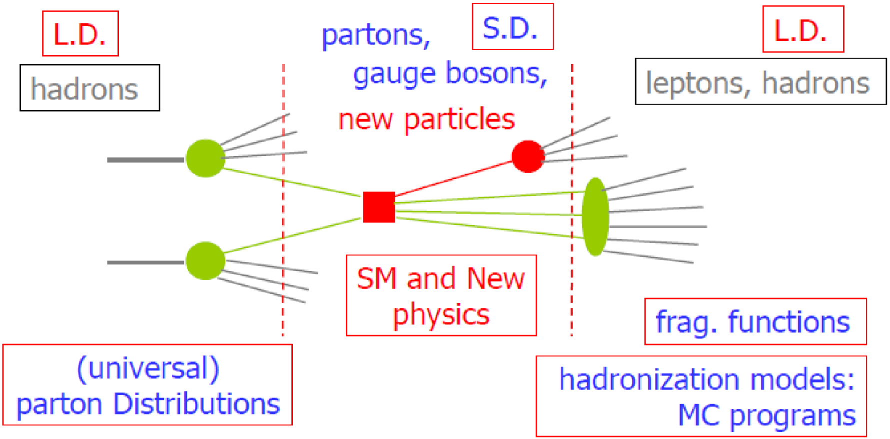
Recent development on PDF at the LHC

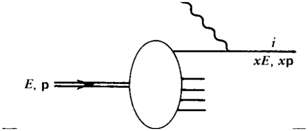
Tie-Jiun Hou

Northeastern University (Shenyang)

January 9, 2020

NCTS Dark Physics Workshop, Hsinchu



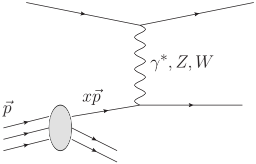


Parton distribution function (PDF)

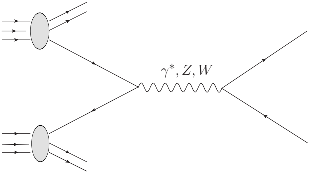
$f_{j/A}(x,Q)$ describe the possibility to find a parton j, i.e. quark and gluon, in a nucleon A.

PDF $f(x,Q)$ is universal

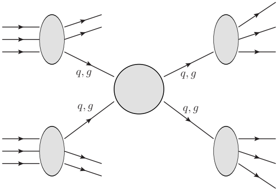
DIS



Drell-Yan



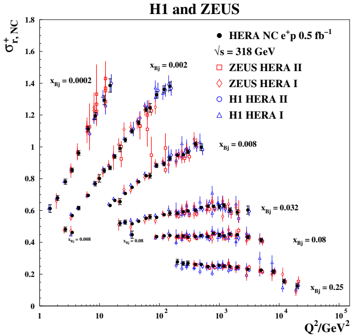
jet

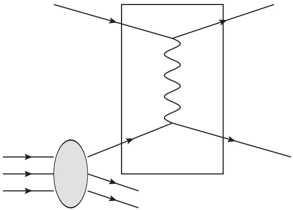


PDF(Parton distribution function) tell us the probability to find out a parton in a proton with particular momentum fraction x and energy Q.

PDF is determined by comparing data and hard cross section

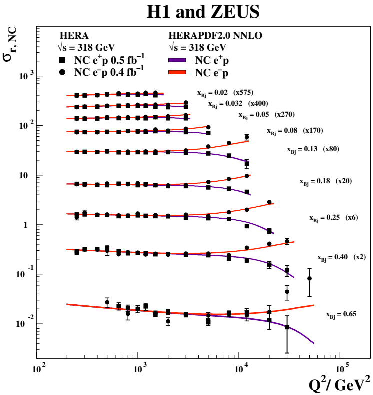
$$\sigma = f(x, Q^2, \{a\}) \otimes \hat{\sigma}$$

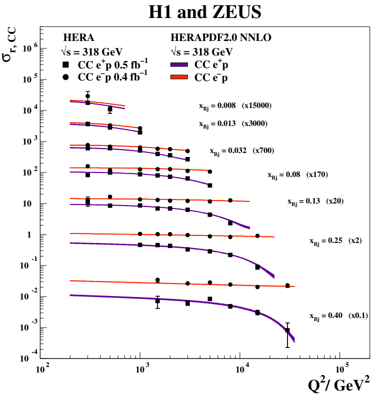




PDF	Hard part
LO	LO
NLO	NLO
NNLO	NNLO

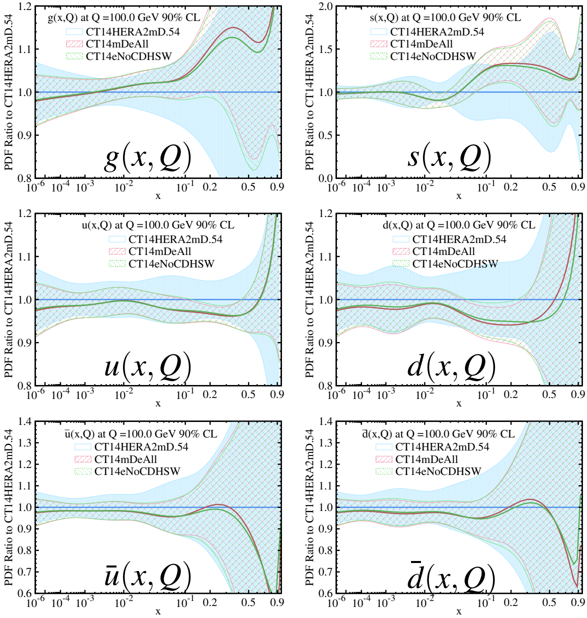
HERA I+II





There are 3287 data point from HERA I+II data, after the criteria $Q > 2$ GeV and $W^2 > 12.5$ GeV^2, 1120 points are included in global analysis. It constraint PDF from $10^{-4} < x < 0.6$. The HERA I+II is one of the most firm foundation of global analysis of PDF now a day.

EPJC75 (2015) no.12, 580



CDHSW F_2^p and F_3^p

Z. Phys. C49 (1991) 187-224

CT14HERA2mD.54:
CT14HERA2 without DIS data but HERA.

CT14mDeAll:
CT14HERA2mD.54 + all DIS.

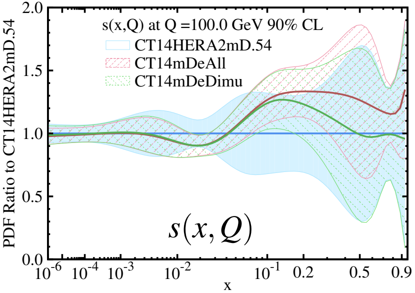
CT14eNoCDHSW:
Including CDHSW and HERA data for DIS only.

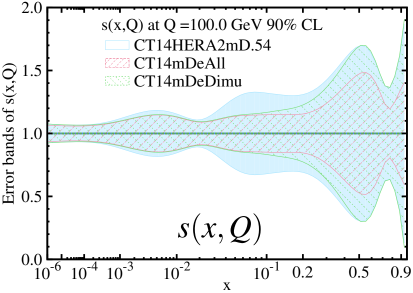
Arguement about it.
EPJC12, 243 (2000)

arXiv: 1907.12177

Di-muon

NuTeV $\nu\mu\mu$ and $\bar{\nu}\mu\mu$ SIDIS: D. A. Mason, Ph.D. thesis, Oregon U. (2006)
CCFR $\nu\mu\mu$ and $\bar{\nu}\mu\mu$ SIDIS: PRD64, 112006 (2001)

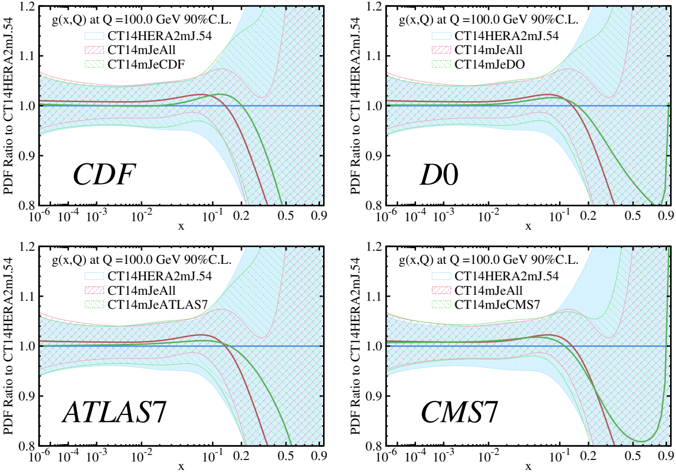




CT14mDeDimu: CT14HERA2mD.54 + di-muon.

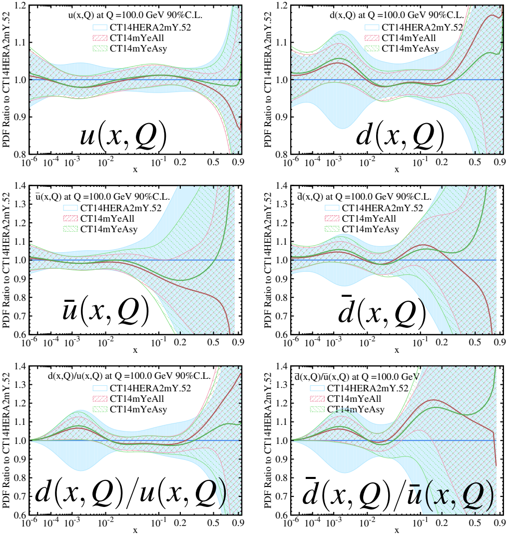
The NuTeV and CCFR di-muon data solely constraint strange PDFs for $0.01 < x < 0.4$.

Inclusive Jet



Jet data have dominant constraint on gluon, especially for $0.02 < x < 0.5$. Among the jet data from Tevetron and LHC Run 1, the CMS 7 TeV jet data dominate the constraint.

W charge asymmetry A_{ch}



D0 A^{μ}_{ch}: PRD77, 011106 (2008)

ATL 7 *WZ*: PRD85, 072004 (2012)

CMS 7 A^{e}_{ch}: PRL109, 111806 (2012)

CMS 7 A^{μ}_{ch}: PRD90, 032004 (2014)

CT14HERA2mY.54:

CT14HERA2 without DY

CT14mYeAll:

CT14HERA2mY.54 + all DY.

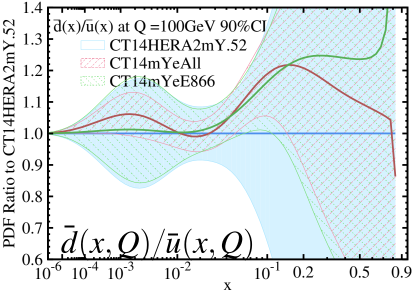
CT14mYeAsy:

CT14HERA2mY.54 + A_{ch}.

Asymmetry data play a dominant role among DY data.

E866 $\sigma_{pd}/(2\sigma_{pp})$

PRD64, 052002 (2001)



CT14mYeE866: CT14HERA2mY.54 + E866.

The E866 data solely dominate the constraint for $\bar{d}/\bar{u}, 0.01 < x < 0.2$.

Weak mixing angle $\sin^2\theta_W$

LEP/SLD: PR427, 257 (2006)

$\sin^2\theta_W = 0.23153 \pm 0.00016$

Tevetron: PRD97 (2018) no11, 112007

$\sin^2\theta_W = 0.23148 \pm 0.00027\text{(Stat.)} \pm 0.00005\text{(syst.)} \pm 0.00018\text{(PDF.)}$

CMS: EPJC78 (2018) no.9, 701

$$\sin^2\theta_W = 0.23101 \pm 0.00036\text{(Stat.)} \pm 0.00018\text{(syst.)} \pm 0.00016\text{(theo.)} \pm 0.00031\text{(PDF.)}$$

For pp collider LHC, the PDF uncertainty for a $q\bar{q} \to Z$ is non-negoligible.

$\Longrightarrow$We reduce the PDF uncertainty for $\sin^2\theta_W$ by updating PDF through A_{FB} and $A_{\pm}$ in DY process $f\bar{f} \to Z/\gamma^* \to ll$!

Reduce PDF uncertainty for $\sin^2\theta_W$ by A_{FB}

Samples:

- Pseudo-data: CT14NNLO+ResBos, $\sim$ LHC Run 2, (130 fb^{-1} 500 M in full phase space)
- Theory: central and error sets CT14NNLO+ResBos

$\sin^2\theta_w$:

- Pseudo-data: $\sin^2\theta_w = 0.2345$
- Theory: $\sin^2\theta_w = 0.2315$

ATLAS acceptance:

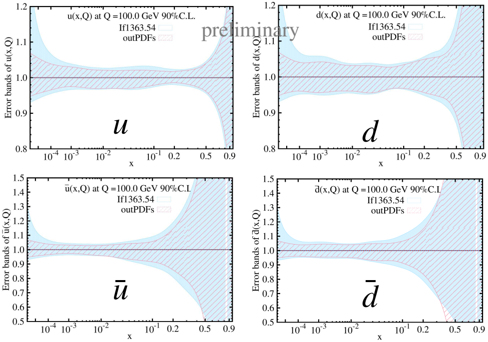
- lepton $p_T > 25$ GeV
- CC: both lepton $\eta < 2.5$ (double luminosity for $ee + \mu\mu$)
- CF: one lepton $\eta < 2.5$ the other $2.5 < \eta < 5.0$ (only for CF)
- Z pole: $\text{M} = [80, 100]$ GeV
- sideband: $\text{M} = [60, 80] + [100, 130]$ GeV

To be published by Yao Fu, Liang Han, Minghui Liu, Tie-Tiun Hou, Chen Wang, Siqi Yang, Hang Yin, C.-P. Yuan.

Update PDF by both Z-pole and sideband region

average A_{FB} at Z pole (preliminary)

Update using CC+CF		central	Stat. unc.	PDF unc.
Theory prediction in CC+CF				
Before update	$\sin^2\theta_w = 0.2315$	0.01846	0.00007	0.00117
after update	$\sin^2\theta_w = 0.2315$	0.01770	0.00007	0.00053(54.7%)
Theory prediction in CC				
Before update	$\sin^2\theta_w = 0.2315$	0.00873	0.00008	0.00083
after update	$\sin^2\theta_w = 0.2315$	0.00824	0.00008	0.00042(49.4%)
Theory prediction in CF				
Before update	$\sin^2\theta_w = 0.2315$	0.04197	0.00017	0.00198
after update	$\sin^2\theta_w = 0.2315$	0.04064	0.00017	0.00092(53.5%)

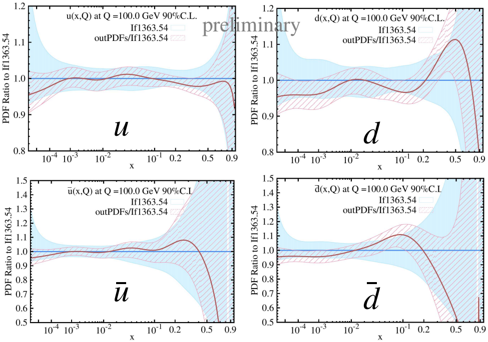


54.7% reduction on PDF uncertainty for A_{FB} for the combined CC+CF case.

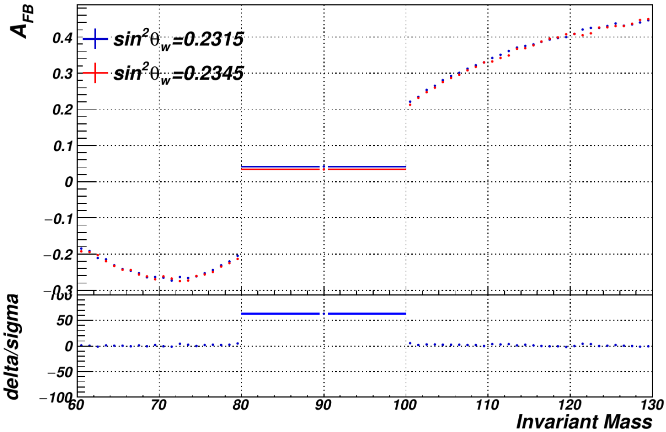
Update PDF by both Z-pole and sideband region

average A_{FB} at Z pole (preliminary)

Update using CC+CF		central	Stat. unc.	PDF unc.
Theory prediction in CC+CF				
Before update	$\sin^2\theta_w = 0.2315$	0.01846	0.00007	0.00117
after update	$\sin^2\theta_w = 0.2315$	0.01770	0.00007	0.00053(54.7%)
pseudo-data	$\sin^2\theta_w = 0.2324$	0.01709	0.00007	-
Theory prediction in CC				
Before update	$\sin^2\theta_w = 0.2315$	0.00873	0.00008	0.00083
after update	$\sin^2\theta_w = 0.2315$	0.00824	0.00008	0.00042(49.4%)
pseudo-data	$\sin^2\theta_w = 0.2324$	0.00793	0.00008	-
Theory prediction in CF				
Before update	$\sin^2\theta_w = 0.2315$	0.04197	0.00017	0.00198
after update	$\sin^2\theta_w = 0.2315$	0.04064	0.00017	0.00092(53.5%)
pseudo-data	$\sin^2\theta_w = 0.2324$	0.03920	0.00017	-



Central PDF receive large impact from the pseudo A_{FB} data, and bias the predicted $\sin^2\theta_w$.



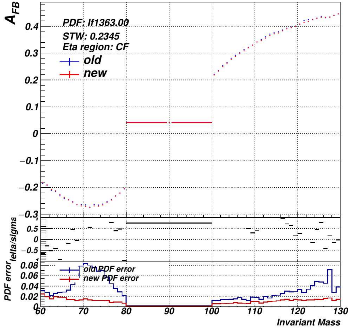
Z-Pole mass region is sensitive to the $\sin^2\theta_w$, the inclusion of Z-Pole mass region for PDF update bias the $\sin^2\theta_w$ prediction.
$\Longrightarrow$ $\sin^2\theta_w$ should be determined by measurement but not PDF!

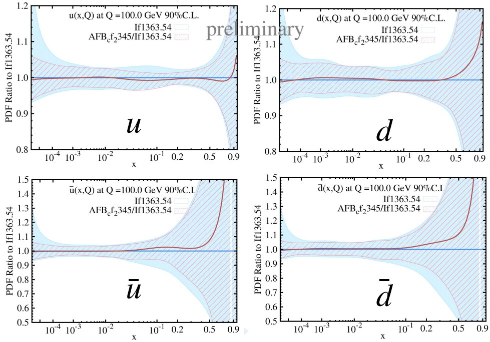
Update PDF by sideband CF region

average A_{FB} at Z pole (preliminary)

Update using CF		central	Stat. unc.	PDF unc.
Theory prediction in CC+CF				
Before update	$\sin^2\theta_w = 0.2315$	0.01846	0.00007	0.00117
after update	$\sin^2\theta_w = 0.2315$	0.01844	0.00007	0.00084(28.2%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.01490	0.00007	-
Theory prediction in CC				
Before update	$\sin^2\theta_w = 0.2315$	0.00873	0.00008	0.00083
after update	$\sin^2\theta_w = 0.2315$	0.00874	0.00008	0.00062(25.3%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.00695	0.00008	-
Theory prediction in CF				
Before update	$\sin^2\theta_w = 0.2315$	0.04197	0.00017	0.00198
after update	$\sin^2\theta_w = 0.2315$	0.04175	0.00017	0.00143(27.8%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.03411	0.00017	-

Less reduction on PDF unc, but $\sin^2\theta_w$ bias is smaller than statistical error for $\sin^2\theta_w = 0.2345$ pseudo-data.



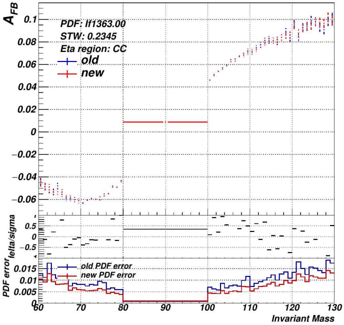


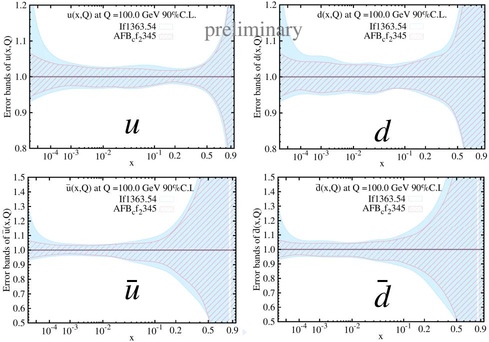
Update PDF by sideband CC region

average A_{FB} at Z pole (preliminary)

Update using CC		central	Stat. unc.	PDF unc.
Theory prediction in CC+CF				
Before update	$\sin^2\theta_w = 0.2315$	0.01846	0.00007	0.00117
after update	$\sin^2\theta_w = 0.2315$	0.01839	0.00007	0.00087(25.6%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.01490	0.00007	-
Theory prediction in CC				
Before update	$\sin^2\theta_w = 0.2315$	0.00873	0.00008	0.00083
after update	$\sin^2\theta_w = 0.2315$	0.00869	0.00008	0.00067(19.3%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.00695	0.00008	-
Theory prediction in CF				
Before update	$\sin^2\theta_w = 0.2315$	0.04197	0.00017	0.00198
after update	$\sin^2\theta_w = 0.2315$	0.04182	0.00017	0.00144(27.3%)
pseudo-data	$\sin^2\theta_w = 0.2345$	0.03411	0.00017	-

Less reduction on PDF unc, but $\sin^2\theta_w$ bias is smaller than statistical error for $\sin^2\theta_w = 0.2345$ pseudo-data.





Reduce PDF uncertainty for $\sin^2\theta_W$ by $A_\pm$

Samples:

- Pseudo-data: CT14NNLO+ResBos, $\sim$ LHC Run 2, (130 fb^{-1} 5000 M in full phase space)
- Theory: central and error sets CT14NNLO+ResBos

ATLAS acceptance:

- lepton $p_T > 25$ GeV (including neutrinos)
- charge lepton $|\eta| < 2.5$
- Both electron and muon channel.

Bin size:

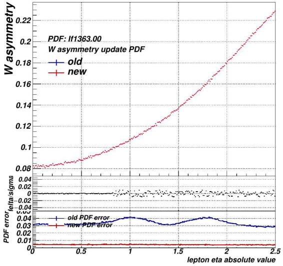
- 0.1 bin size used.

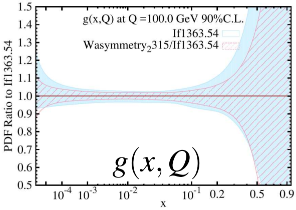
Reduce PDF uncertainty for $\sin^2\theta_W$ by $A_\pm$

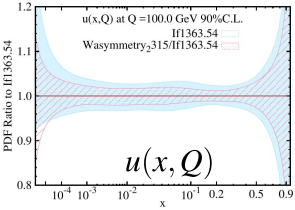
average A_{FB} at Z pole (preliminary)

Update using CC	central	Stat. unc.	PDF unc.
Theory prediction in CC+CF			
Before update	0.01846	0.00007	0.00117
after update	0.01846	0.00007	0.00081(30.8%)
Theory prediction in CC			
Before update	0.00873	0.00008	0.00083
after update	0.00873	0.00008	0.00061(26.5%)
Theory prediction in CF			
Before update	0.04196	0.00017	0.00198
after update	0.04197	0.00017	0.00139(29.8%)

- Less sensitive to energy spectrum
- Larger cross section
- Negligible correlation to weak mixing angle
- PDF unc. reduced by 30% for $A_\pm$.







Reduce PDF uncertainty for $\sin^2\theta_W$ by A_{FB} and $A_{\pm}$ combined

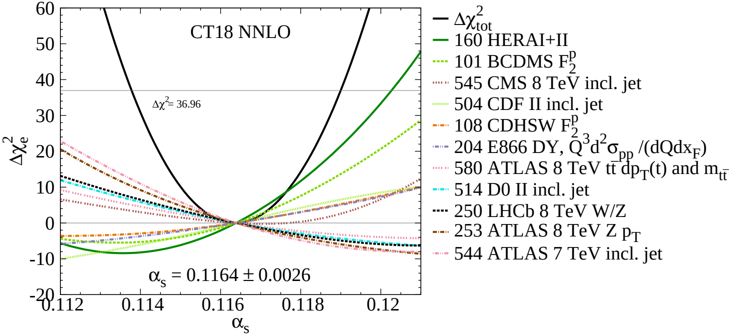
Update PDF by A_{FB} and A_{+} combined (preliminary)	
Theory prediction in CC	PDF unc.
Before update	0.00156
After update	0.00077(50.6%)
Theory prediction in CF	
Before update	0.00079
After update	0.00035(55.7%)

Overall PDF unc. reduced by 56% by both A_{FB} and $A_{\pm}$.

Update PDF by A_{FB} (preliminary)			
bin size unc.	Update by CC event	Update by CF event	Update by CC+CF event
Before update	0.00079	0.00079	0.00079
After update(1 GeV)	0.00057(27.8%)	0.00056(29.1%)	0.00048(50.6%)
After update(2 GeV)	0.00060(24.0%)	0.00060(24.0%)	-
After update(5 GeV)	0.00065(17.7%)	0.00062(21.5%)	-
Update PDF by $A_{\pm}$ (preliminary)			
After update(0.1)	0.00055(30.4%)		
After update(0.2)	0.00063(20.2%)		

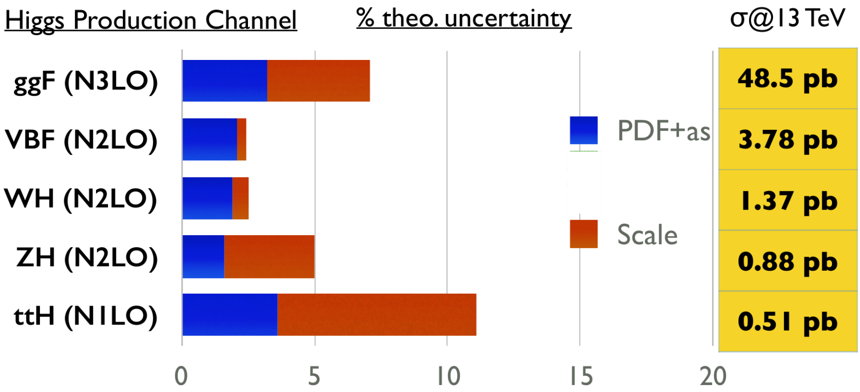
- Larger bin size: lower sensitivity
- Smaller bin size: larger systematic uncertainties

Study of physical parameter through global analysis: $\alpha_s(M_z)$



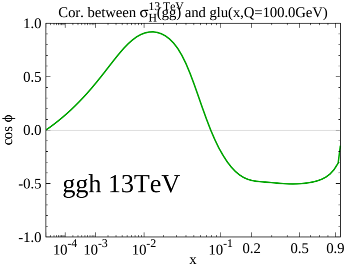
- The global fitting of $\alpha_s(M_z)$ value need to change all the $\alpha_s(M_z)$ value at the same time.
- The fixed target F_2 data and HERA DIS data prefer smaller α_s value.
- The ATLAS 8TeV Z p_T, ATLAS 7 TeV incl. jet data, bring the central value of $\alpha_s(M_z)$ from $0.115^{+0.006}_{-0.004}$ (CT14) to 0.1164 ± 0.0026 (CT18).

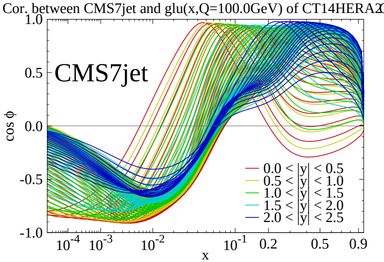
PDF uncertainty in Higgs production

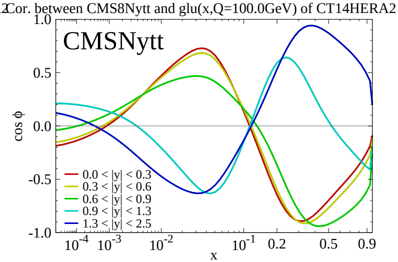


Yellow Report 4 (2016)

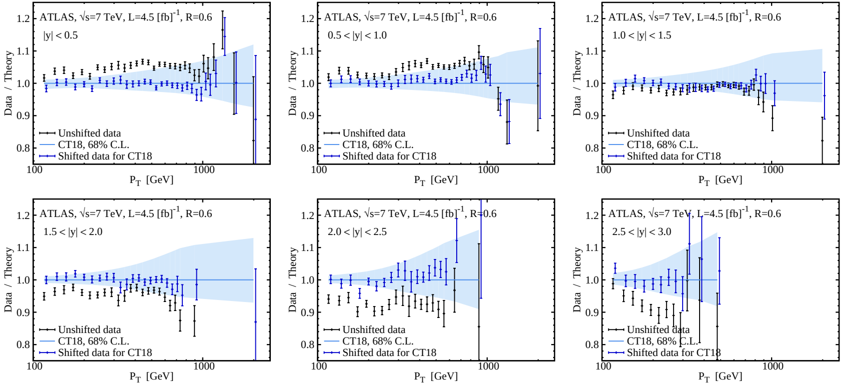
Correlation between Higgs production and gluon PDF through gluon fusion







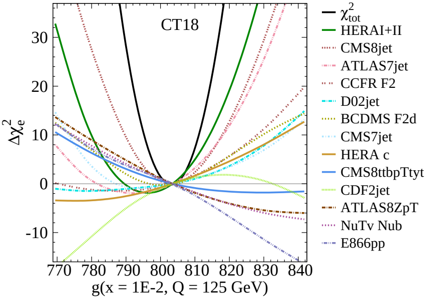
De-correlation for incl. jet

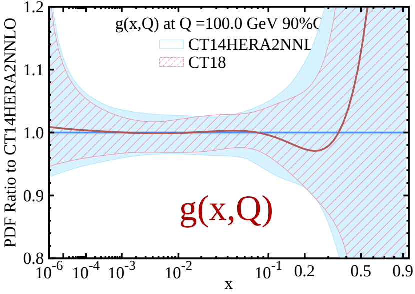


JHEP 1502 (2015) 153, Erratum: JHEP 1509 (2015) 141

- The corr. error "jes16" and "jes62" of ATLAS 7 TeV incl. jet data are de-correlated according to Table 6 of 1706.03192. Its χ^2/Npts reduces from 2.34 to 1.68 for CT14HERA2NNLO.
- Precise systematic error analysis help on the reduction of global analysis of PDFs.

gluon PDF of CT18

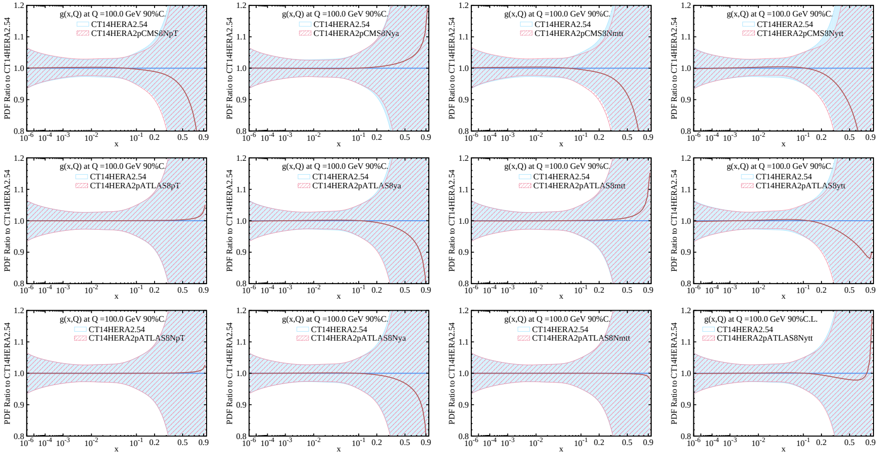




Lagrange Multiplier Scans for gluon at x around 0.01:

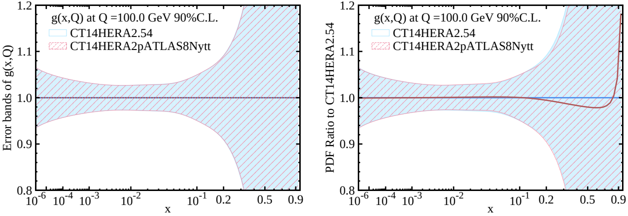
- ATLAS 8 TeV Z p_T data prefer a slightly larger gluon PDF.
- ATLAS 7 TeV and CMS 8TeV incl. jet data prefer a slightly smaller gluon PDF.
- HERA I+II data prefer a slightly smaller gluon PDF.
- Including all the contribution in global analysis, the reduction of PDF uncertainty for $gg \rightarrow h$ production reduced by 5% comparing to CT14.

Impact of top-quark pair production on CT14HERA2

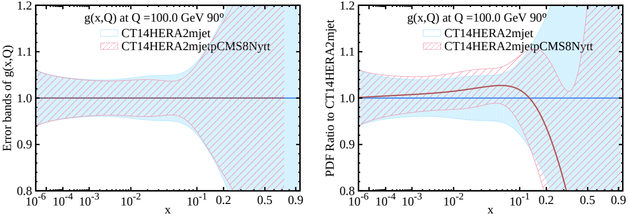


- No significant impact on the uncertainty of PDFs.
- Minor impact on gluon in large x region.

PoS DIS2019 (2019) 017



Impact from ATLAS 8TeV norm. y_{tt} $t\bar{t}$ data on CT14HERA2

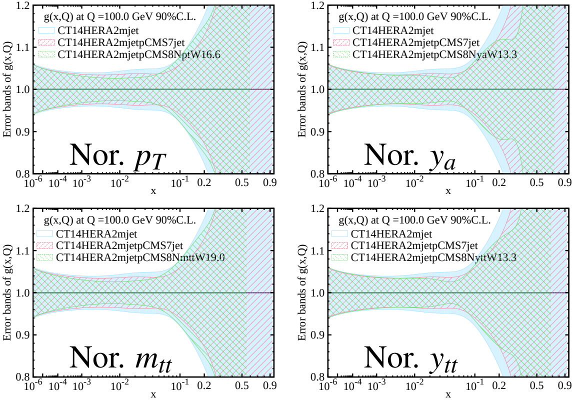


Impact from CMS 8TeV norm. y_{tt} $t\bar{t}$ data on CT14HERA2mjet

- CT14HERA2mjet: CT14HERA2 without all the jet data included.
- Without the jet data included in global analysis, $t\bar{t}$ data have rather obvious impact on both central predictions and error bands of PDFs.

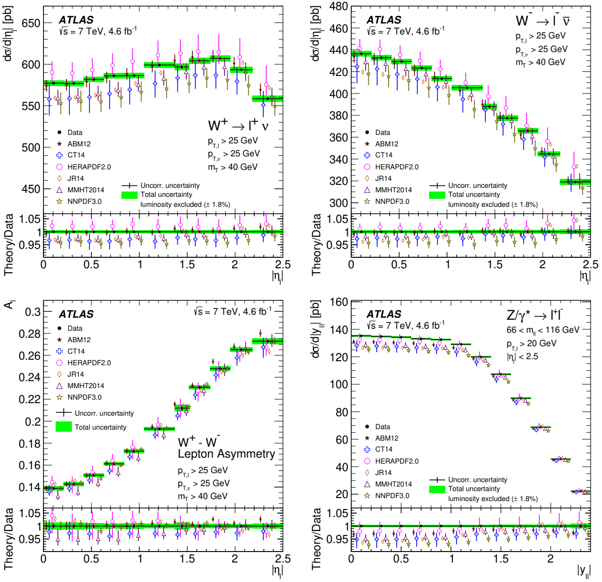
Distribution	Detector	Npts	χ^2/N
inclusive jet	CDF	72	1.50
inclusive jet	D0	110	1.03
inclusive jet	ATLAS	90	0.57
inclusive jet	CMS	133	0.93
$\frac{1}{\sigma}\frac{d\sigma}{dp_T^t}$	ATLAS, CMS	8,8	0.39, 3.88
$\frac{1}{\sigma}\frac{d\sigma}{dy_t}$	ATLAS, CMS	5,10	2.70, 2.53
$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	ATLAS, CMS	7,7	0.25, 8.67
$\frac{1}{\sigma}\frac{d\sigma}{dy_{t\bar{t}}}$	ATLAS, CMS	5,10	2.46, 3.67

The Npts of $t\bar{t}$ data is smaller by a factor of 10 than jet data.



With the assumption of a higher weight for $t\bar{t}$ data. By weighting the $t\bar{t}$ data by the ratio of Npts of CMS 7 jet(133) to Npts of $t\bar{t}$ data, Impact from weighted $t\bar{t}$ data on gluon PDF is as strong as jet data.

ATLAS 7 TeV WZ production



The statistical error of measurment is about 1%, while the uncertainty from PDF is about 5% ∼ 7%.

EPJC77 (2017) no.6, 367

- In the era of LHC, PDF uncertainty used to be one of the dominant uncertainty in a measurement.
- To reduce the uncertainty from PDF need precise SM measurements.
- We need to know the SM observable which can really help on reducing the uncertainty of PDF about the region sensitive to the target measurement, and then analysis the SM observable ahead.
- The communication between experimentalist and global fitter and the real global analysis take times

PDFsense

PDFsense predicts that the CMS data will have the largest impact

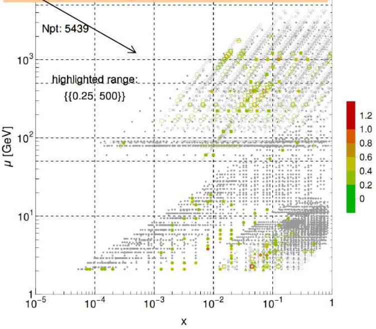
PDFSense estimates…

- ranking of strength of sensitivities of experimental data sets to PDF flavors without (re-)doing the full global fit
- impact on global fit requires both correlation and sensitivity

…kinematical distributions of sensitivities to the PDFs in the $\{x,\mu\}$ plane

for example, HERAII, CMS jets provide information on gluon and on Higgs σ

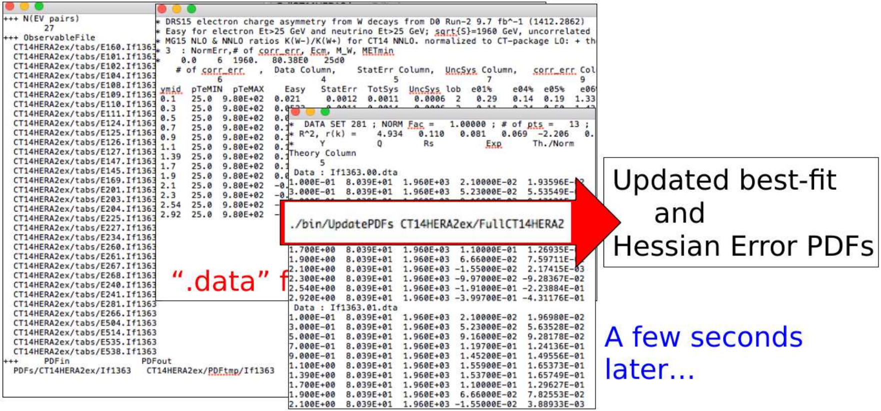
Sensitivity to the PDF error on $\sigma(pp \to HX)$ at 14 TeV



					Rankings											
No.	Exp. ID	N_d	$\sum_f \lvert S_f^E\rvert$	$\langle\sum_f \lvert S_f^E\rvert\rangle$	$\lvert S_{\bar d}^E\rvert$	$\langle\lvert S_{\bar d}^E\rvert\rangle$	$\lvert S_{\bar u}^E\rvert$	$\langle\lvert S_{\bar u}^E\rvert\rangle$	$\lvert S_g^E\rvert$	$\langle\lvert S_g^E\rvert\rangle$	$\lvert S_u^E\rvert$	$\langle\lvert S_u^E\rvert\rangle$	$\lvert S_d^E\rvert$	$\langle\lvert S_d^E\rvert\rangle$	$\lvert S_s^E\rvert$	$\langle\lvert S_s^E\rvert\rangle$
1	160	1120.	620.	0.0922	B		**A**	3	**A**	3	**A**	3	B		C	
2	111	86	218.	0.423	C	**1**	C	**1**		3	B	**1**	C	2		
3	101	337	184.	0.0909			C		C		B	3	C			
4	104	123	169.	0.229	C	2					C	2	B	2		
5	102	250	141.	0.0938	C				C	3	C	3	C	3		
6	109	96	115.	0.199	C	2	C	2		3	C	2	C	3		
7	201	119	113.	0.158	C	2	C	2				3				
8	204	184	103.	0.0935		3	C	3			C	3				
9	110	69	89.3	0.216		3		3	C	2		3		2		3
10	**545**	185	86.4	0.0779					C	3						
11	108	85	82.4	0.161		3		3		3		3	C	3		
12	538	133	66.2	0.0829					C	3						
13	**542**	158	59.	0.0622					C	3						
14	124	38	58.9	0.258		3		3				3		3	C	**1**
15	127	38	49.4	0.217		3		3				3		3	C	**1**
16	**544**	140	48.7	0.058						3						
17	126	40	48.	0.2		3		3				3		3	C	**1**
18	**250**	42	41.5	0.165		3		3		3		3		2		
19	268	41	39.6	0.161		3		3				3		3		3
20	**249**	33	39.2	0.198		2		3				3		2		3
21	514	110	36.8	0.0557						3						
22	125	33	36.7	0.185		3		3				3		3		2

PRD98 (2018) no.9, 094030 http://metapdf.hepforge.org/PDFSense/

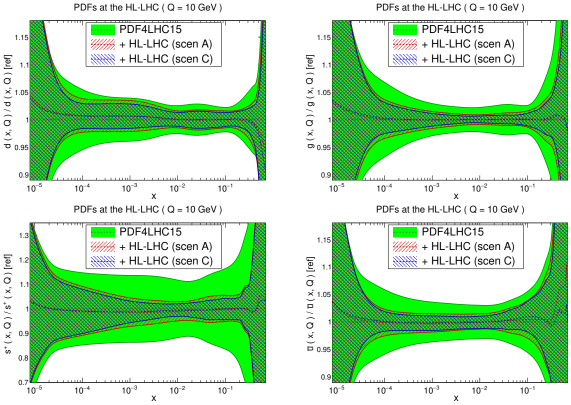
ePump



".in" file

".theory" file

PRD98(2018) no.9, 094005,1907.12177 http://hep.pa.msu.edu/epump/



- Base on Hessian reweighting method, similar to ePump.
- The prediction of PDFs at HL-LHC concerned no systematic error. Well-control systematic error is important to reach this goal.
- Physical parameters, such like α_s and $\sin^2\theta_w$, and PDF parametrization are fixed in reweighting method. Therefore, the PDF updated through reweighting method by the measurements which probing the phase space of PDF (in flavor and x) that has not been well constrained will not be reliable.

arXiv:1810.03639

CT18 parton distribution functions

This page provides numerical table files for the computation of CT18 next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) parton distribution functions. They can be interpolated with the help of a standalone Fortran interface and demonstration program , as well as the tables with interpolated values of the QCD coupling alpha_s and PDFs. A simple C++ interface for the CTEQ-TEA PDFs with CTEQ6.6 or later can be found here.

ATTENTION: The format of the PDF table files (with the extension .pds) and the Fortran interface were revised in 2012 to provide interpolation of the energy dependence of the QCD running coupling alpha_s using a table of alpha_s values stored inside the .pds file. The pre-2012 PDF table files are partly incompatible with the new format. They can be interpolated by the 2012 interface, but without computing alpha_s.

Available PDF sets

See the header of the Fortran interface for further explanations.

PDF set	Description and links to the table files	Auth
CT18 NNLO	*General-purpose NNLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . **The recommended 90% C.L. uncertainty estimate is 0.116 - 0.120.**	T.-J. Hou
CT18 NLO	*General-purpose NLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . **The recommended 90% C.L. uncertainty estimate is 0.116 - 0.120.**	T.-J. Hou
CT18Z NNLO	*General-purpose NNLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . **The recommended 90% C.L. uncertainty estimate is 0.116 - 0.120.**	T.-J. Hou
CT18Z NLO	*General-purpose NLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . **The recommended 90% C.L. uncertainty estimate is 0.116 - 0.120.**	T.-J. Hou
CT18A NNLO	*General-purpose NNLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . **The recommended 90% C.L. uncertainty estimate is 0.116 - 0.120.**	T.-J. Hou
CT18A NLO	*General-purpose NLO central set + 58 eigenvector sets (LHAPDF) *PDF sets with a varied strong coupling alpha_s(M_Z) in the ranges 0.116-0.120 and 0.110-0.124 . The recommended 90% C.L. uncertainty	T.-J. Hou

Summary

- In the LHC era, the PDF uncertainty become a dominant contribution to a measurement. It is getting more crucial for precision measurement and also new physics search.
- Precise SM measurements are key to reduce PDF uncertainty.
- PDFsense and reweighting base methods, such like ePump and xfitter profiling, are the efficient ways to help experimentalist to find out the SM observable which can help on reducing PDF uncertainty for target measurement, and analysis the SM observable first.
- Updated PDF obtained by the reweighting method share the fixed physical parameters and PDF parametrization, and thus is not equivalent to the PDF done by real global analysis. Extend data analysis need PDFs from real global analysis.