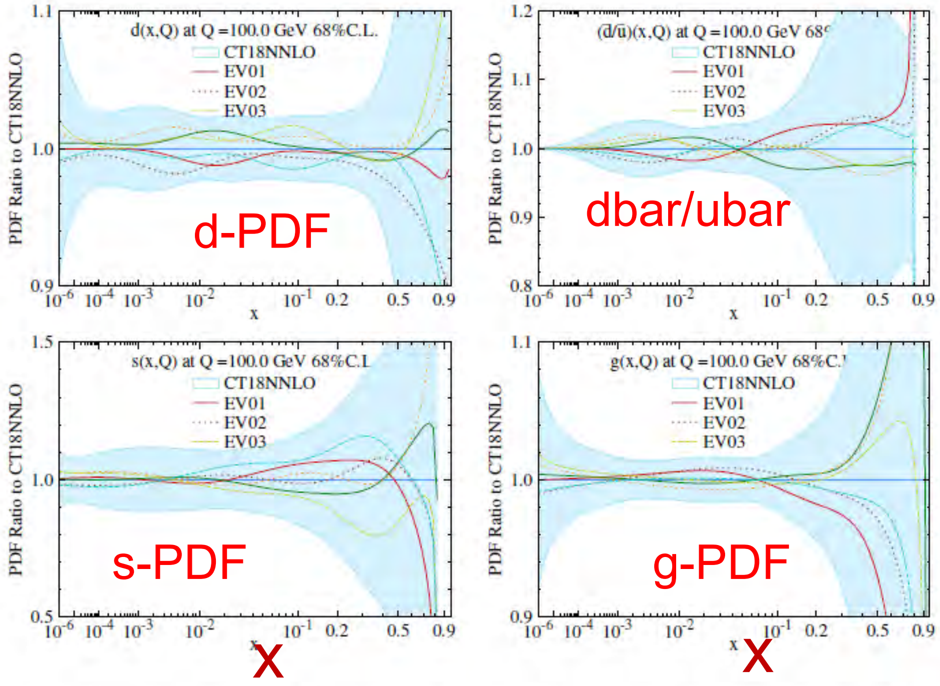


FIG. 13. Ratios of the top three pairs of eigenvector PDFs and the original CT18 NNLO error PDFs, at $Q = 100$ GeV, to the CT18 NNLO central value of d, $\bar{d}/\bar{u}$, s and g PDFs. These eigenvector PDFs were obtained after applying the ePump-optimization to the original CT18 NNLO PDFs with respect to the m_T distribution.

The three eigenvalues are 44.5, 3.0 and 2.4, respectively, with 50 bins in the m_T distribution.