PROBLEMS IN PDF DETERMINATION

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN

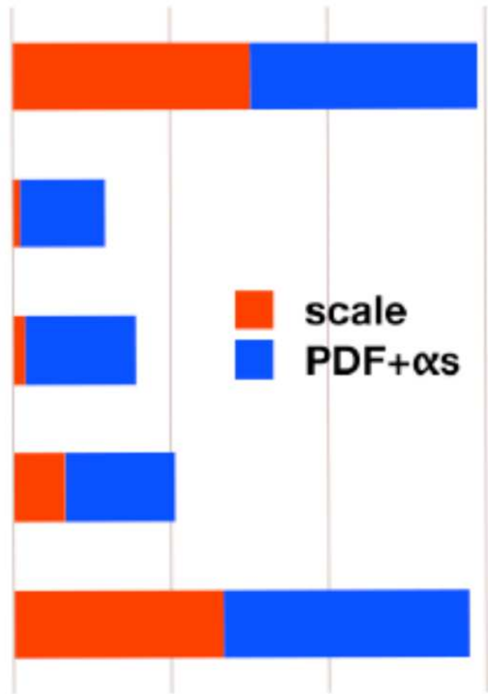
WÜRZBURG JANUARY 26, 2017

PROLOGUE

NOT SO LONG AGO

Higgs production

		σ (8 TeV)	uncertainty
NNLL QCD +NLO EW	gg→H	19.5 pb	14.7%
	VBF	1.56 pb	2.9%
NNLO QCD +NLO EW	WH	0.70 pb	3.9%
	ZH	0.39 pb	5.1%
NLO QCD	ttH	0.13 pb	14.4%



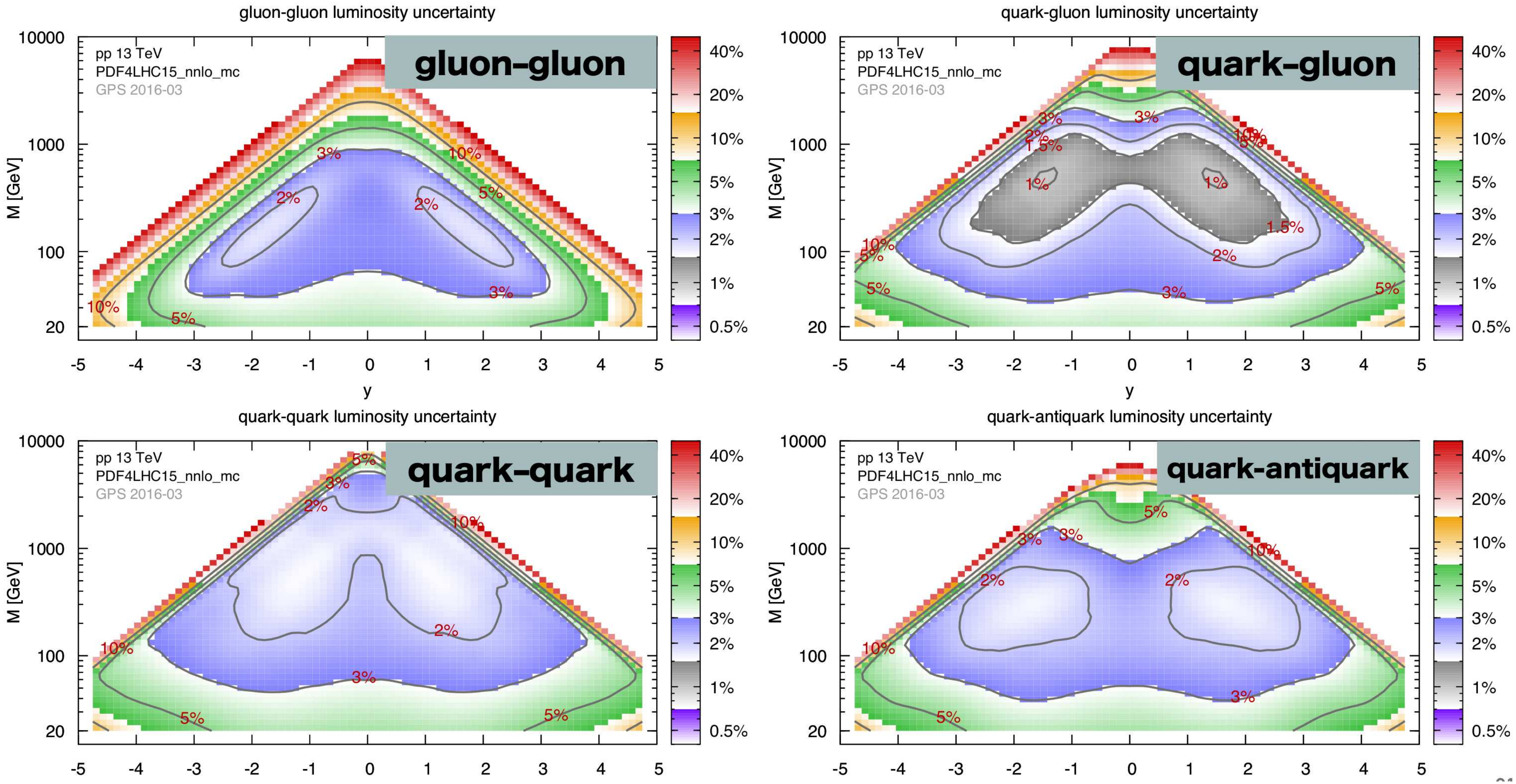
(J. Campbell, HCP2012)

PDF uncertainty either dominant, or very large, or both

typical PDF uncertainty $\sim 5 - 10\%$

NOW: THE PDF4LHC SET

LUMINOSITY UNCERTAINTIES VS RAPIDTY & MASS



G.P. Salam, LHCP2016

TYPICAL PDF UNCERTAINTY DOWN TO $\sim 2-5\%$

CAN WE BELIEVE IN 1% PDF UNCERTAINTIES?

SUMMARY

PROBLEMS SOLVED

- GLOBAL PDF FITS & WHY THE AGREE
- MONTE CARLO, HESSIAN & COMPRESSION
- COMBINED SETS
- QED PDFS

OPEN PROBLEMS

- GLOBAL FITS & WHY THEY DISAGREE
- THE ROLE OF NEW DATA
- COLLIDER PDFS: OPPORTUNITIES & PROBLEMS

NEW PROBLEMS

- THEORETICAL UNCERTAINTIES ON PDFS
- PDFS BEYOND NNLO QCD
- THE TREATMENT OF HEAVY QUARKS

DISCLAIMER

- THIS IS NOT AN UNBIASED REVIEW TALK
- PROBLEMS SOLVED ⇒ CONSENSUS OF THE COMMUNITY (PDF4LHC)
- OPEN & NEW PROBLEMS ⇒ MY OWN BIASED OPINION
  (DOES NOT REPRESENT PDF4LHC, OR EVEN NNPDF)

PROBLEMS SOLVED

CONTEMPORARY PDF TIMELINE (PUBLISHED ONLY)

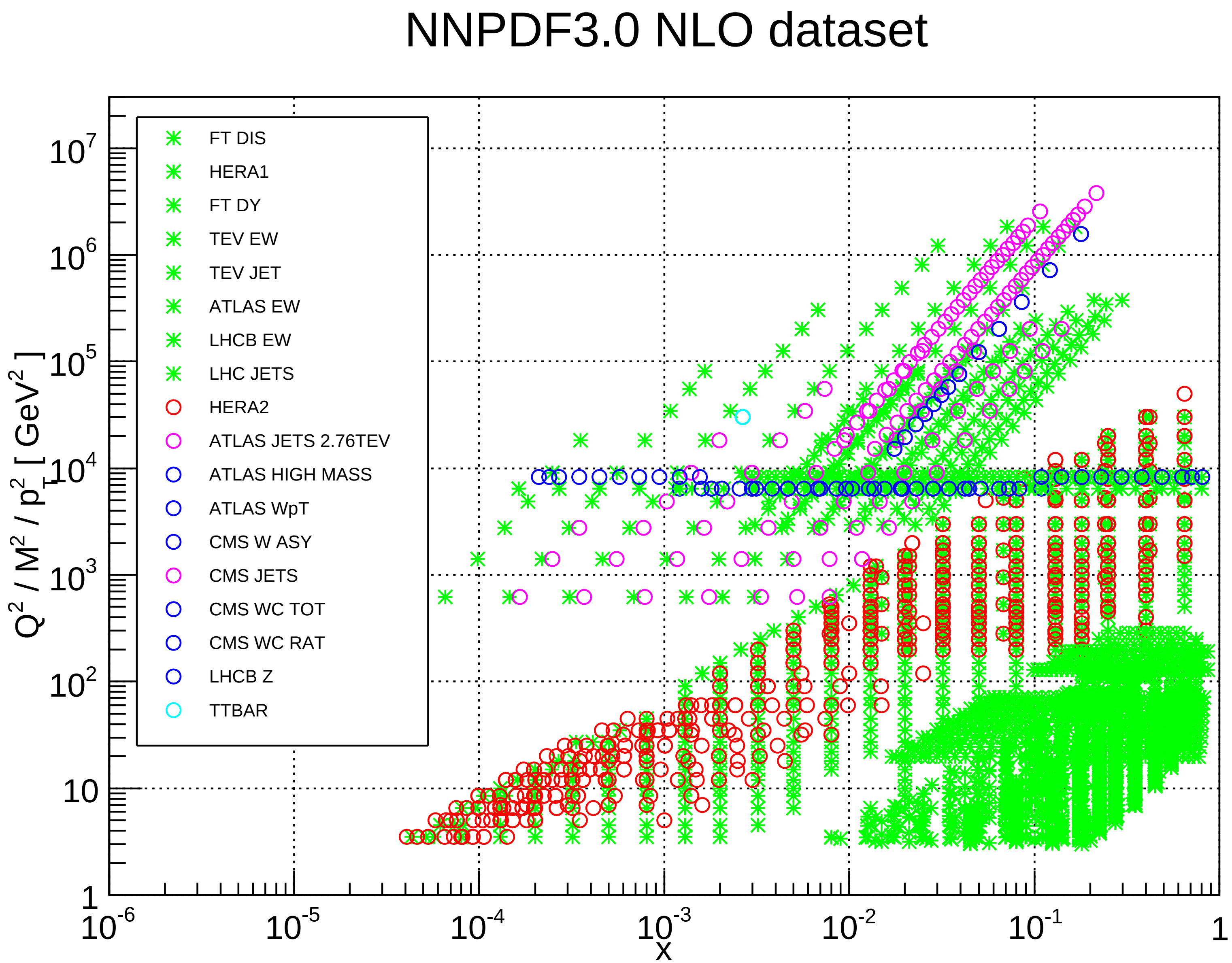
	2008		2009		2010		2011	2012		2013		2014		2015
SET MONTH	CT6.6 (02)	NN1.0 (08)	MSTW (01)	ABKM09 (08)	NN2.0 (02)	CT10(N) (07)	NN2.1(NN) (07)	ABM11 (02)	NN2.3 (07)	CT10(NN) (02)	ABM12 (10)	NN3.0 (10)	MMHT (12)	CT14 (06)
F. T. DIS	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ZEUS+H1-HI	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
COMB. HI	✗	✗	✗	✗	✔	✗	✔	✗	✔	✗	✔	✔	✗	✗
ZEUS+H1-HII	✗	✗	✗	✗	✗	✗	some	✗	✗	some	✗	✔	✗	✗
HERA JETS	✗	✗	✔	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗
F. T. DY	✔	✗	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
TEV. W+Z	✔	✗	✔	✗	✔	✔	✔	✗	✔	✔		✔	✔	✔
TEV. JETS	✔	✗	✔	✗	✔	✔	✗	✔	✔	✔	✗	✔	✔	✔
LHC W+Z	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	some	✔	✔	✔
LHC JETS	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗	✔	✔	✔
TOP	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✔	✗	✗
W+C	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗
W p_T	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗

- INCREASINGLY WIDE DATASET USED FOR PDF DETERMINATION
- HERAPDF: ONLY HERA STRUCTURE FUNCTION DATA ⇒ EXTREME CONSISTENCY
- MANY THEORETICAL AND METHODOLOGICAL IMPROVEMENTS:
  - MSTW, ABKM: ALL NNLO; NNPDF NNLO SINCE 07/11 (2.1), CT SINCE 02/13 (CT10)
  - MSTW, CT ALL MATCHED HEAVY QUARK SCHEMES; NNPDF GM-VFN SINCE 01/11 (2.1)

GLOBAL FITS: THE DATASET IN DETAIL

	NNPDF3.0	MMHT14	CT14
SLAC P,D DIS	✔	✔	✗
BCDMS P,D DIS	✔	✔	✔
NMC P,D DIS	✔	✔	✔
E665 P,D DIS	✗	✔	✗
CDHSW NU-DIS	✗	✗	✔
CCFR NU-DIS	✗	✔	✔
CHORUS NU-DIS	✔	✔	✗
CCFR DIMUON	✗	✔	✔
NUTEV DIMUON	✔	✔	✔
HERA I NC,CC	✔	✔	✔
HERA I CHARM	✔	✔	✔
H1,ZEUS JETS	✗	✔	✗
H1 HERA II	✔	✗	✗
ZEUS HERA II	✔	✗	✗
E605 & E866 FT DY	✔	✔	✔
CDF & D0 W ASYM	✗	✔	✔
CDF & D0 Z RAP	✔	✔	✔
CDF RUN-II JETS	✔	✔	✔
D0 RUN-II JETS	✗	✔	✔
D0 RUN-II W ASYM	✗	✗	✔
ATLAS HIGH-MASS DY	✔	✔	✔
CMS 2D DY	✔	✔	✗
ATLAS W,Z RAP	✔	✔	✔
ATLAS W p_T	✔	✗	✗
CMS W ASY	✔	✔	✔
CMS W +C	✔	✗	✗
LHCB W,Z RAP	✔	✔	✔
ATLAS JETS	✔	✔	✔
CMS JETS	✔	✔	✔
TTBAR TOT XSEC	✔	✔	✗
TOTAL NLO	4276	2996	3248
TOTAL NNLO	4078	2663	3045

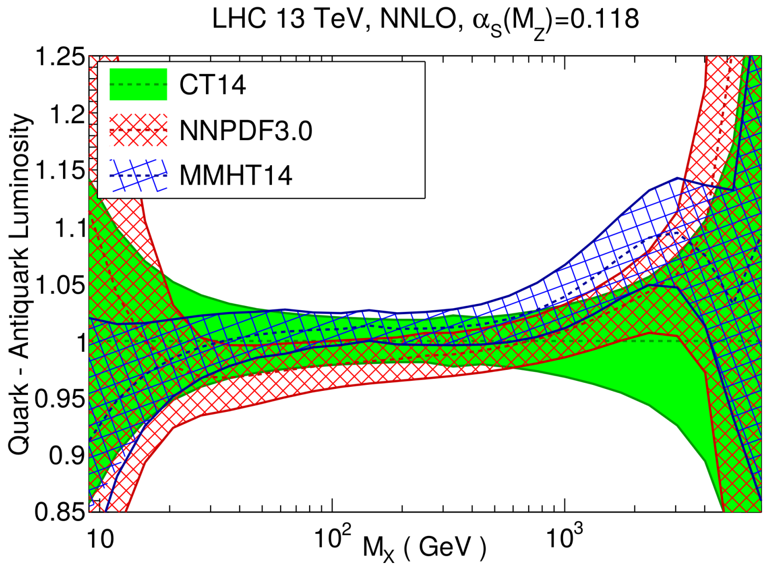
THE NNPDF3.0 DATASET



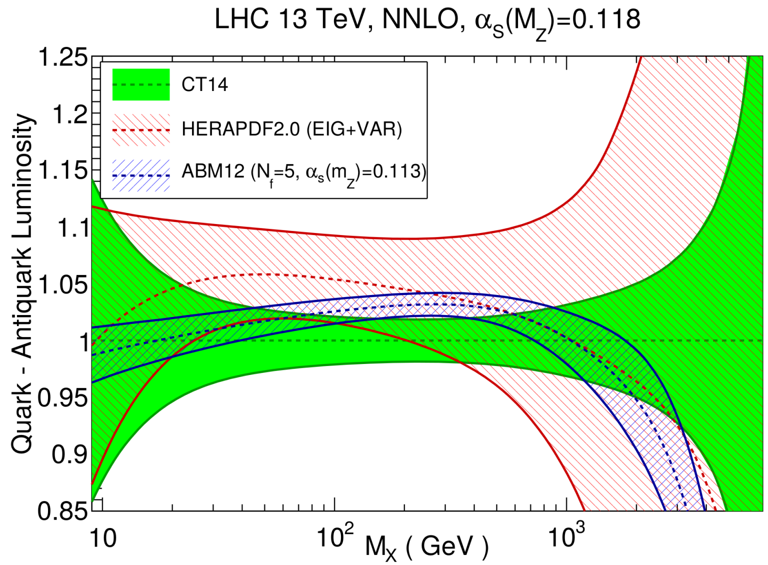
PARTON LUMINOSITIES (LHC 13)

QUARK-ANTIQUARK

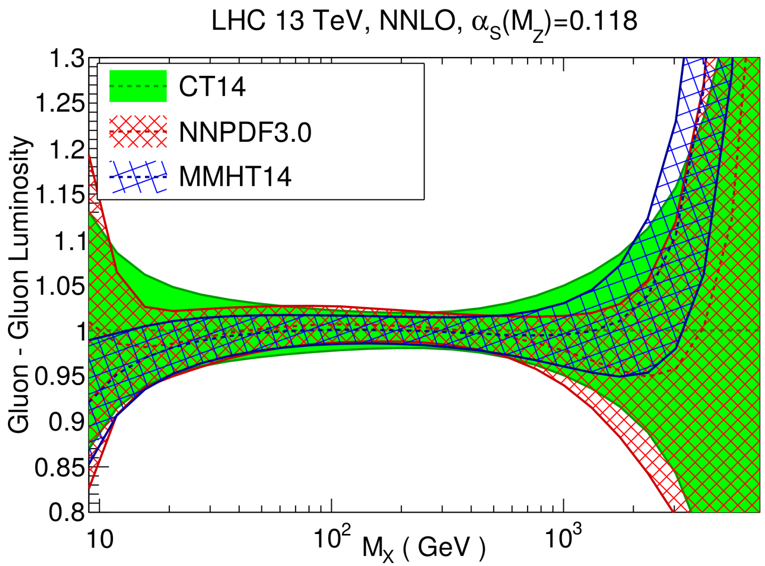
GLOBAL

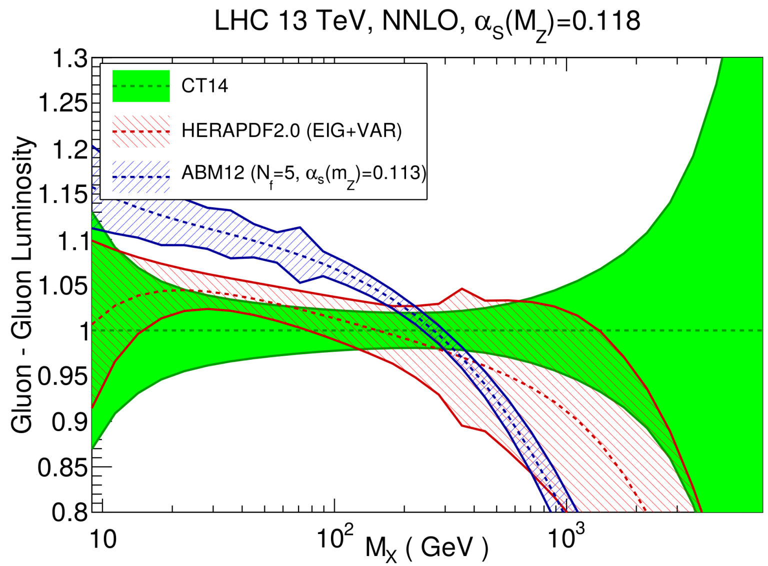


REDUCED



GLUON-GLUON



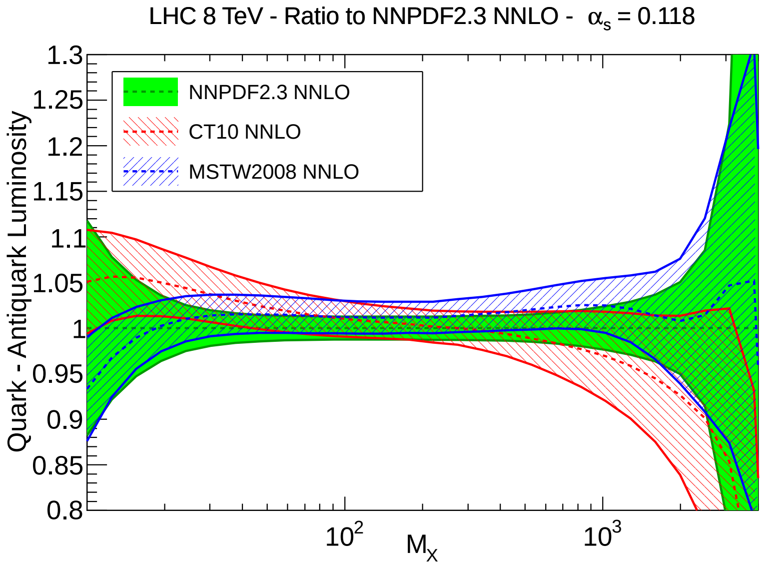


- GLOBAL FITS AGREE WELL
- FITS BASED ON REDUCED DATASET HAVE EITHER LARGE UNCERTAINTIES OR SHOW SIZABLE DEVIATIONS

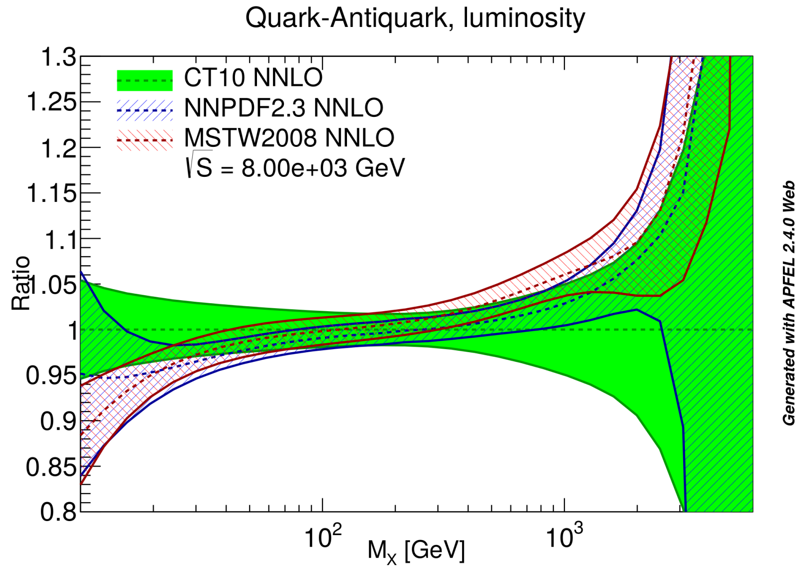
PARTON LUMINOSITIES (LHC 8)

QUARK-ANTIQUARK

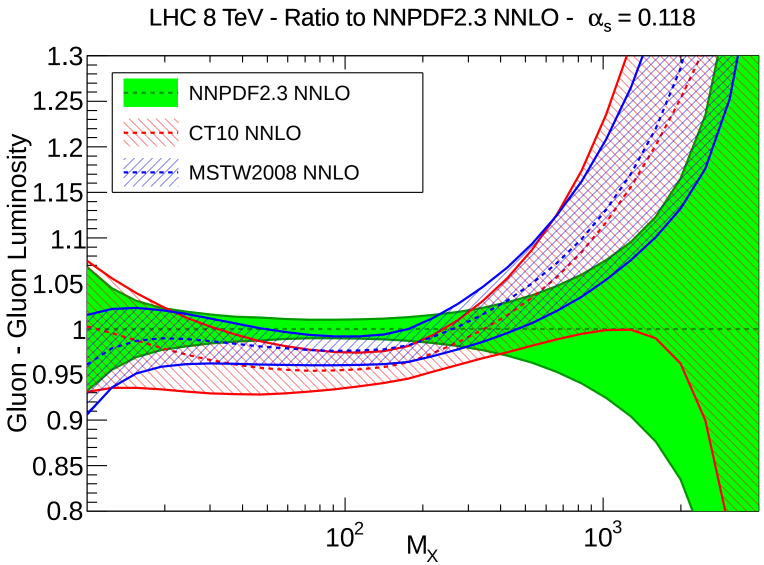
2012

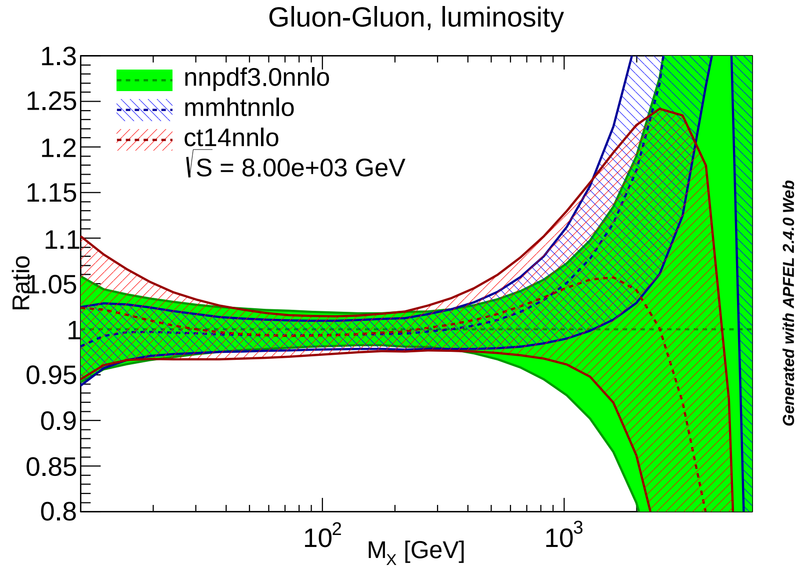


2015



GLUON-GLUON





- LONGSTANDING DISCREPANCY BETWEEN GLUON LUMINOSITIES IS GONE $\Rightarrow$ IMPACT ON HIGGS
- UNCERTAINTIES BLOW UP FOR LIGHT ($\lesssim 10$ GeV) OR HEAVY ($\gtrsim 1$ TeV) FINAL STATES $\Rightarrow$ IMPACT ON SEARCHES

PROGRESS

- Q: WHY ARE PDF UNCERTAINTIES ON GLOBAL FITS OF SIMLAR SIZE?
  - SIMILAR DATASETS
  - BUT DIFFERENT PROCEDURES
- A: UNCERTAINTY TUNED TO DATA THROUGH TOLERANCE (MMHT & CT) OR CLOSURE TESTING (NNPDF)
- Q: WHAT HAS DRIVEN THE IMPROVED AGREEMENT OF GLOBAL FITS
  - SIMILAR DATASETS
  - BUT DIFFERENT PROCEDURES
- A: DATA+METHODOLOGY

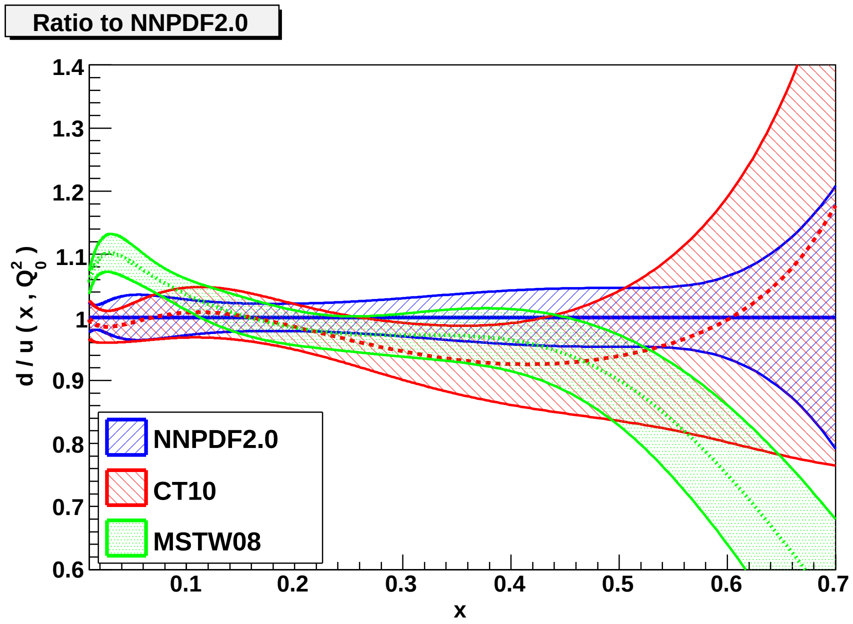
METHODOLOGY

	NNPDF3.0	MMHT14	CT14
No. OF FITTED PDFs	7	7	6
PARAMETRIZATION	NEURAL NETS	$x^a(1-x)^b\times$ CHEBYSCHEV	$x^a(1-x)^b\times$BERNSTEIN
FREE PARAMETERS	259	37	30-35
UNCERTAINTIES	REPLICAS	HESSIAN	HESSIAN
TOLERANCE	NONE	DYNAMICAL	DYNAMICAL
CLOSURE TEST	✔	✗	✗
REWEIGHTING	REPLICAS	EIGENVECTORS	EIGENVECTORS

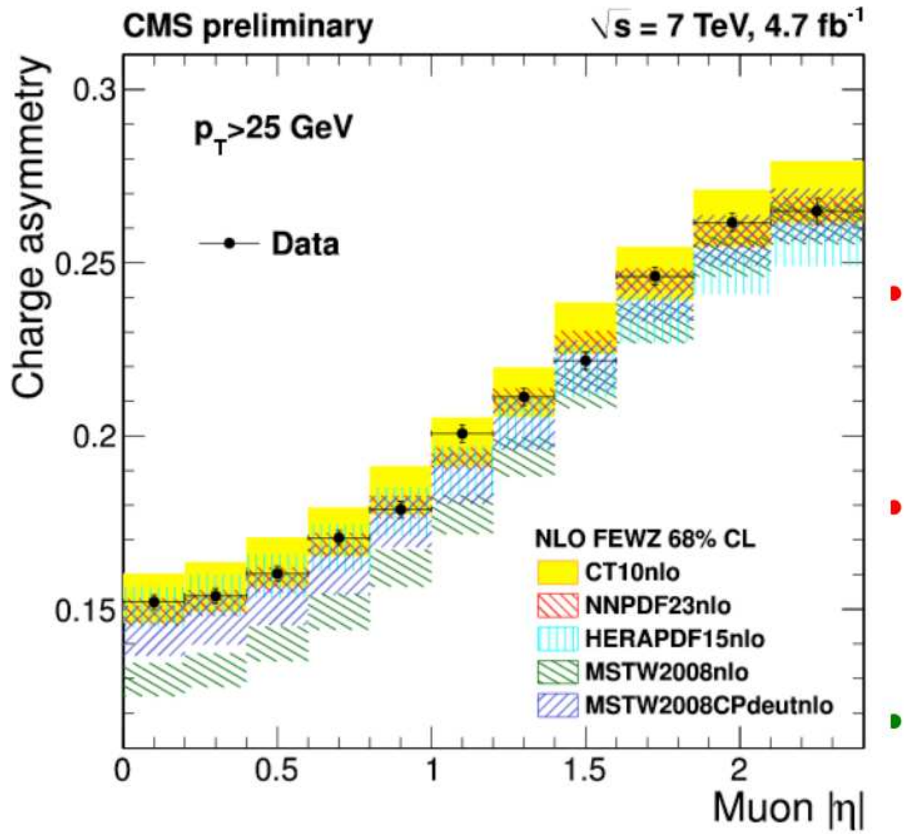
- MMHT, CT10 LARGER # OF PARMS., ORTHOGONAL POLYNOMIALS
- NNPDF CLOSURE TEST

EXAMPLE OF DATA-DRIVEN PROGRESS
MSTW/MMHT: THE d/u RATIO

THE d/u RATIO



THE CMS W ASYMMETRY

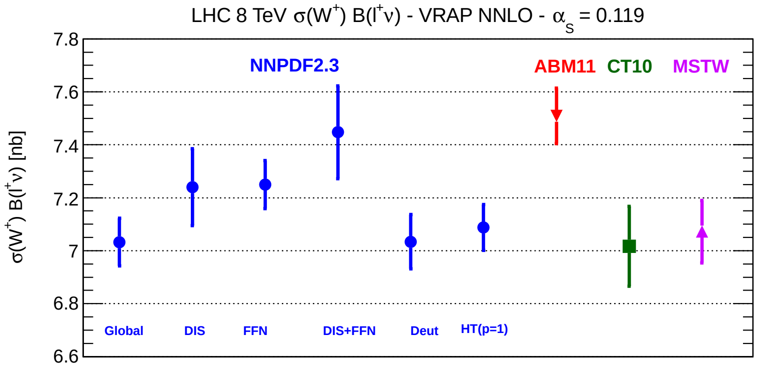


- LONG-STANDING DISCREPANCY IN THE d/u RATIO BETWEEN MSTW AND OTHER GLOBAL FITS
- RESOLVED BY W ASYMMETRY DATA
- EXPLAINED BY INSUFFICIENTLY FLEXIBLE PDF PARAMETRIZATION $\Rightarrow$ FIXED IN MSTW08DEUT/MMHT

WHY DO SOME PDF SETS DISAGREE?
FFN PDFs

- SOME PDF SETS ADOPT A FFN SCHEME (ABM, JR)
- ABM ALSO INCLUDES HIGHER TWIST & NUCLEAR CORRECTIONS
- ALSO, ABM MOSTLY BASED ON DIS DATA
- NNPDF WITH FFN &DIS DATA SET AGREES WITH ABM;
  HIGHER TWIST & NUCLEAR CORRRECTIONS HAVE SMALL & LOCALIZED EFFECT;
- FFN EVOLUTION WEAKER $\Rightarrow$ GLUON DISTORTED analytic argument by R.Thorne, 2012

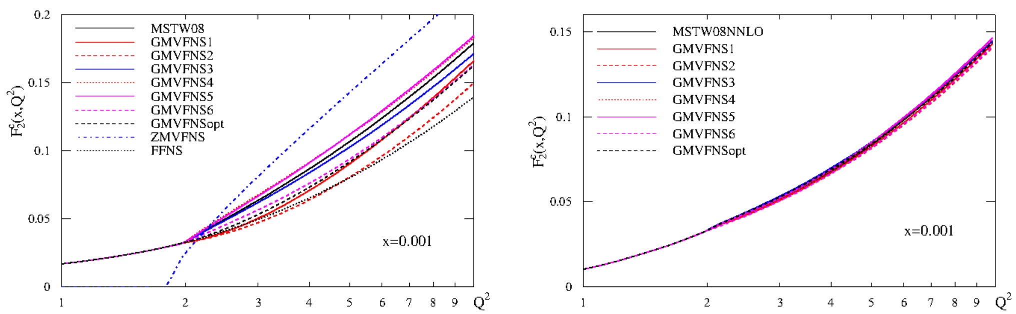
W^+



(NNPDF, 2013)

F_2 FFN VS VFN

NLO NNLO

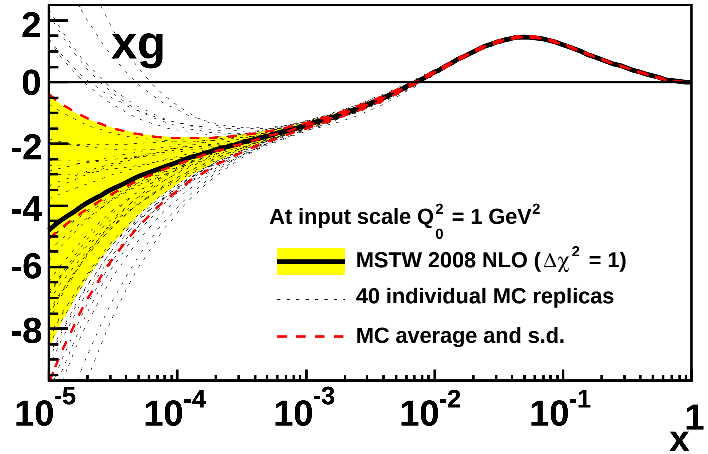


(Thorne, 2012)

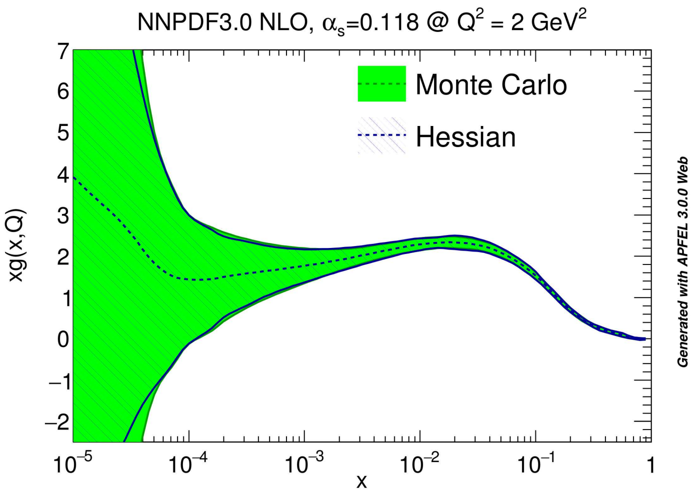
- ABM $\Rightarrow$ NO TOLERANCE
- ABM $\Rightarrow \alpha_s(M_z) = 0.113 \pm 0.001$

MC ⇔ HESSIAN

- TO CONVERT HESSIAN INTO MONTECARLO GENERATE MULTIGAUSSIAN REPLICAS IN PARAMETER SPACE
- ACCURATE WHEN NUMBER OF REPLICAS SIMILAR TO THAT WHICH REPRODUCES DATA



(Thorne, Watt, 2012)

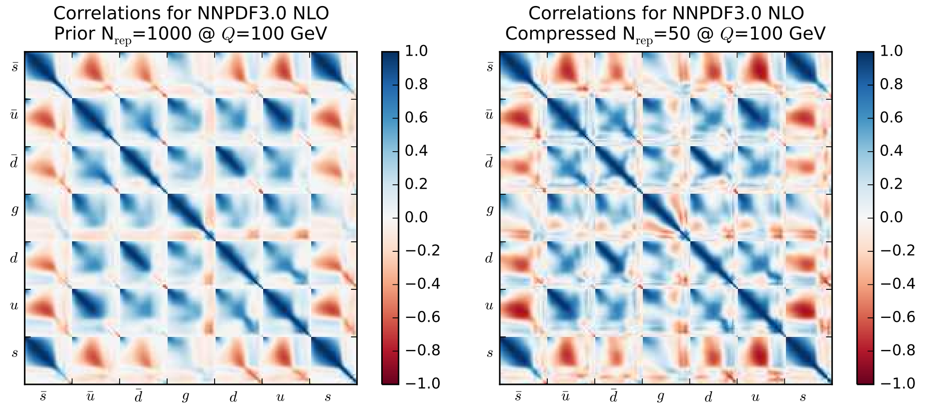


(Carrazza, SF, Kassabov, Rojo, 2015)

- TO CONVERT MONTE CARLO INTO HESSIAN, SAMPLE THE REPLICAS $f_i(x)$ AT A DISCRETE SET OF POINTS & CONSTRUCT THE ENSUING COVARIANCE MATRIX
- EIGENVECTORS OF THE COVARIANCE MATRIX AS A BASIS IN THE VECTOR SPACE SPANNED BY THE REPLICAS BY SINGULAR-VALUE DECOMPOSITION
- NUMBER OF DOMINANT EIGENVECTORS SIMILAR TO NUMBER OF REPLICAS ⇒ ACCURATE REPRESENTATION

COMPRESSION

MONTECARLO



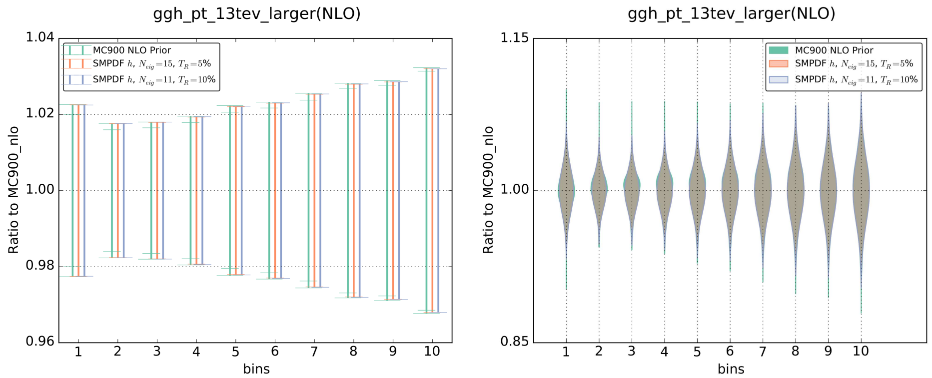
(Carrazza, Latorre, Kassabov, Rojo, 2015)

- CONSTRUCT A VERY LARGE REPLICA SAMPLE
- SELECT (BY GENETIC ALGORITHM) A SUBSET OF REPLICAS WHOSE STATISTICAL FEATURES ARE AS CLOSE AS POSSIBLE TO THOSE OF THE PRIOR
- $\Rightarrow$ FOR ALL PDFs ON A GRID OF POINTS// MINIMIZE DIFFERENCE OF: FIRST FOUR MOMENTS, CORRELATIONS; OUTPUT OF KOLMOGOROV-SMIRNOV TEST (NUMBER OF REPLICAS BETWEEN MEAN AND σ, 2σ, INFINITY)
- 50 COMPRESSED REPLICA REPRODUCE 1000 REPLICA SET TO PRECENT ACCURACY

CAN REPRODUCE NONGAUSSIAN FEATURES WITH REASONABLY SMALL REPLICA SAMPLE

HESSIAN

- SELECT SUBSET OF THE COVARIANCE MATRIX CORRELATED TO A GIVEN SET OF PROCESSES
- PERFORM SVD ON THE REDUCED COVARIANCE MATRIX, SELECT DOMINANT EIGENVECTOR, PROJECT OUT ORTHOGONAL SUBSPACE
- ITERATE UNTIL DESIRED ACCURACY REACHED
- CAN ADD PROCESSES TO GIVEN SET; CAN COMBINE DIFFERENT OPTIMIZED SETS
- 15 EIGENVECTORS DESCRIBE ALL HIGGS MODES + JETS + W, Z PRODUCTION



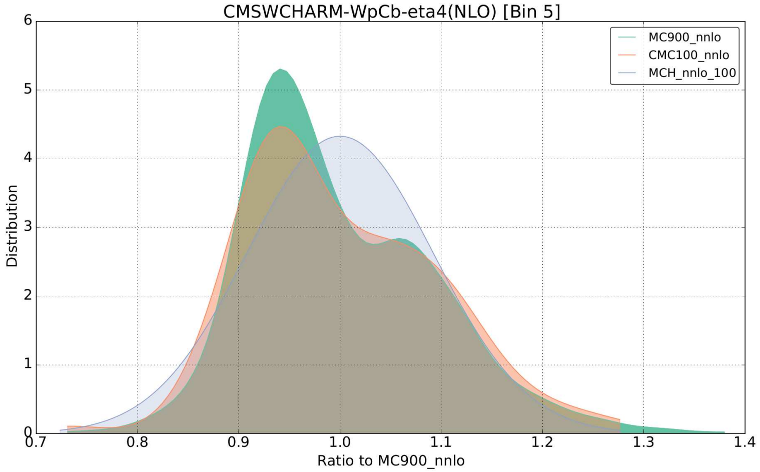
(Carrazza, SF, Kassabov, Rojo, 2016)

VERY SMALL NUMBER OF EVECS; CAN COMBINE WITH NUISANCE PARMS

NONGAUSSIAN BEHAVIOUR

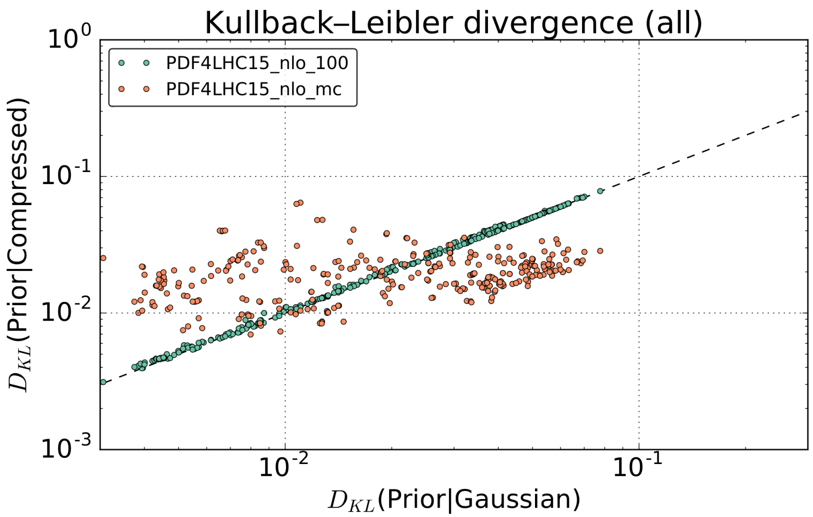
Monte Carlo compared to Hessian

CMS $W + c$ production



- DEVIATION FROM GAUSSIANITY E.G. AT LARGE x DUE TO LARGE UNCERTAINTY + POSITIVITY BOUNDS
$\Rightarrow$ RELEVANT FOR SEARCHES
- CANNOT BE REPRODUCED IN HESSIAN FRAMEWORK
- WELL REPRODUCED BY COMPRESSED MC

- DEFINE KULLBACK-LEIBLER DIVERGENCE $D_{\rm KL} = \int_{-\infty}^{\infty} P(x) \frac{\ln P(x)}{\ln Q(x)}\, dx$ BETWEEN A PRIOR P AND ITS REPRESENTATION Q
- $D_{\rm KL}$ BETWEEN PRIOR AND HESSIAN DEPENDS ON DEGREE OF GAUSSIANITY
- $D_{\rm KL}$ BETWEEN PRIOR AND COMPRESSED MC DOES NOT



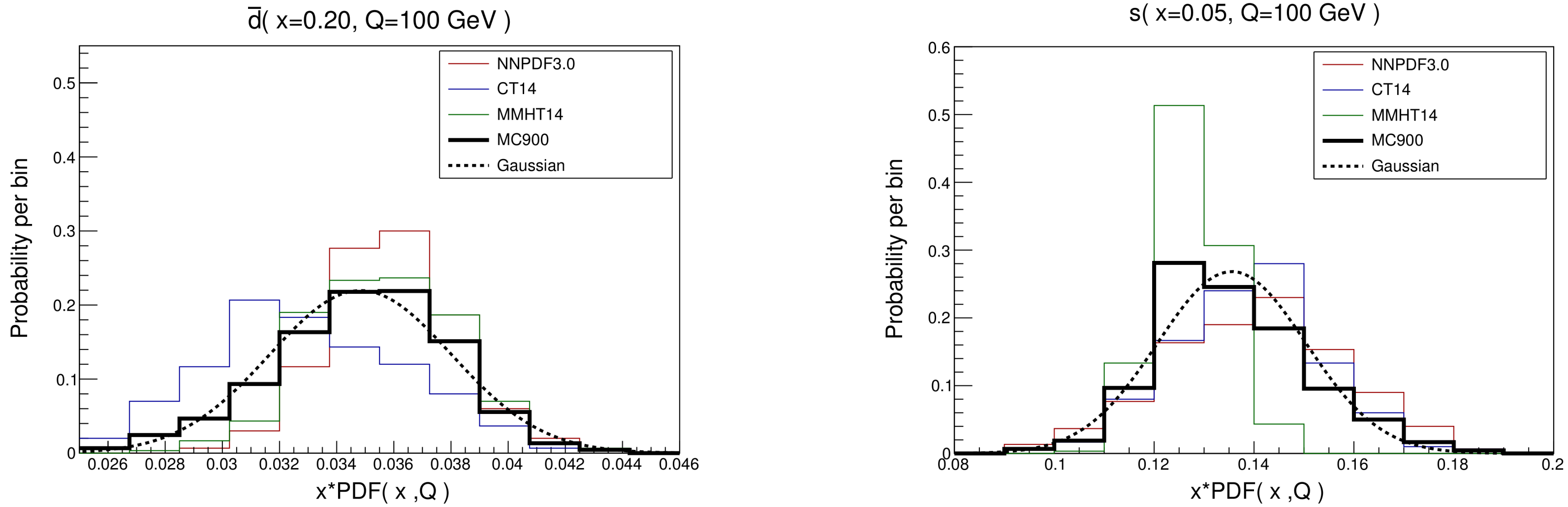
CAN GAUGE WHEN MC IS MORE ADVANTAGEOUS THAN HESSIAN!

MONTE CARLO COMBINATION

(Watt, S.F., 2010-2013)

- MAY COMBINE DIFFERENT PDF SETS, AFTER MC CONVERSION OF HESSIAN SETS
- COMBINE MONTE CARLO REPLICAS INTO SINGLE SET
- COMBINED SET APPROXIMATELY GAUSSIAN

COMBINED PDF4LHC SETS FOR ANTIDOWN & STRANGE

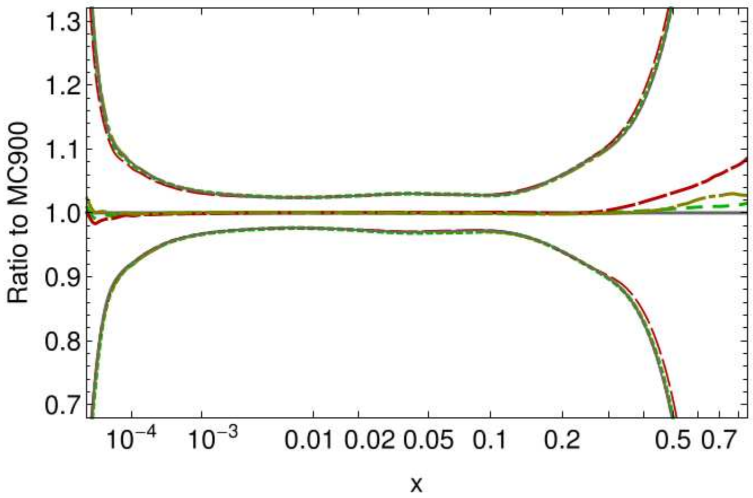


PDF4LHC15 COMBINATION

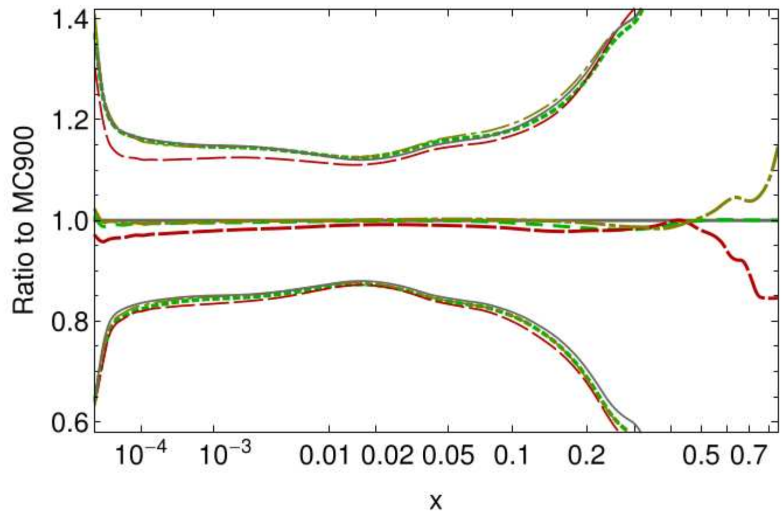
- INCLUDES CT14, MMHT, NNPDF3.0
- 900 REPLICAS (300 FOR EACH SET) ENSURE PRECENTAGE ACCURACY ON ALL QUANTITIES

300, 900,1800 REPLICAS
(RATIO TO 900)

gluon



strange



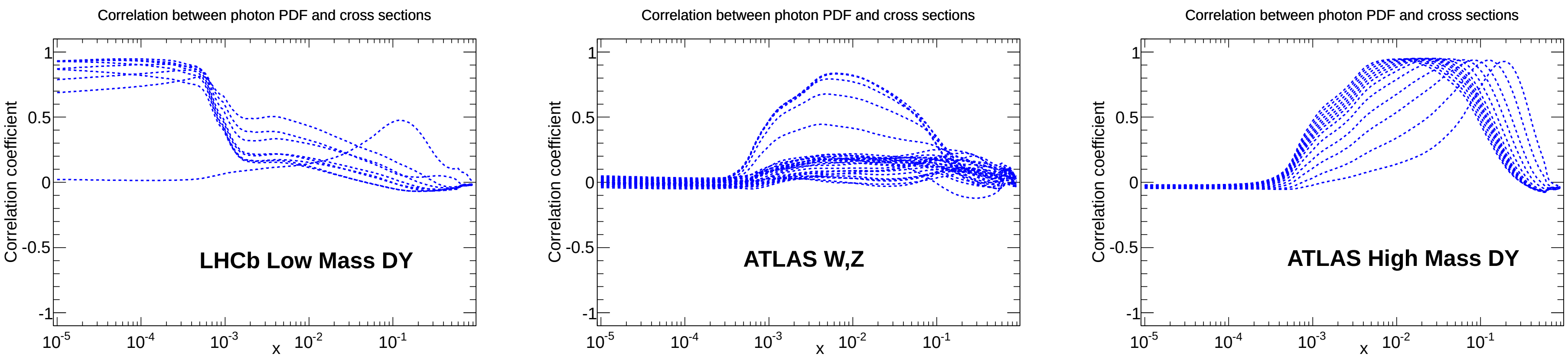
THE PHOTON PDF

NNPDF2.3QED/NNPDF3.0QED DATASET

Dataset	Observable	N_{dat}	$[\eta_{\text{min}}, \eta_{\text{max}}]$	$[M_{\text{ll}}^{\text{min}}, M_{\text{ll}}^{\text{max}}]$
LHCb γ^*/Z Low Mass	$d\sigma(Z)/dM_{ll}$	9	[2,4.5]	[5,120] GeV
ATLAS W, Z	$d\sigma(W^{\pm}, Z)/d\eta$	30	[-2.5,2.5]	[60,120] GeV
ATLAS γ^*/Z High Mass	$d\sigma(Z)/dM_{ll}$	13	[-2.5,2.5]	[116,1500] GeV

IMPACT

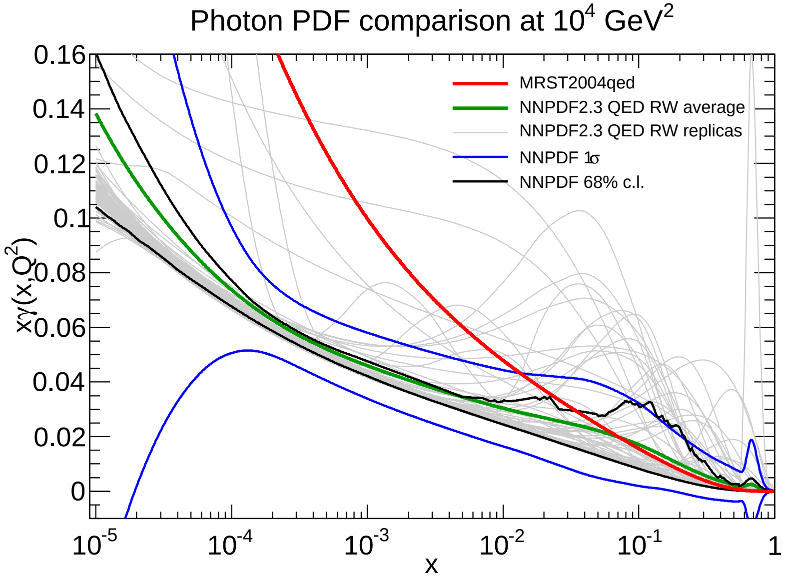
CORRELATION BETWEEN DATA AND γ PDF

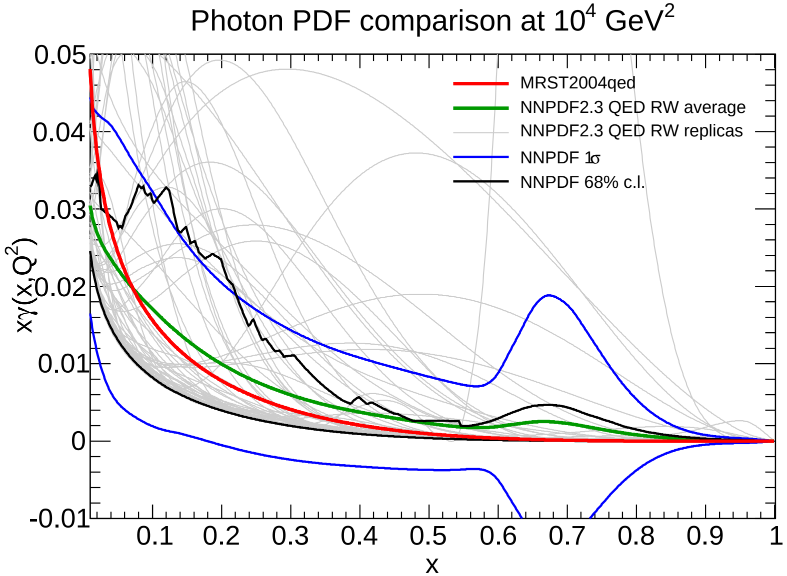


THE PHOTON PDF FROM DATA

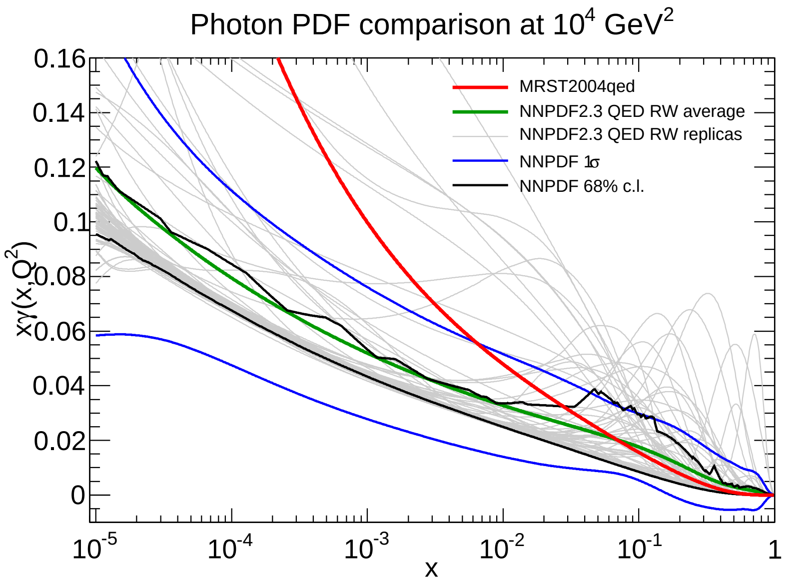
NNPDF2.3QED-NNPDF3.0QED

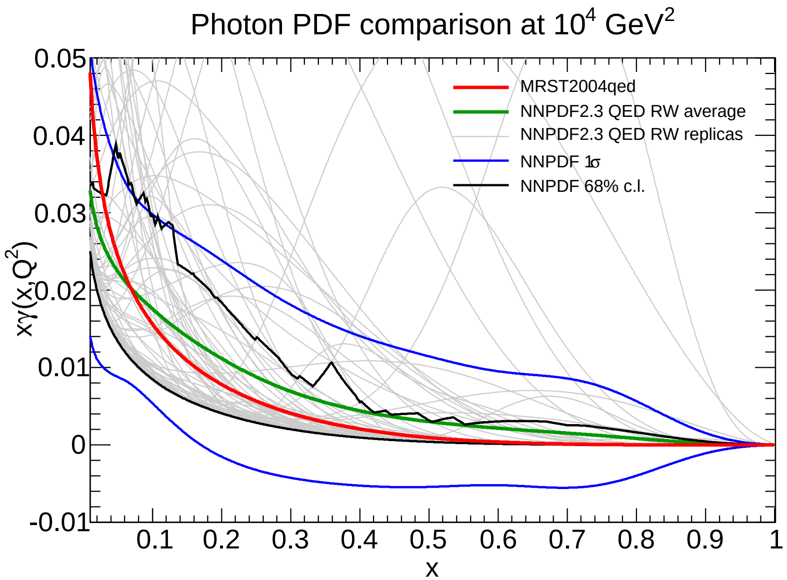
NLO RESULTS





NNLO RESULTS



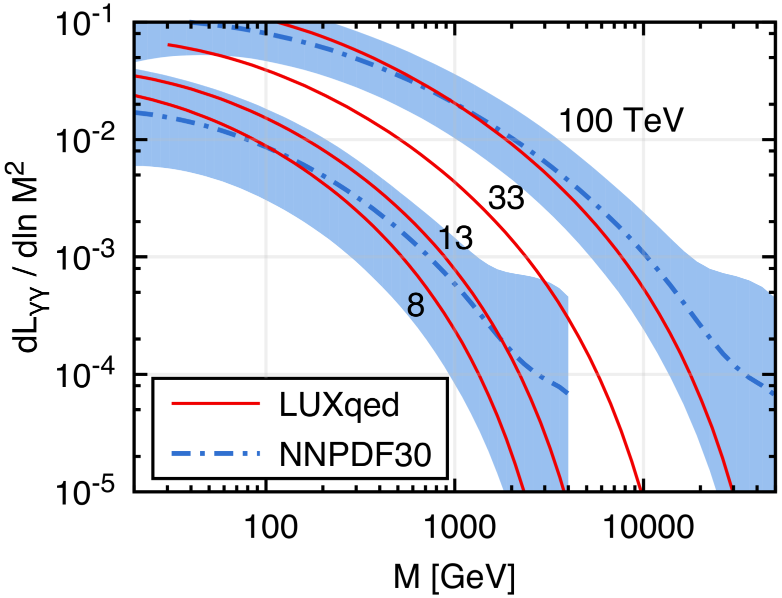


THE PHOTON PDF BREAKTHROUGH

(Manohar, Nason, Salam, Zanderighi, 2016)

- QED IS PERTURBATIVE DOWN TO LOW SCALES ⇒ THE PHOTON PDF MUST BE COMPUTABLE IF THE INPUT QUARK SUBSTRUCTURE IS KNOWN
- WRITE THE CROSS-SECTION FOR A CHOSEN PROCESS: SUSY PRODUCTION IN EP COLLISION (Drees, Zeppenfeld, 1989)
- COMPUTE IT DIRECTLY, OR USING THE PHOTON PDF
- ⇒ PDF EXPRESSED IN TERMS OF THE STRUCTURE FUNCTION INTEGRATED OVER ALL SCALES, INCLUDING ELASTIC FORM FACTORS

$$xf_{\gamma/p}(x,\mu^2) = \frac{1}{2\pi\alpha(\mu^2)}\int_x^1 \frac{dz}{z}\left\{\int_{\frac{x^2m_p^2}{1-z}}^{\frac{\mu^2}{1-z}} \frac{dQ^2}{Q^2}\alpha^2(Q^2)\left[\left(zp_{\gamma q}(z)+\frac{2x^2m_p^2}{Q^2}\right)F_2(x/z,Q^2)-z^2F_L\left(\frac{x}{z},Q^2\right)\right] - \alpha^2(\mu^2)z^2F_2\left(\frac{x}{z},\mu^2\right)\right\},$$



OLD PROBLEMS

LESSONS LEARNT

- THE RELIABILITY OF PDF SETS CRUCIALLY DEPENDS ON THE INCLUSION OF A WIDE ENOUGH DATASET
- FFN PDFs UNRELIABLE
- CAN FREELY CHOOSE MC OR HESSIAN PDF UNCERTAINTIES:
  - MC ALLOWS REPRODUCING NON-GAUSSIAN BEHAVIOUR
  - HESSIAN ALLOWS FOR OPTIMAL ACCURATE REPRESENTATION OF GAUSSIAN UNCERTAINTIES
  - COMPRESSION METHODS AVAILABLE IN BOTH CASES
  - THE PHOTON PDF CAN BE COMPUTED FROM THE PROTON STRUCTURE FUNCTION

OPEN PROBLEMS

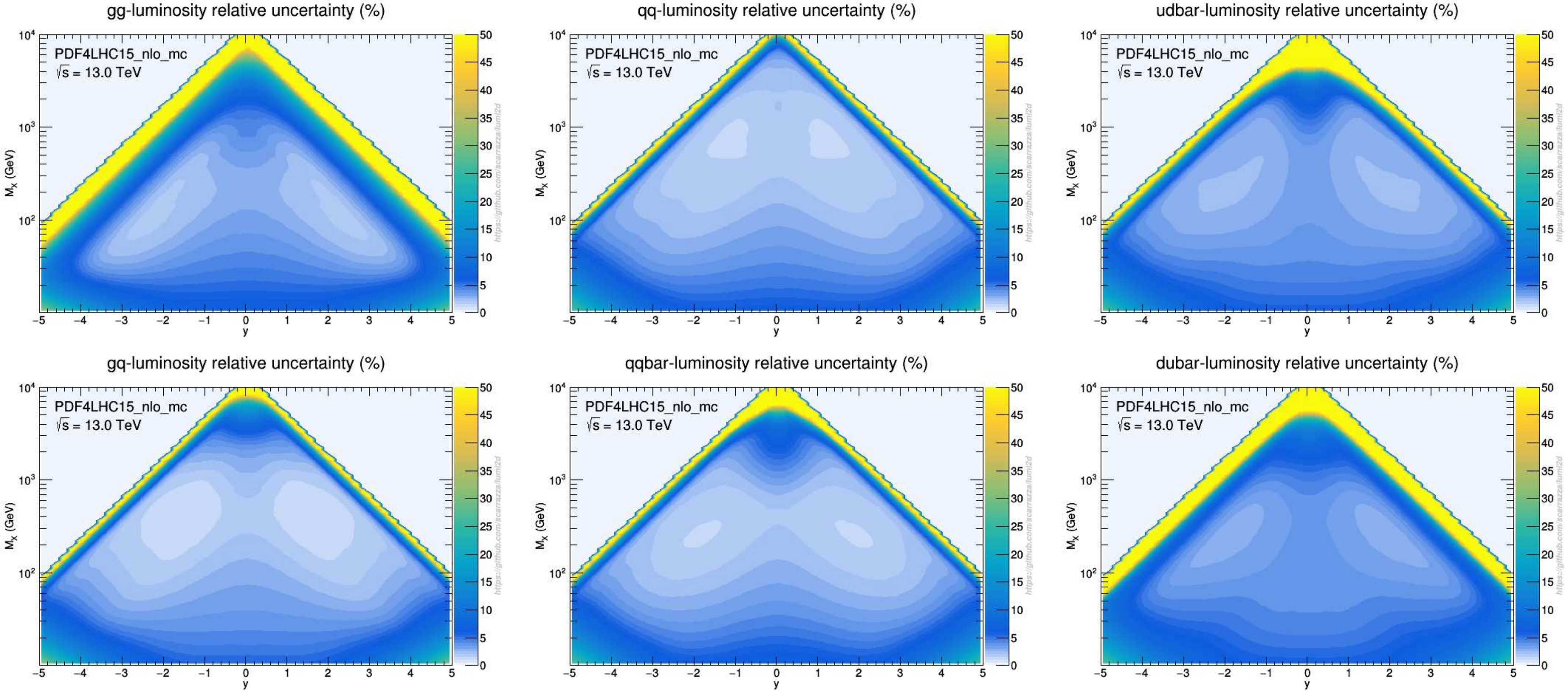
PDF UNCERTAINTIES: THE STATE OF THE ART

THE KARLSRUHE PLOTS

PDF UNCERTAINTIES: THE STATE OF THE ART

(PDF4LHC15, NLO)

GLUON SINGLET FLAVORS

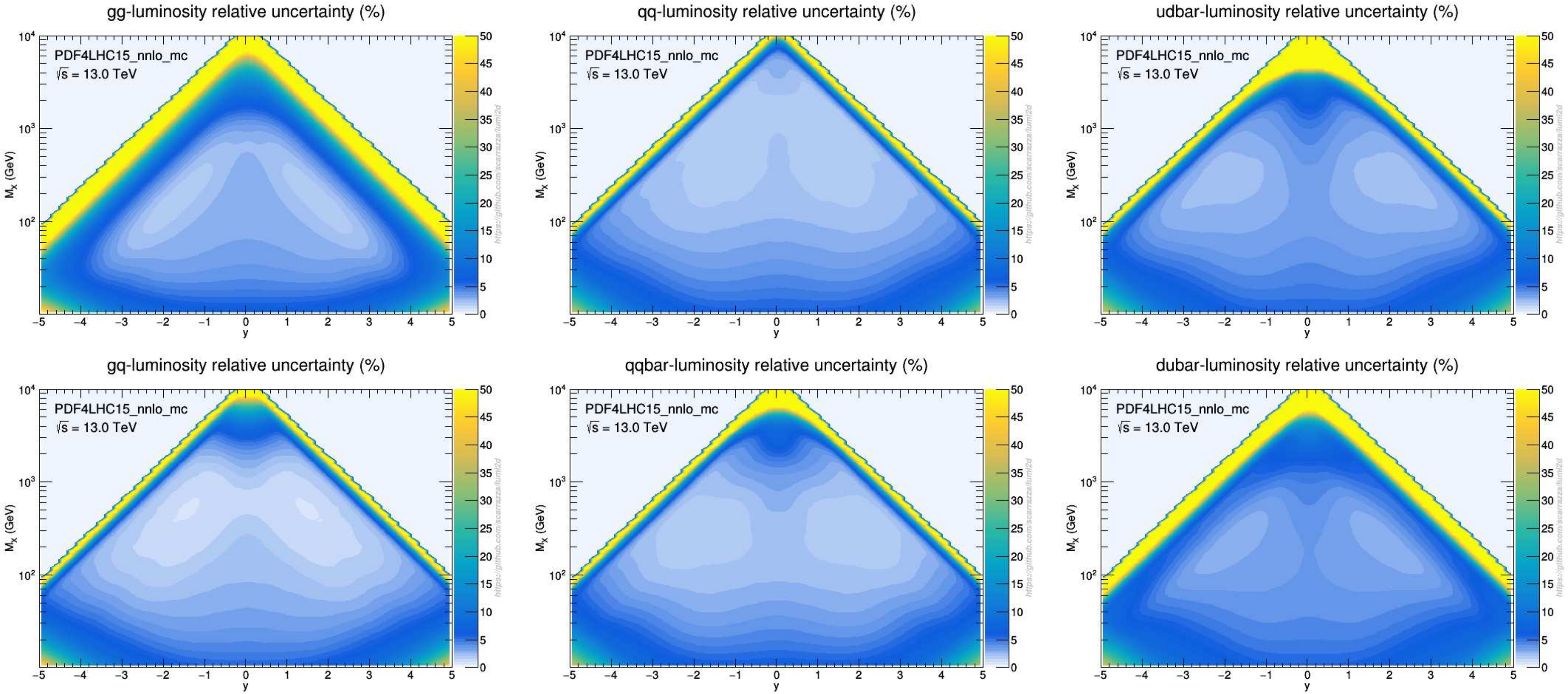


- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- SWEET SPOT: VALENCE Q - G; UNCERTAINTIES DOWN TO 1%
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS

• PDF UNCERTAINTIES: THE STATE OF THE ART

(PDF4LHC15, NNLO)

GLUON SINGLET FLAVORS



- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- SWEET SPOT: VALENCE Q - G; UNCERTAINTIES DOWN TO 1%
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS
- NO QUALITATIVE DIFFERENCE BETWEEN NLO AND NNLO

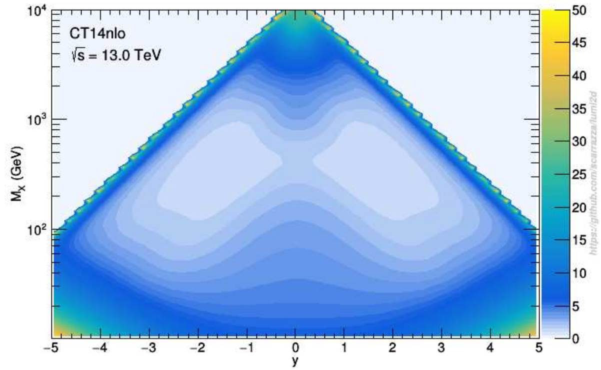
CAN WE BELIEVE IN 1% PDF UNCERTAINTIES?

WHAT'S THE UNCERTAINTY ON THE PDF UNCERTAINTY?

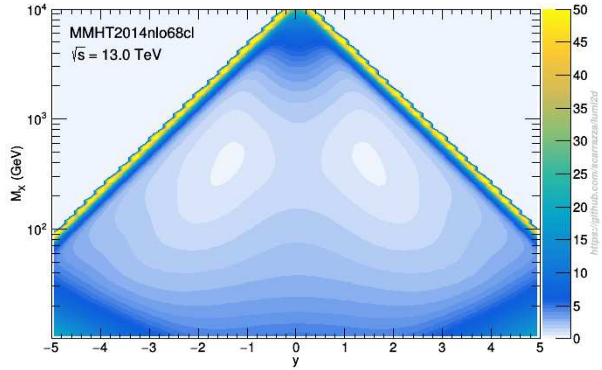
qg

CT14	MMHT	NNPDF3.0

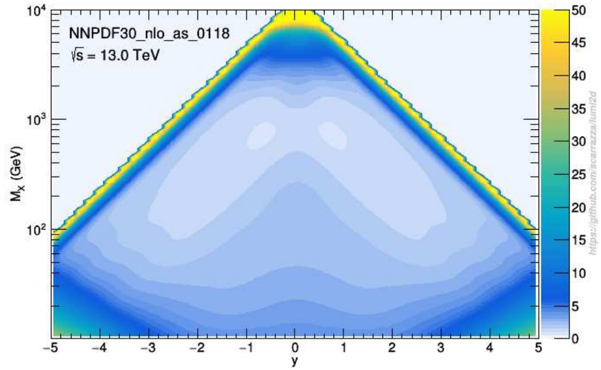
gq-luminosity relative uncertainty (%)



gq-luminosity relative uncertainty (%)

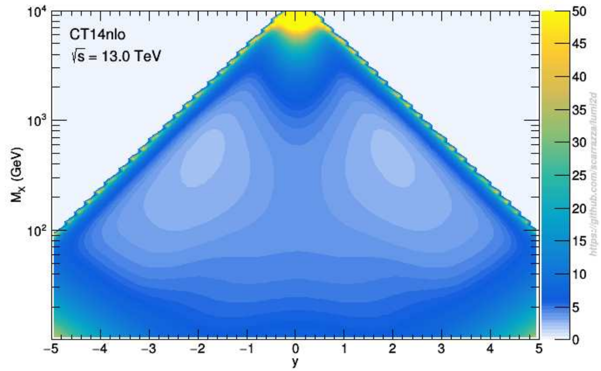


gq-luminosity relative uncertainty (%)

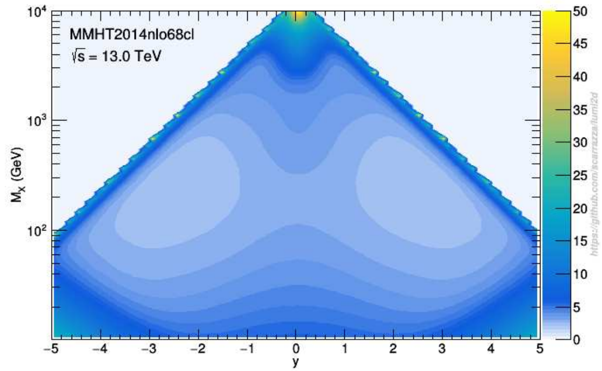


$u\bar{d}$

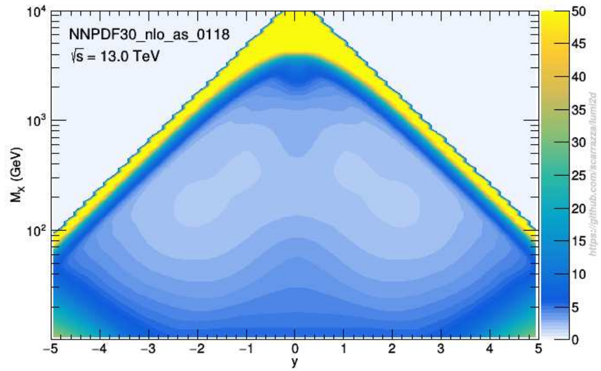
udbar-luminosity relative uncertainty (%)



udbar-luminosity relative uncertainty (%)



udbar-luminosity relative uncertainty (%)

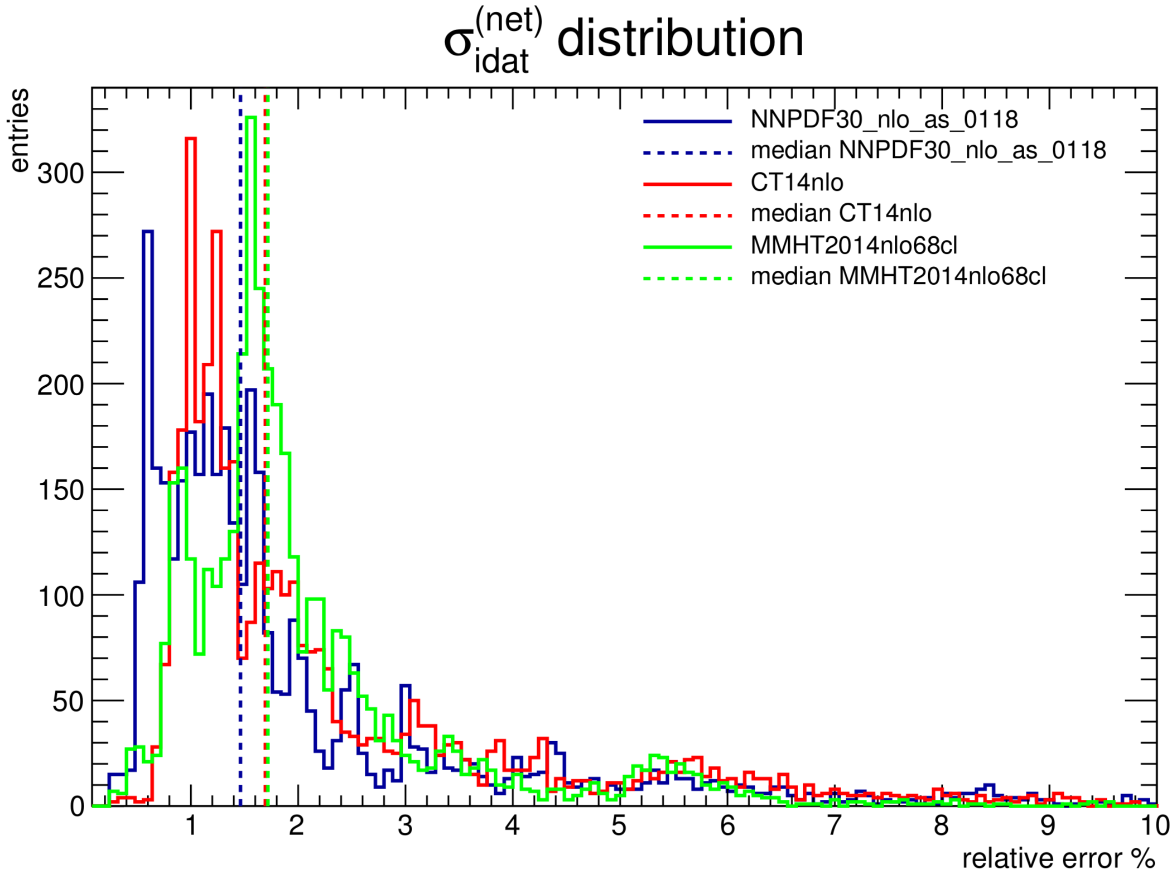


- DIFFERENCES IN UNCERTAINTIES BETWEEN GLOBAL FITS ALWAYS OF ORDER OF SEVERAL PERCENT
- UNCERTAINTY ON UNCERTAINTY SMALLER???

PDF UNCERTAINTIES: HOW MUCH DO THEY VARY?

- COMPUTE PERCENTAGE PDF UNCERTAINTY ON ALL DATA INCLUDED IN GLOBAL FIT
- COMPARE GLOBAL FITS

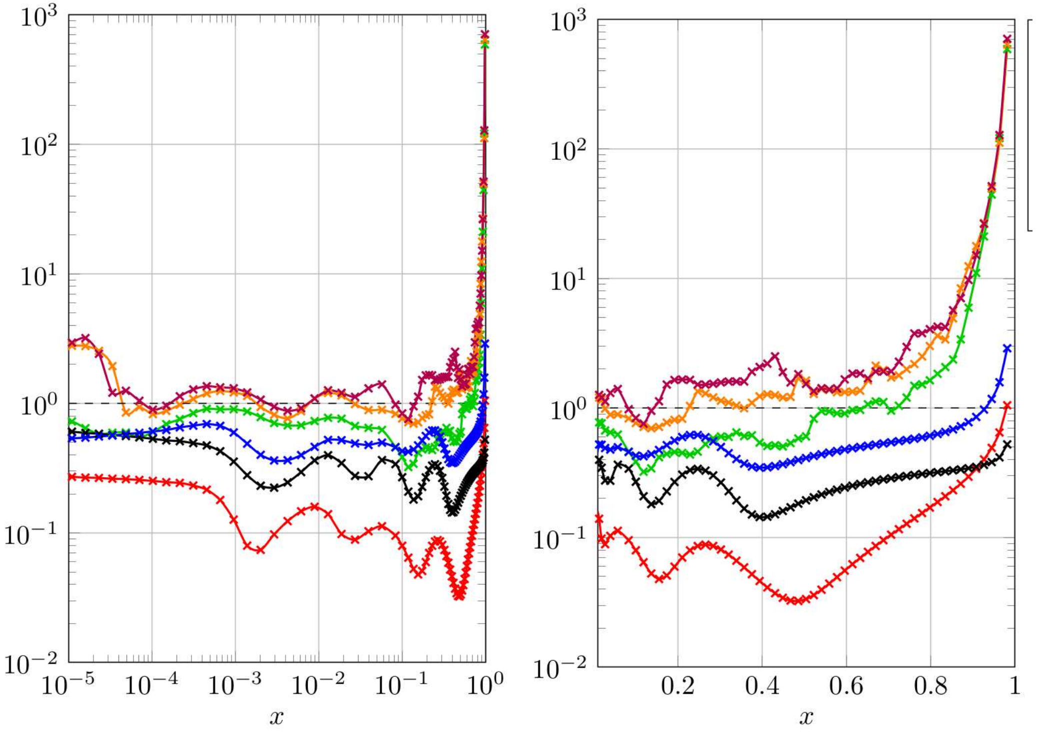
PERCENTAGE PDF UNCERTAINTY ON PREDICTIONS



- MEDIAN SIMILAR
- DISTRIBUTION VERY DIFFERENT!
- NNPDF: SMALLER MODE, BUT FAT TAIL ⇔ GREATER FLEXIBILITY

WHY MORE FLEXIBLE IS BETTER

GLUON PDF UNCERTAINTY NORMALIZED TO MSTW08



(C. Mascaretti, 2016)

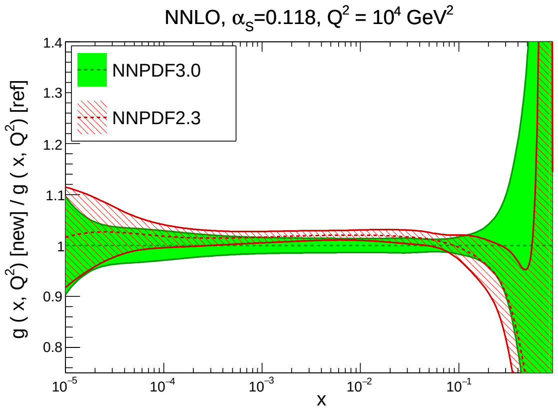
- CLOSURE TEST PERFORMED WITH DATA GENERATED BASED ON MST08 FUNCTIONAL FORM
- REFITTED EITHER WITH NNPDF OR MSTW FUNCTIONAL FORM
- LEVEL 0: VANISHING DATA UNCERTAINTY
  - MSTW-CT: FIT HAS ZERO UNCERTAINTY
  - NNPDF: ABOUT HALF OF TOTAL UNCERTAINTY
- LEVEL 1: NOMINAL DATA UNCERTAINTY, BUT REPLICAS FITTED W/O PSEUDODATA
  - MSTW-CT: FIT HAS SMALL UNCERTAINTY
  - NNPDF: ABOUT 2/3 OF FINAL UNCERTAINTY
- LEVEL 2
  - NNPDF UNCERTAINTY LARGER THAN MSTW-CT
  - NNPDF UNCERTAINTY SIMILAR TO TRUE MSTW

WHY MORE FLEXIBLE IS MORE DANGEROUS
NNPDF: 3.0 VS. 2.3

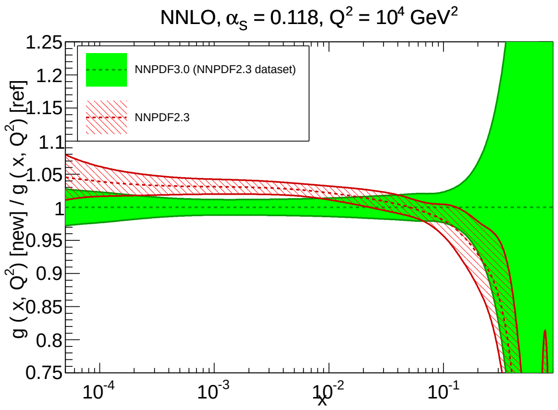
- REPEAT THE 3.0 FIT BUT WITH 2.3 DATASET
- COMPARE WITH 2.3 DATA & METHODOLOGY; 2.3 DATA BUT 3.0 METHODOLOGY; 3.0 DATA & METHODOLOGY

THE GLUON DISTRIBUTION

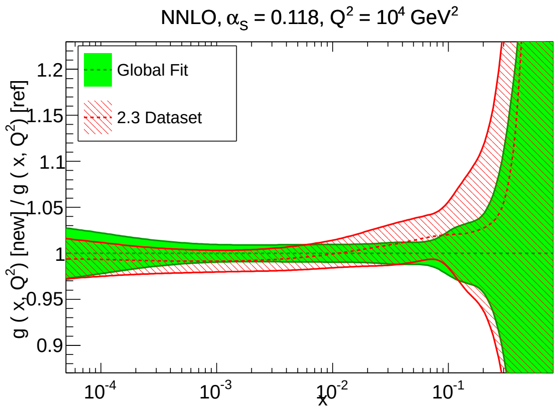
DEFAULT 3.0 VS DEFAULT 2.3



DEFAULT 2.3 VS. 2.3 DATA



DEFAULT 3.0 VS 2.3 DATA



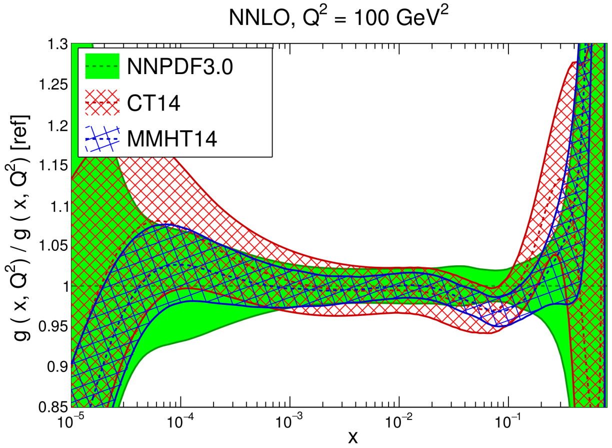
- MAIN METHDOLOGICAL DIFFERENT: MORE EFFICIENT GENETIC MINIMIZATION
- THE METHODOLOGY HAS A DOMINANT EFFECT ON THE GLUON

CAN LHC RUN II DATA DATA HELP?

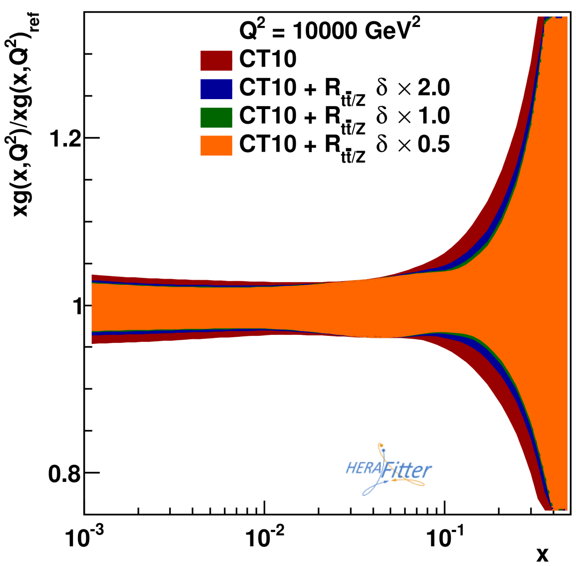
- DATA AT HIGHER CM ENERGY & INFO ON CORRELATION TO LOW ENERGY
  $\to$ EXTENDED KINEMATIC COVERAGE & REDUCED SYSTEMATICS
- REDUCED STAT. UNCERTAINTIES
- PDF4LHC STUDY $\Rightarrow$ MODERATE REDUCTION IN PDF UNCERTAINTY
  EXAMPLE: IMPACT ON GLUON OF $t\bar{t}/Z$ RATIO

THE GLUON

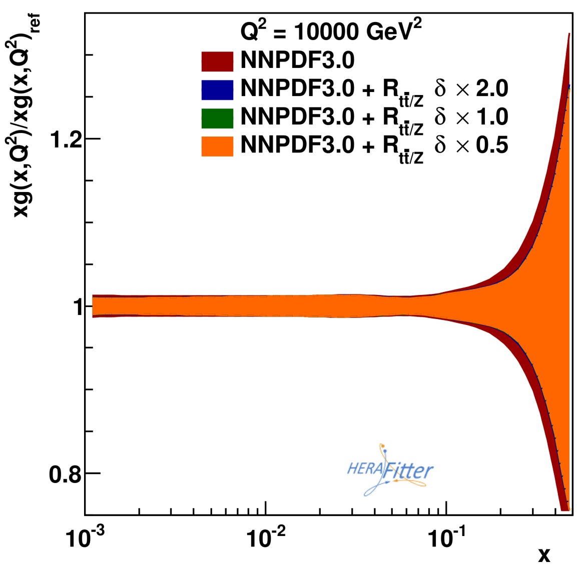
NOW



CTEQ AFTER RUN II



NNPDF AFTER RUN II

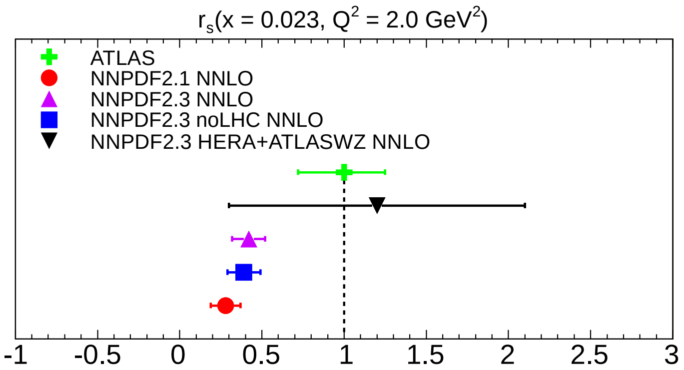


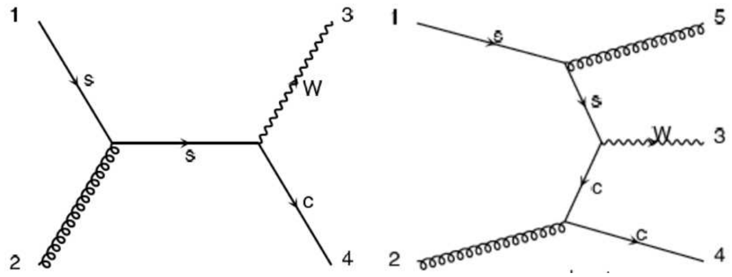
(PDF4LHC: 1507.00556)

NOT VERY MUCH, IF EVERYTHING IS CONSISTENT

HOW LHC DATA MIGHT HELP

STRANGENESS & W PRODUCTION





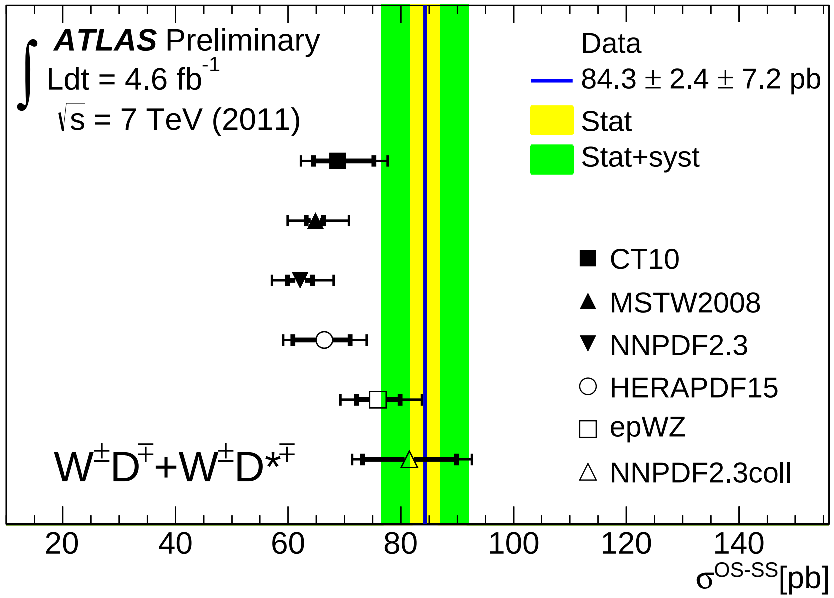
- ATLAS INCLUSIVE W PRODUCTION DATA (2012) SUGGEST LARGE (≈ 1) STRANGE FRACTION $r_s(x, Q^2) = \frac{s(x,Q^2)+\bar{s}(x,Q^2)}{2\bar{d}(x,Q^2)}$
- HOWEVER LARGE UNCERTAINTIES (NNPDF2.3) $\Rightarrow$ CONSISTENT WITH PREVIOUS DET. (FROM NEUTRINO DATA) WITHIN UNCERTAINTIES
- STRANGENESS PROBED DIRECTLY IN $W + c$ PRODUCTION

$W + c$ production at the LHC

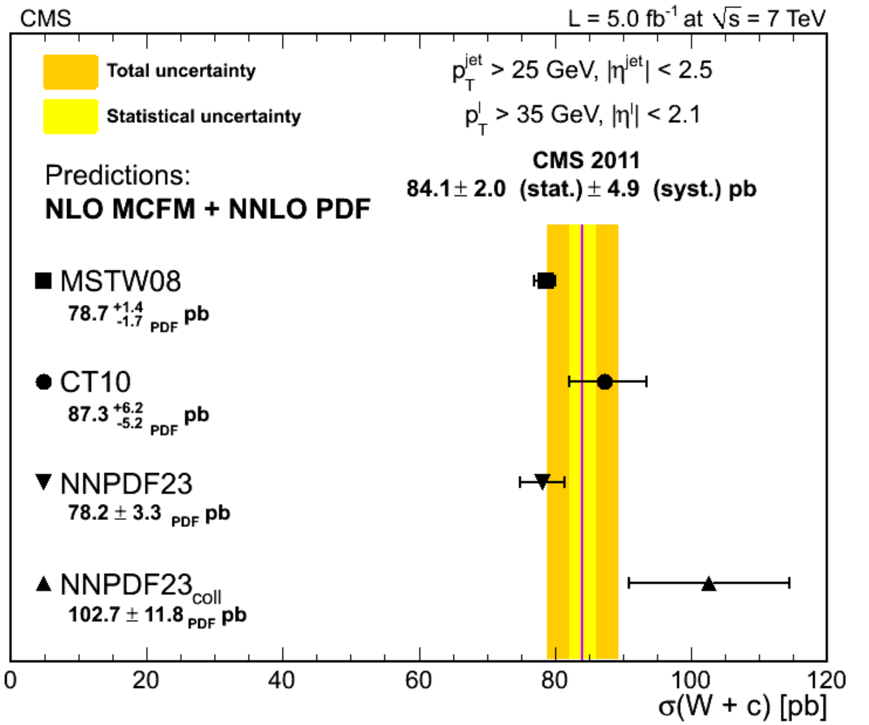
- measurements by ATLAS and CMS consistent within uncertainties
- compare to available default fits & to NNPDF "collider only" fit (no neutrino data) ⇒ tension between DY and neutrino data
- ATLAS central vaue favors DY measurement, CMS central value favors neutrino data

LHC measurements

ATLAS



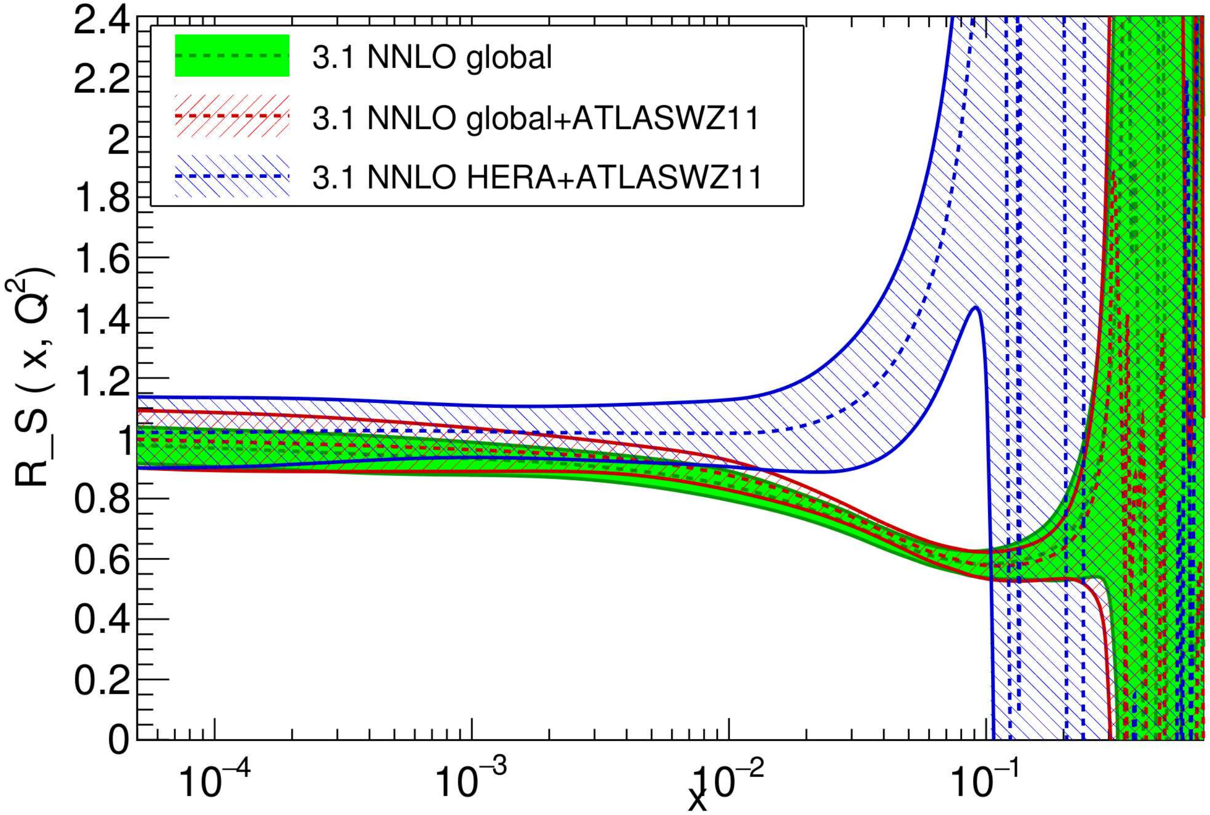
CMS



W PRODUCTION AT THE LHC: 2016

STRANGENESS RATIO

Q=100 GeV



- MORE PRECISE W PRODUCTION DATA PUBLISHED IN 12/2016 TENFOLD INTEGRATED LUMI WR TO 2012 RESULTS
- LEADS TO $R_s(0.023) = 1.19 \pm 0.08$ BASED ON HERA+ATLAS FIT (ATLAS, CONFIRMED BY NNPDF)
- BUT ONLY MINOR IMPACT ON GLOBAL FIT

THE ROLE OF NEW DATA:
LESSONS LEARNT

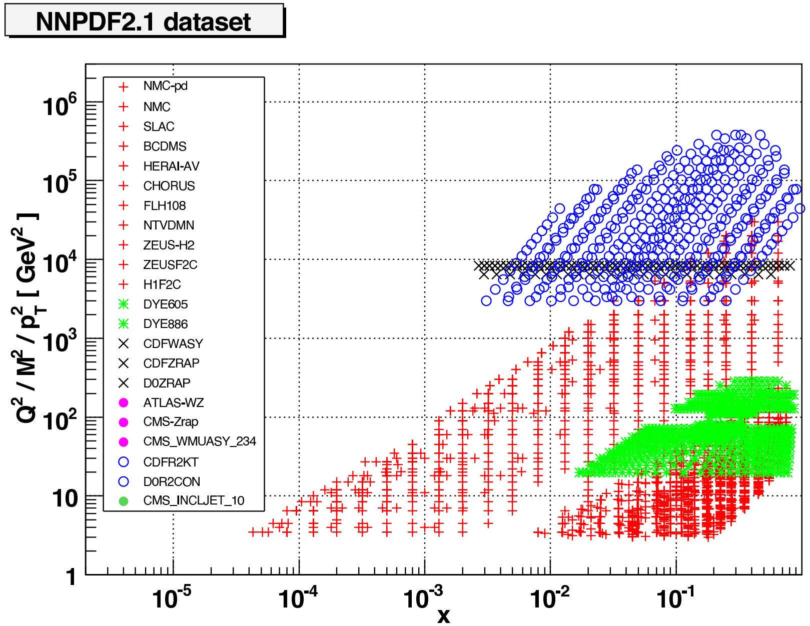
- MISTRUST ASSESSMENTS OF THE IMPACT OF NEW DATA X BASED ON A HERA+X FIT
- NEW DATA ARE UNLIKELY TO REDUCE UNCERTAINTIES IF CONSISTENT WITH EXISTING DATA, BUT MAY HAVE A SIZABLE IMPACT IF NOT CONSISTENT WITH THEM
- NEW DATA CAN HELP IN LEAVING BEHIND DATA BASED ON OBSOLETE PHENOMENOLOGY OR LESS RELIABLE THEORY:
  - OLD FIXED TARGET DIS DATA
  - DIS AND DY DATA ON NUCLEAR (DEUTERIUM) TARGETS
  - NEUTRINO DIS DATA (ALL ON NUCLEAR TARGETS)

LHC DATA CAN HAVE A SIGNFICANT IMPACT ON FLAVOR SEPARATION
VERY IMPORTANT FOR NEW PHYSICS SEARCHES

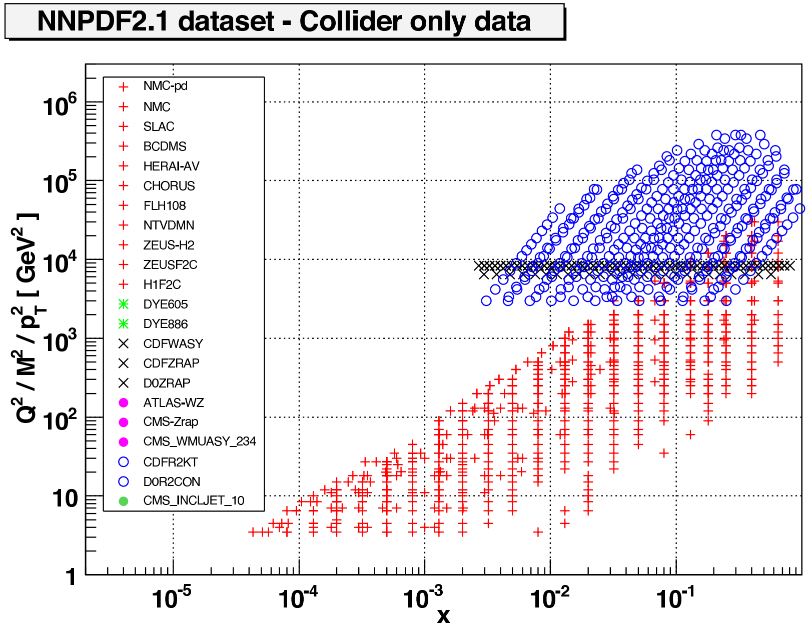
COLLIDER ONLY PDFs
BEFORE THE LHC

LEAVING BEHIND UNSOUND DATA

THE NNPDF2.1 DATASET



NNPDF2.1 - COLLIDER ONLY

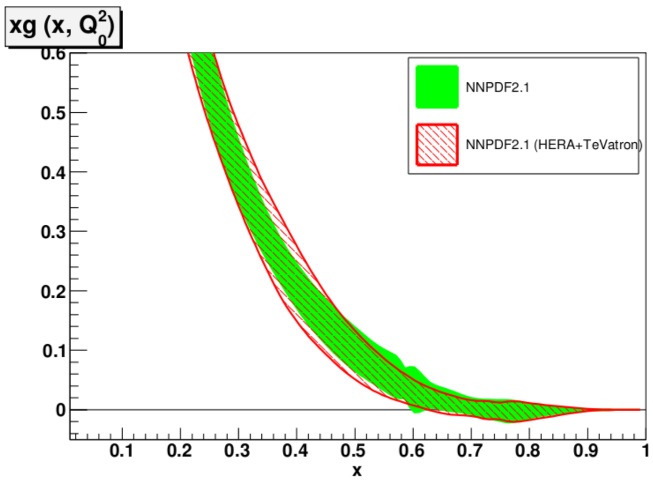


PDFs FROM HERA+TEVATRON DATA?

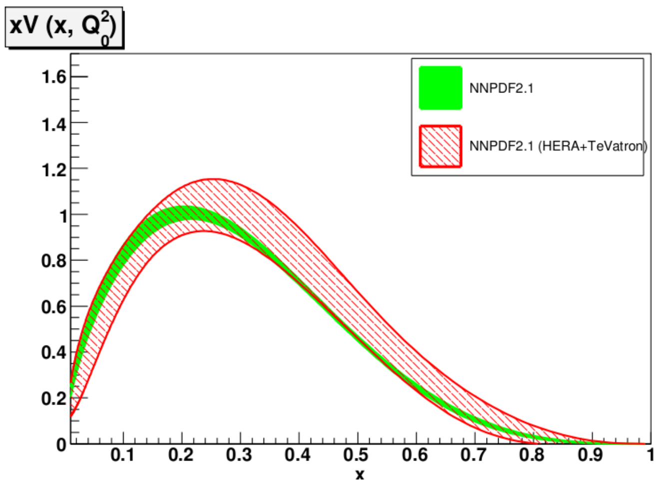
COLLIDER ONLY PDFs
BEFORE THE LHC

NOT A REALISTIC OPTION

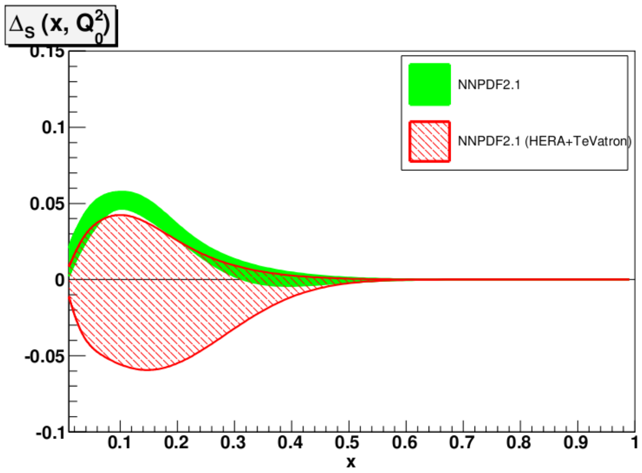
GLUON



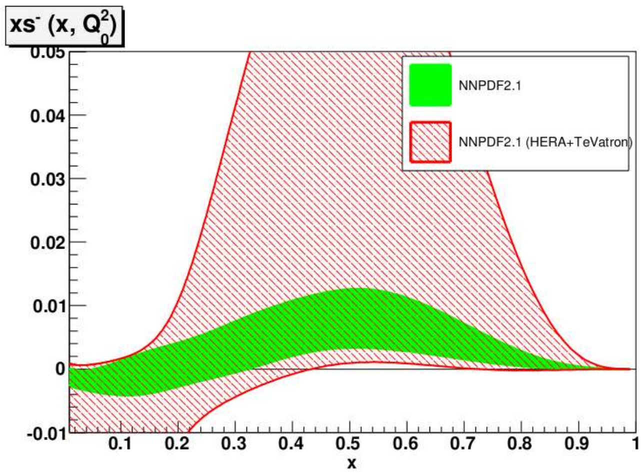
VALENCE



SEA ASYM: $\bar{u} - \bar{d}$



STRANGE: $s - \bar{s}$

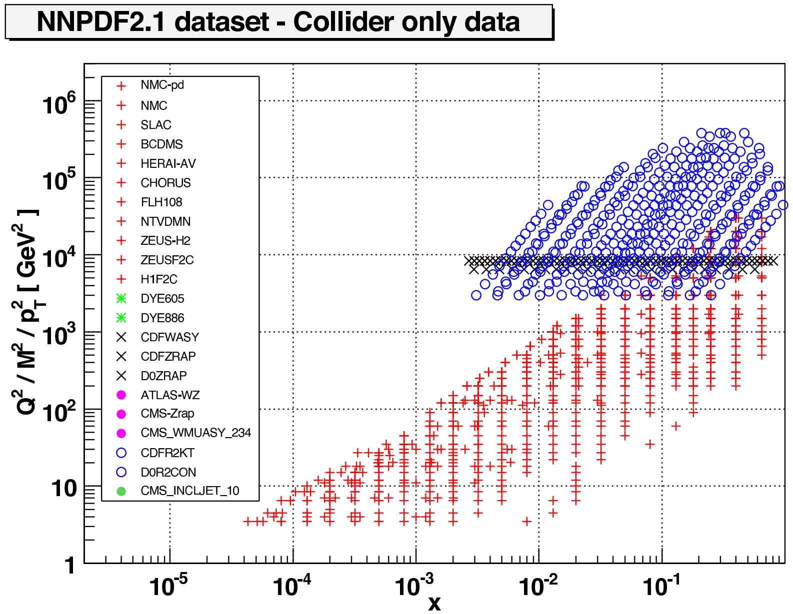


- GOOD ACCURACY FOR GLUON
- GREAT LOSS OF ACCURACY FOR FLAVOR SEPARATION

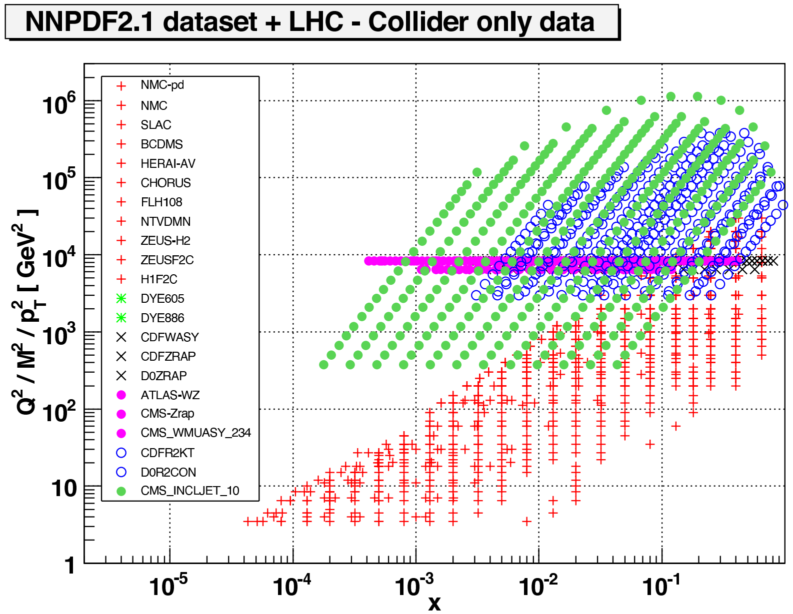
COLLIDER ONLY PDFs
AT THE LHC START

2012 DATA

NNPDF2.1 - COLLIDER ONLY

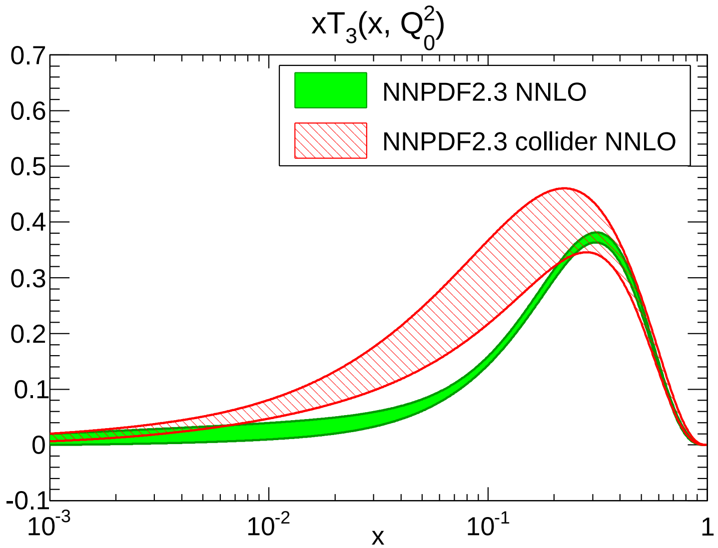


NNPDF2.3 - COLLIDER ONLY

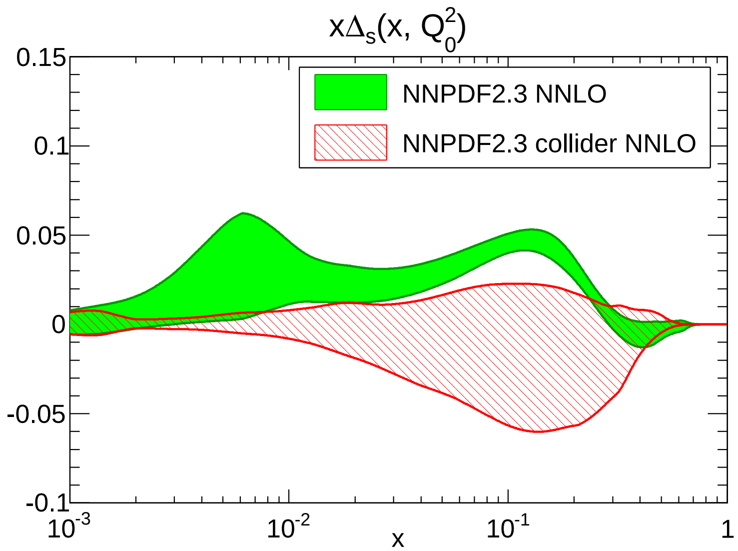


COLLIDER ONLY PDFs
AT THE LHC START

ISOTRIPLET:
$(u+\bar{u})-(d+\bar{d})$



SEA ASYMMETRY:
$\bar{d}-\bar{u}$

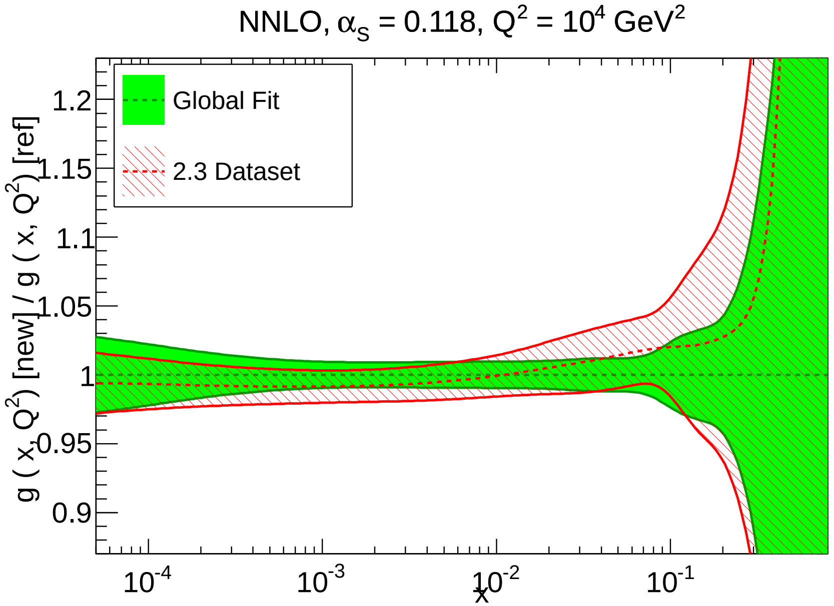


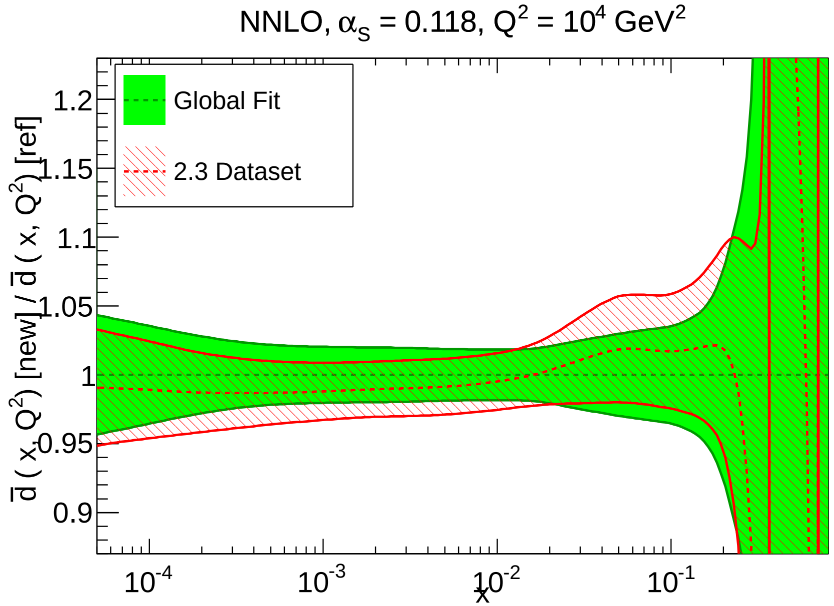
BETTER, BUT STILL NOT VIABLE:

- NNPDF2.3COLLIDER PDFs AFFECTED BY LARGE UNCERTAINTIES
- CRUCIAL MISSING INFORMATION FROM NEUTRINO AND DIS+DY WITH DEUTERON TARGETS
- POOR DETERMINATION OF LIGHT FLAVOR DECOMPOSITION

COLLIDER ONLY PDFs
NOW: CURRENT GLOBAL FITS

NNPDF3.0: DEFAULT VS NNPDF2.3 DATASET





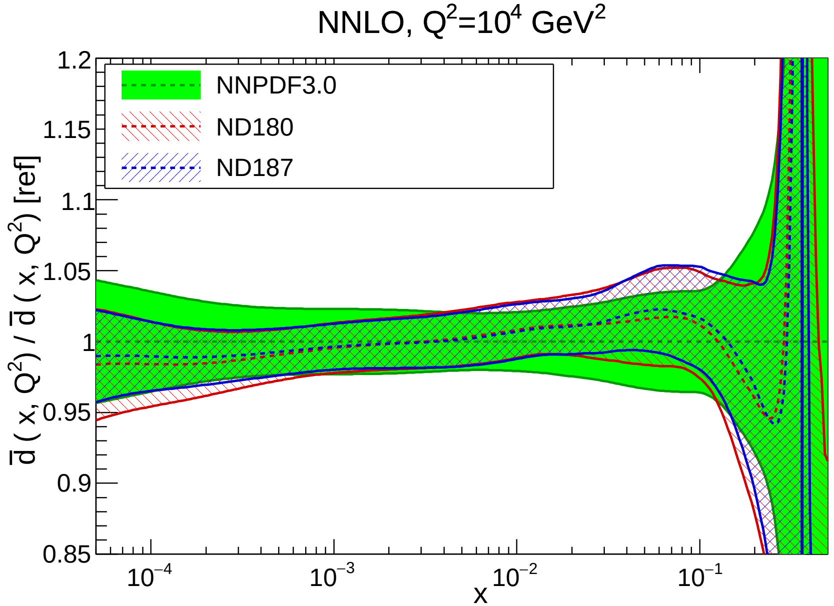
- SOME IMPROVEMENT IN GLUON UNCERTAINTIES DUE TO LHC JET DATA
- NEGLIGIBLE IMPROVEMENT IN FLAVOR DECOMPOSITION FROM EARLY LHC DATA

COLLIDER-ONLY PDFs STILL NOT REALISTIC

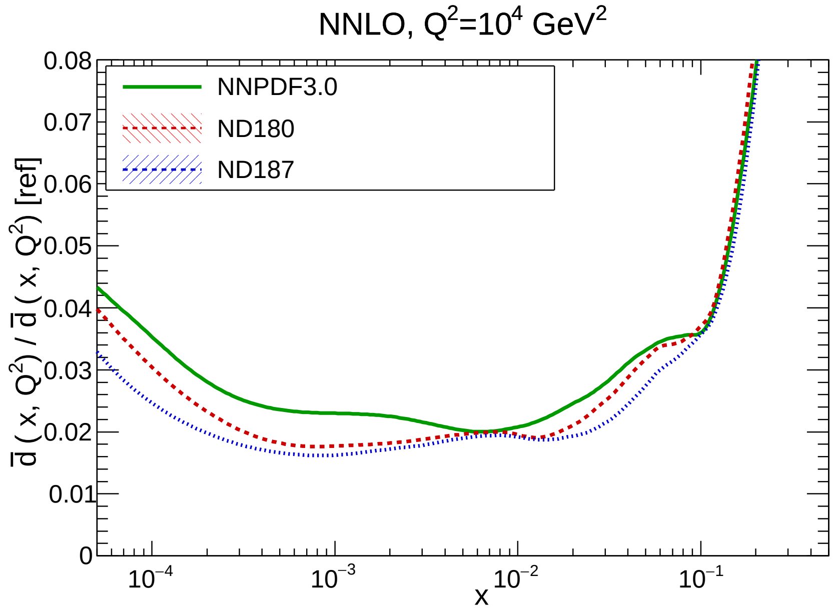
PDFs from the LHC
behind the corner: next generation PDFs

NNPDF3.1 (prelim.) impact of CMS $W\mu$: antidown

PDF



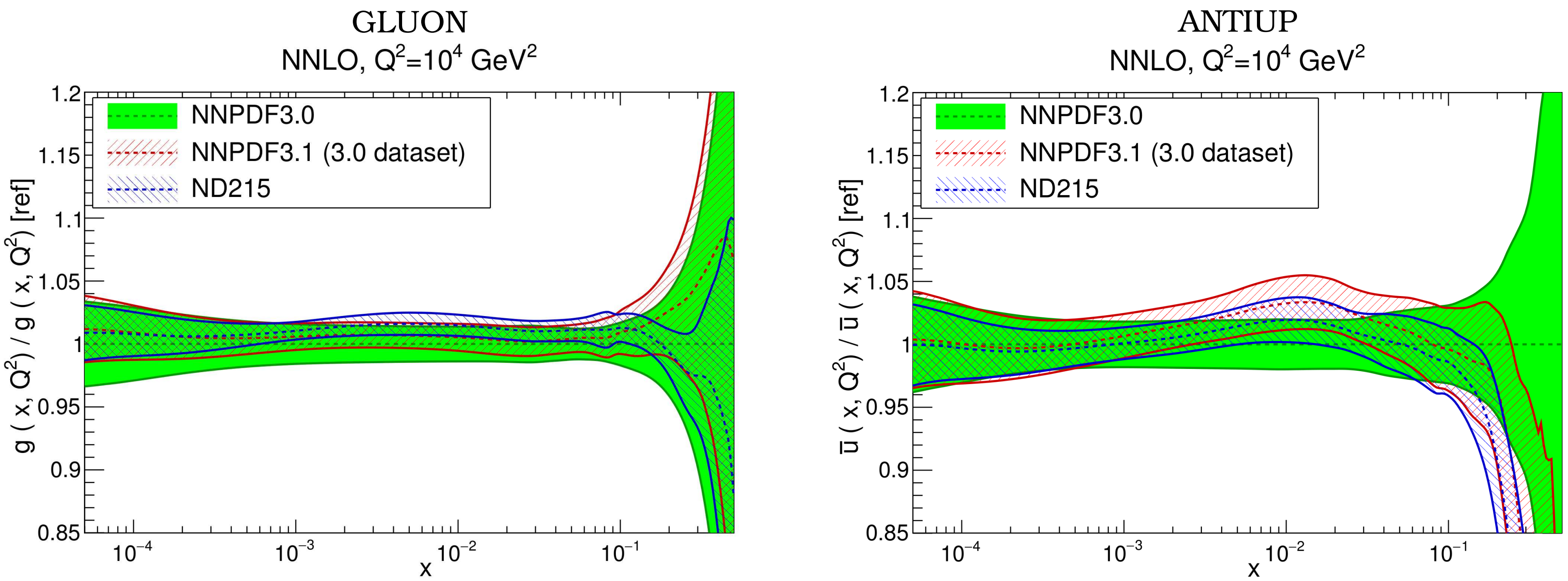
uncertainty



- NNPDF3.1 will include:
  - Tevatron legacy Z rapidity, W asymmetry & jet data
  - ATLAS W, Z rapidity, and total xsect (incl. 13TeV), high and low mass DY, jets
  - CMS W asymmetry, $W + c$ total & ratio, double-differential DY and jets
  - LHCb W and Z rapidity distributions
  - ATLAS and CMS Z p_T distributions
  - ATLAS and CMS top total cross-section & differential rapidity distribution
- no tension ⇒ each dataset small impact plot compares effect of adding CMS W μ 8TeV, overall impact not negligible

collider-only PDFs possible?

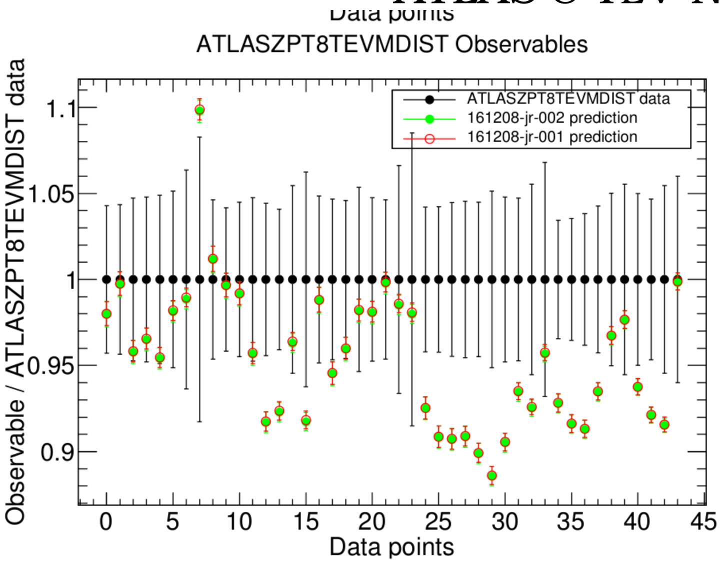
PDFs from the LHC: opportunities
the impact of the Z transverse-momentum distributions

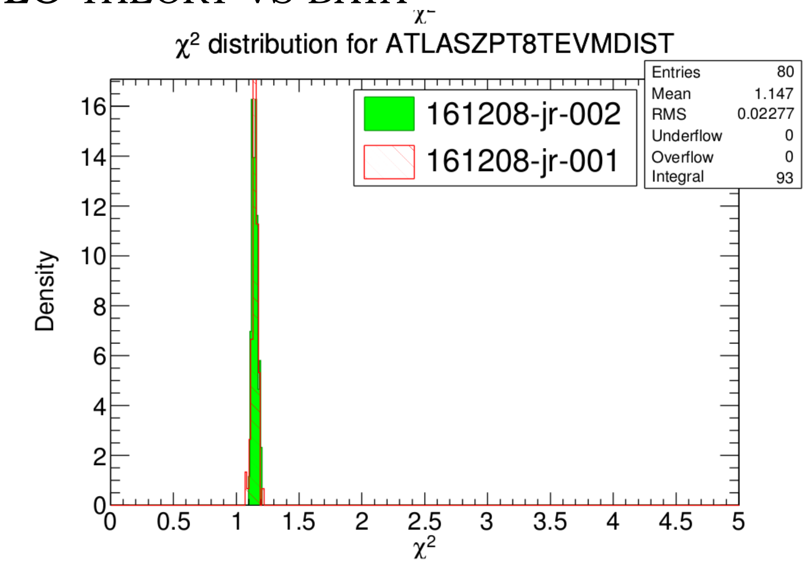


- impact both on gluon and quarks
- improvement in precision & impact on accuracy
- CMS and ATLAS data consistent

PDFs FROM THE LHC: CHALLENGES
THE Z TRANSVERSE-MOMENTUM DISTRIBUTIONS AT THE 1% LEVEL

ATLAS 8 TeV NNLO THEORY VS DATA



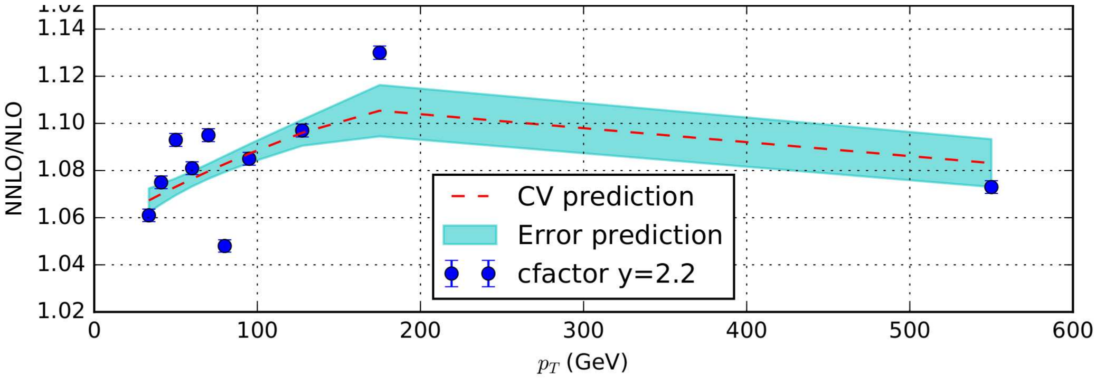


- CORRELATED SYSTEMATICS DOMINATES UNCERTAINTY, OF ORDER $\sim 5\%$
  $\Rightarrow$ GOOD $\chi^2 \sim 1$ per dof DESPITE UNDERSHOOT
- UNCORRELATED UNCERTAINTIES BELOW OR MUCH BELOW 1%
- SIZABLE POINT-TO-POINT FLUCTATIONS IN THEORY PREDICTION

PDFs from the LHC: challenges

the Z transverse-momentum distributions at the 1% level

ATLAS 7 TeV K factors



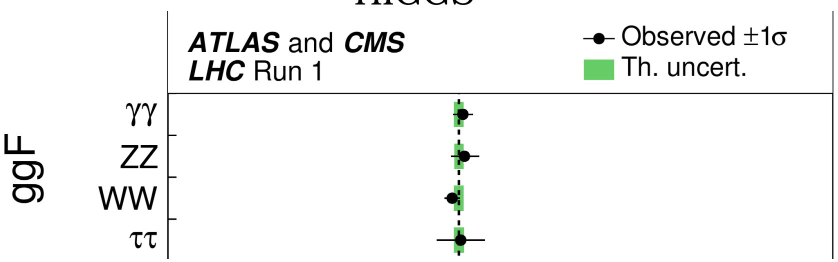
(Boughezal, Liu, Petriello, 2016-2017)

- large NNLO corrections $\sim 10\%$
- uncorrelated uncertainties below or much below 1%
- nominal K-factor uncertainties very small: underestimated?
- extra 1% theory uncertainty estimated based on fluctuations w.r. to interpolation (dashed in plot)

PDFs from the LHC: when does a signal become a standard candle?

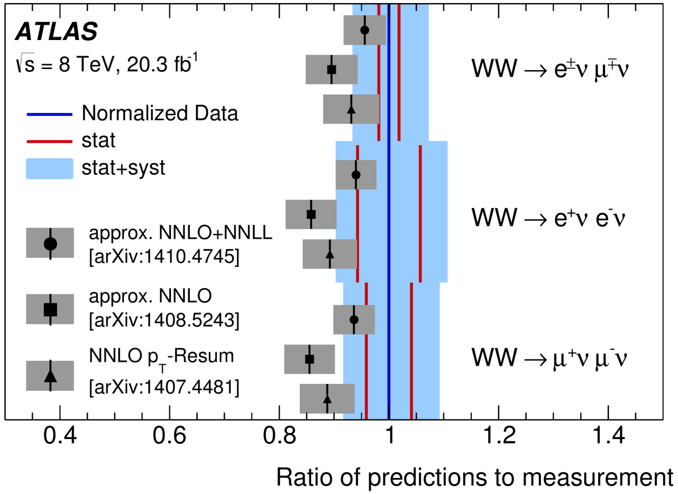
- should we be using double gauge production or Higgs production for PDF determination?
- experimental & theoretical accuracy are or soon to be competitive

HIGGS



ATLAS+CMS (2016)

WW



PROSPECTS & DESIDERATA

- NNLO jets!
- NNLO computations for $W + c$ & $Z + c$ (see below)
- K-factors to 1% accuracy or theory uncertainties
- PDF determination is lagging behind (single top! prompt photon!)

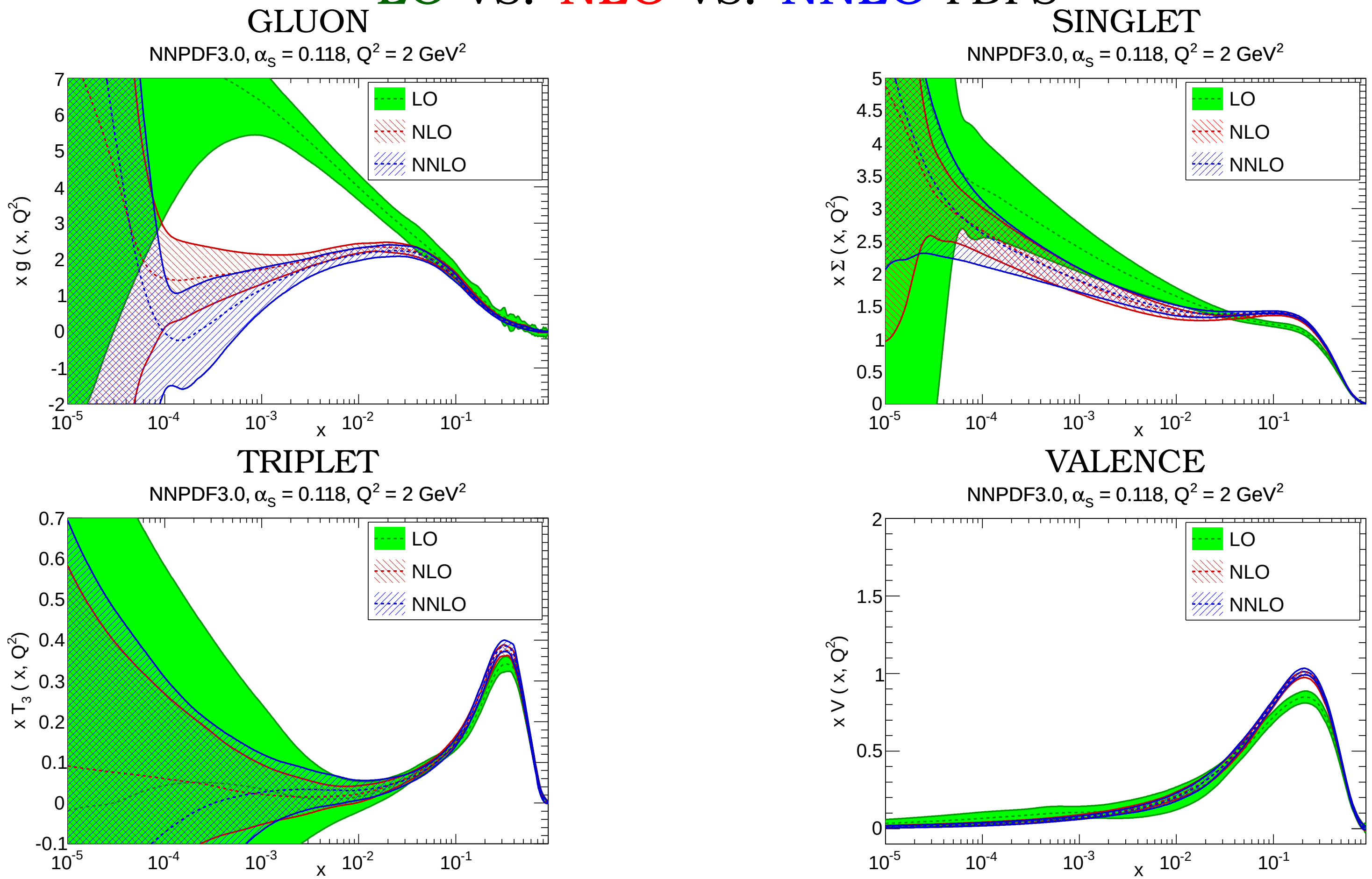
COLLIDER PDFs:

LESSONS LEARNT

- COLLIDER PDFs ARE BEHIND THE CORNER
- CORRELATED SYSTEMATICS DOMINATE THE UNCERTAINTY
- MUST CONTROL POINT-TO-POINT FLUCTUATIONS TO BETTER THAN 1%
- MUST INCLUDE AN ACCURATELY ESTIMATES THEORY UNCERTAINTY

NEW PROBLEMS

PERTURBATIVE STABILITY I
LO vs. NLO vs. NNLO PDFS



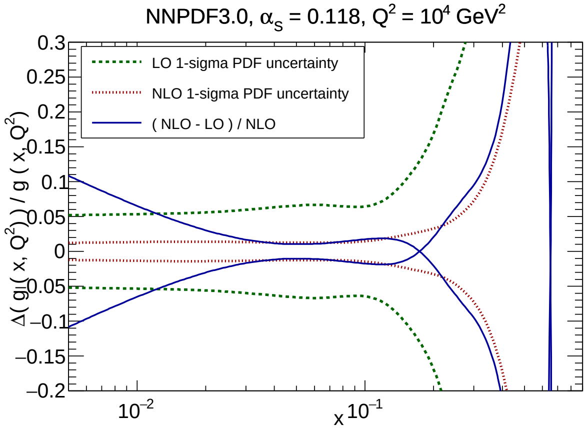
- PERTURBATIVE ACCURACY OF PREDICTION LIMITED BY PERTURBATIVE ACCURACY OF PDF
- $\alpha_s(M_z) \sim 0.1$, $\alpha_s(M_p) \sim 1/2$; $\alpha_s(Q_1^2) = \alpha_s(Q_2^2)(1 + O(\alpha_s^2))$
  $\Rightarrow$ LO: QUALITATIVE; NLO: QUANTITATIVE; NNLO: PRECISION

PERTURBATIVE STABILITY II

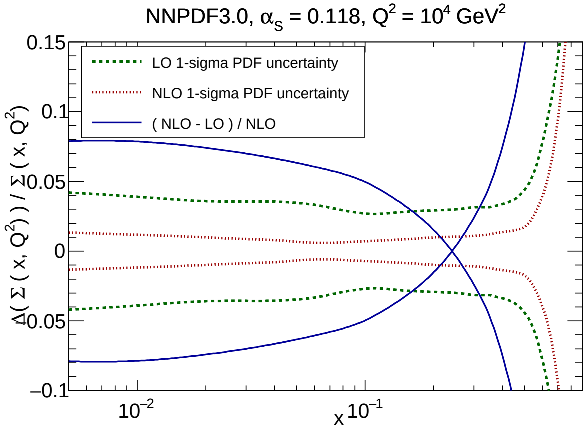
THEORY UNCERTAINTIES VS PDF UNCERTAINTIES

LO VS NLO VS LO-NLO SHIFT

GLUON

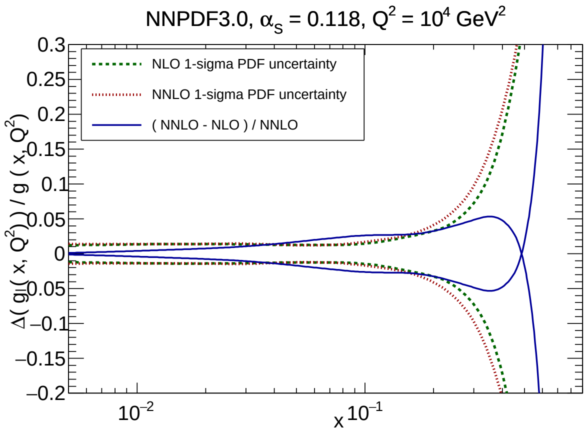


SINGLET

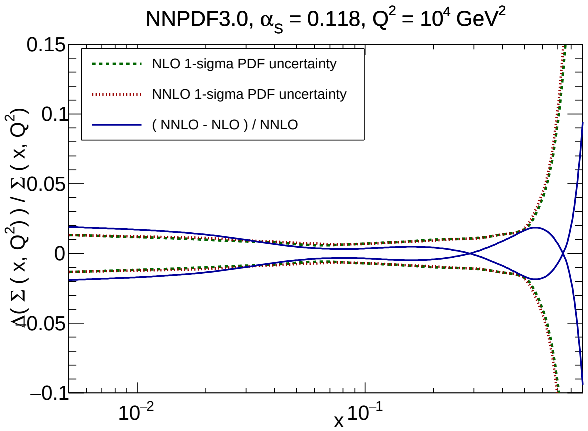


NLO VS NNLO VS NLO-NNLO SHIFT

GLUON



SINGLET



- “DATA” PDF UNCERTAINTY INDEP. OF PERTURBATIVE ORDER
- NLO TH UNCERTAINTY COMPARABLE TO PDF UNCERTAINTY
- TH. UNCERTAINTY (MHOU) VS DATA UNCERTAINTY $\Rightarrow$ LO: DOMINANT; NLO, COMPARABLE; NNLO: SUBDDOMINANT

THEORETICAL UNCERTAINTIES ON PDFs:

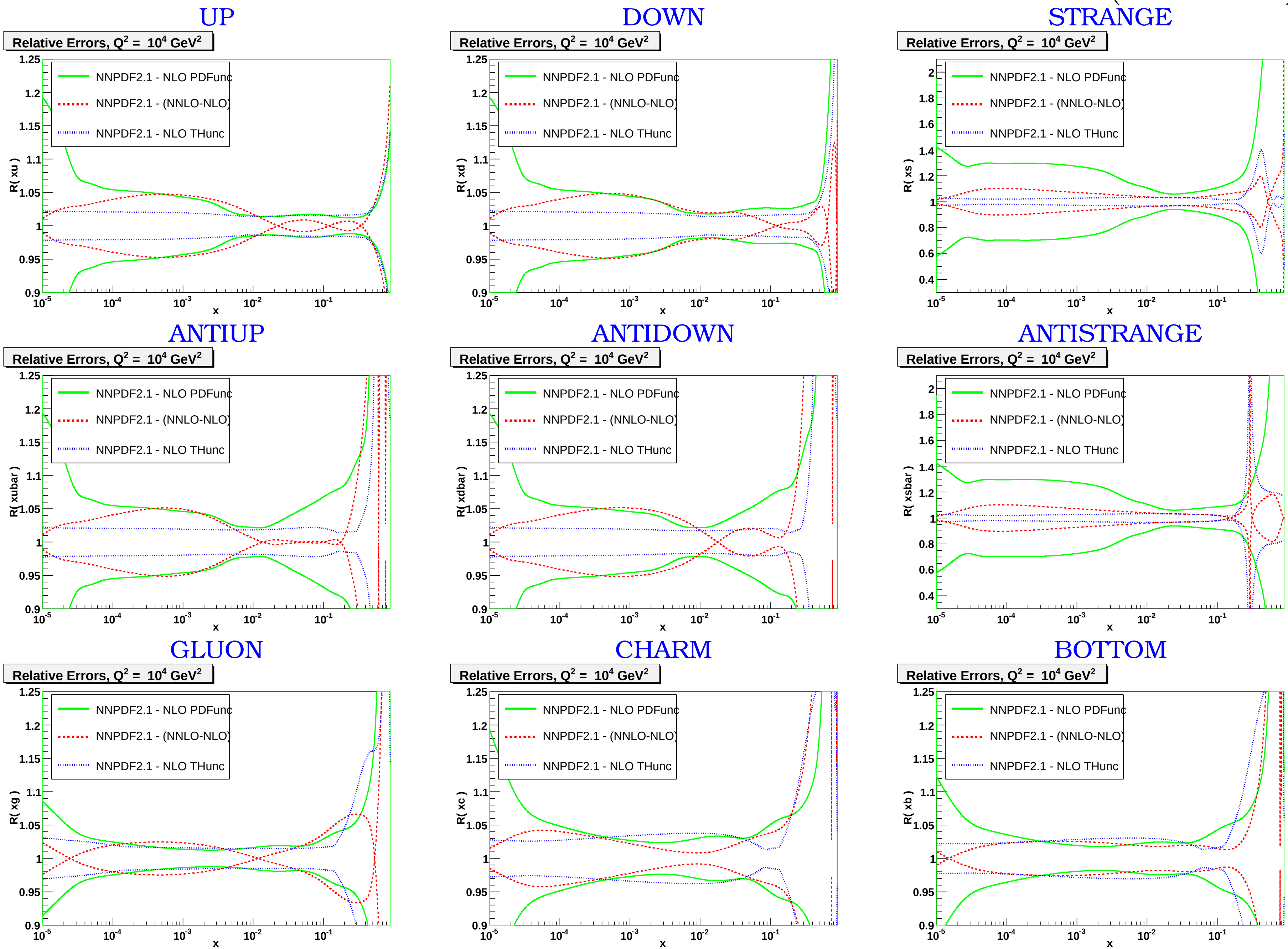
- PDFs ARE DETERMINED BY COMPARING TO DATA THEORY AT SOME FINITE ORDER
- AFFECTED BY THEORETICAL UNCERTAINTY JUST LIKE HARD CROSS-SECTIONS
- NOT INCLUDED IN CURRENT "PDF UNCERTAINTY"
  (ACCOUNTS ONLY DATA & METHODOLOGY)

CAN WE ESTIMATE THEM?

- SCALE VARIATION DIFFICULT:
  CORRELATED BETWEEN PROCESSES? HOW DOES IT CORRELATE WITH PROCESSES IN WHICH PDFs ARE USED?
- AT NLO: WE KNOW THE SHIFT TO NNLO
- AT NNLO: LOOK AT THE BEHAVIOUR OF THE PERTURBATIVE EXPANSION
  (CACCIARI-HOUDEAU)

THEORETICAL UNCERTAINTIES

NLO PDF unc. vs NLO-NNLO shift vs NLO Cacciari-Houdeau (NNPDF2.1)



Cacciari-Houdeau promising?

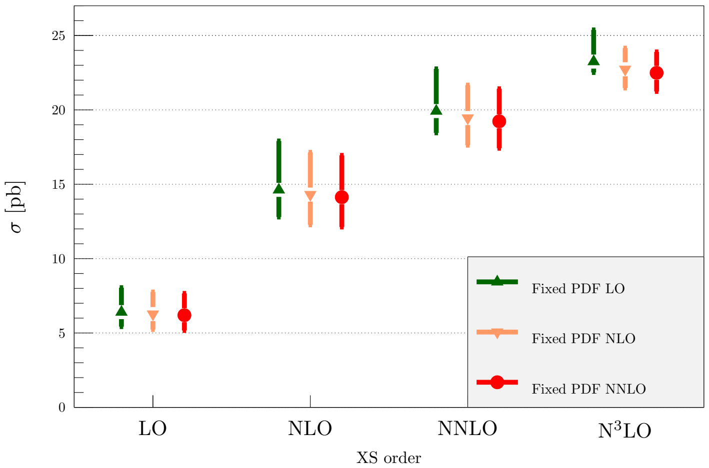
N^{3}LO PDFs:

- NEEDED AT THE 1% ACCURACY LEVEL
- IMPACT OF N^3LO DEPENDS ON PROCESS:
  - HIGGS GLUON FUSION: PERTURBATIVE DEP. OF PDF NEGLIGIBLE IN COMPARISON TO MATRIX ELEMENT ⇒ N^{3}LO NOT NEEDED; UNCERTAINTY ESTIMATED 1% BY ANASTASIOU ET AL.
  - TOP: PERTURBATIVE DEP. OF PDF SMALLER, BUT NOT NEGLIGIBLE IN COMPARISON TO MATRIX ELEMENT, ANTICORRELATED TO IT
    ⇒ N^{3}LO NECESSARY

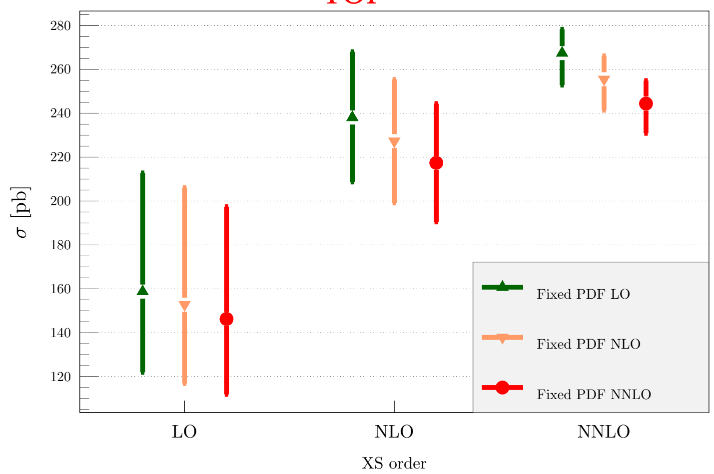
SCALE UNCERTAINTY & DEP. ON PERTURBATIVE ORDER

,

HIGGS



TOP



(s.f., Isgrò, Vita, 2014)

WHEN WILL WE HAVE THEM?

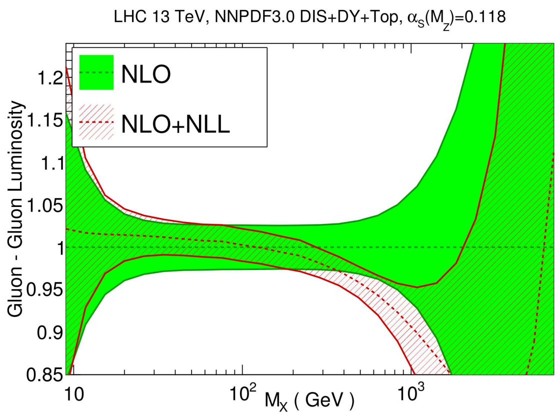
- N^{3}LO DIS COEFFICIENT FUNCTIONS KNOWN
- BOTTLENECK: N^3LO ANOMALOUS DIMENSIONS
- ANOMALOUS DIMENSIONS: LO: 1974; NLO: 1981; NNLO: 2004; N^{3}LO: 2030?

RESUMMED PDFs

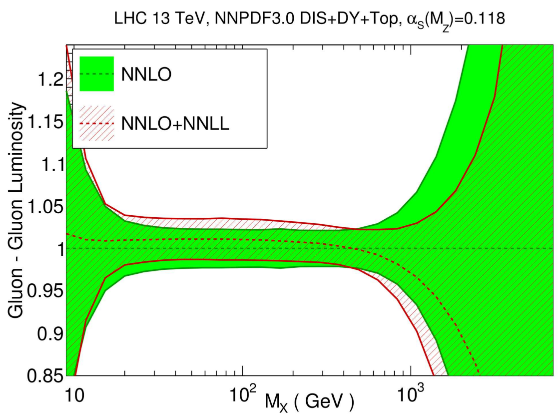
- SO FAR NO RESUMMED PDF SETS AVAILABLE
- PRELIMINARY STUDY: IF THRESHOLD RESUMMATION INCLUDED IN FIT (DIS, DY, TOP DATA), EFFECTS NOT NEGLIGIGLE AT NLLO, LARGE x, MORE MODERATE AT NNLO
- EFFECT ON PDFs COMPARABLE TO EFFECT ON MATRIX ELEMENT, ANTICORRELATED TO IT
- RELEVANT FOR NEW PHYSICS SEARCHES

(Bonvini et al., 2015)

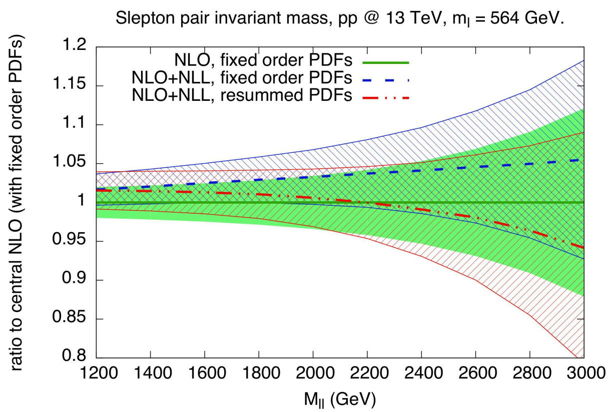
GLUON: NLO vs NLL



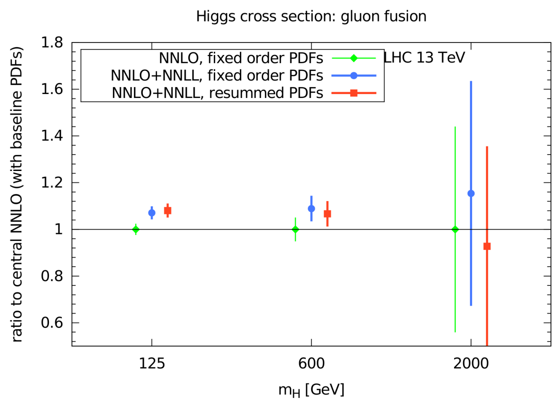
GLUON: NNLO vs NNLL



SLEPTON PAIR PRODUCTION



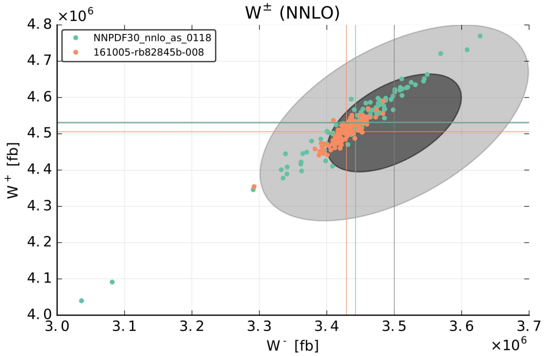
HIGGS IN GLUON FUSION vs m_H



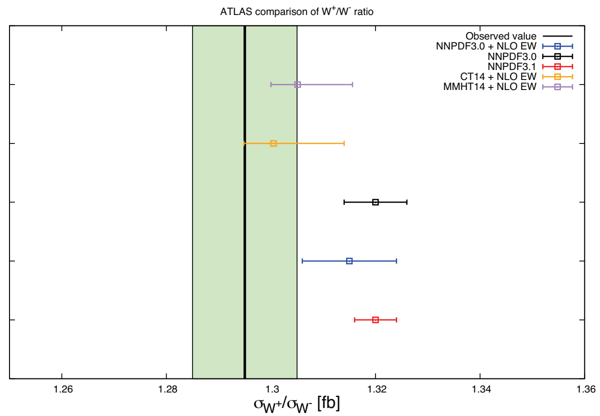
ELECTROWEAK CORRECTIONS

W PRODUCTION RATIOS LHC 13

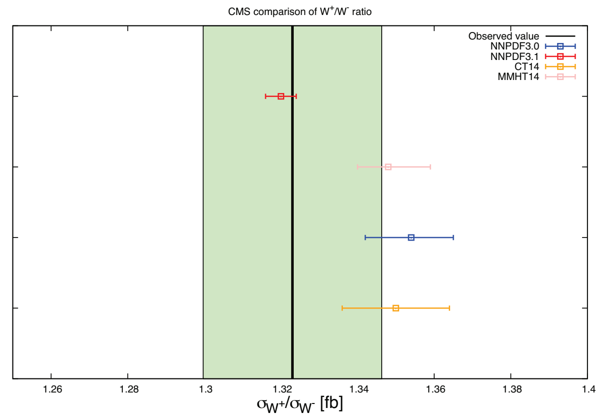
NNPDF3.1 REPLICAS



ATLAS



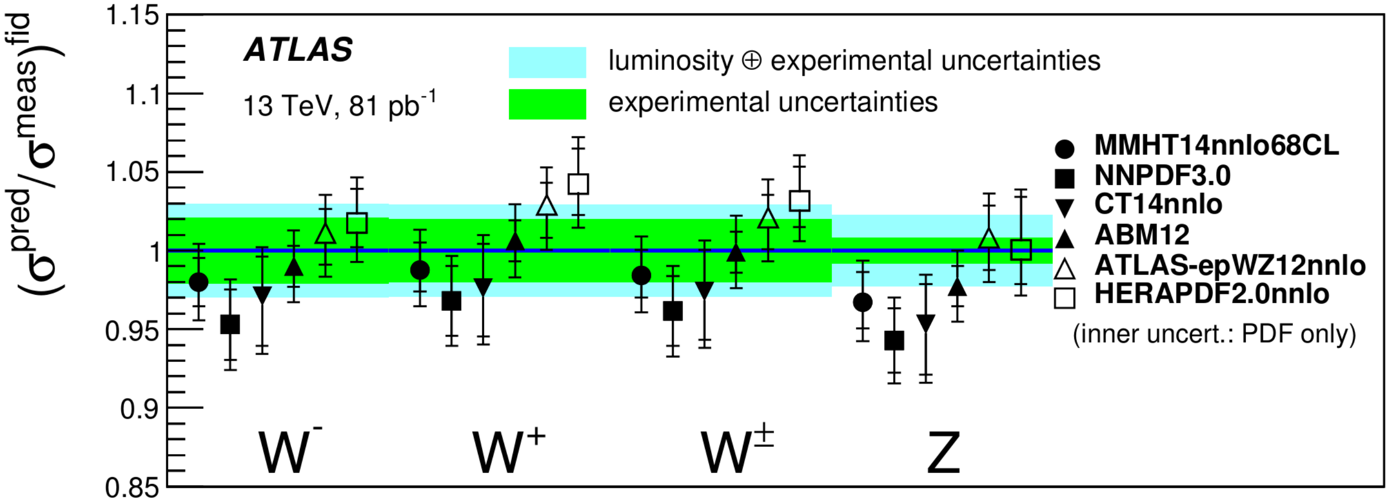
CMS



- GLOBAL PDF SETS OVERSHOOT THE W^+/W^- (AND TO A LESSER EXTENT THE W/Z) TOTAL CROSS-SECTION RATIOS
- EFFECT STATISTICALLY MARGINAL BUT SUGGESTS POSSIBLE SYSTEMATIC PROBLEM
- EW CORRECTIONS ALLEVIATE THE DISCREPANCY (REMOVE IT FOR THE Z CROSS-SECTION)
- EW PS & FSR AT PERCENT OR SUB-PERCENT LEVEL: THEORY CONSISTENTLY TREATED?

HEAVY QUARKS

W AND Z PRODUCTION

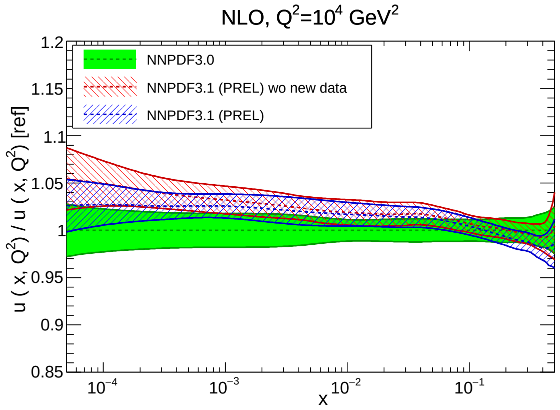


- GLOBAL PDF SETS UNDERSHOOT THE Z (AND TO A LESSER EXTENT $W^{\pm}$) CROSS-SECTION
- EFFECT STATISTICALLY MARGINAL BUT SUGGESTS POSSIBLE SYSTEMATIC PROBLEM

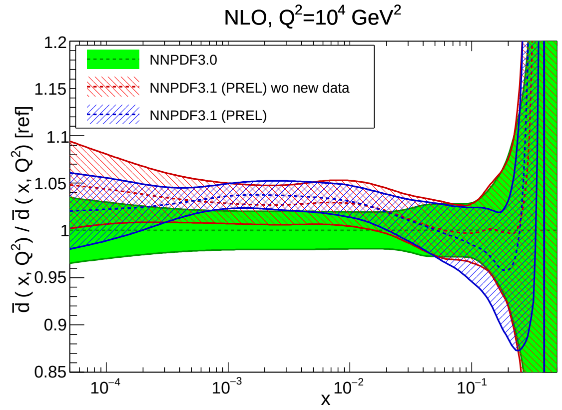
THE IMPACT OF THE CHARM MASS

PERTURBATIVE CHARM

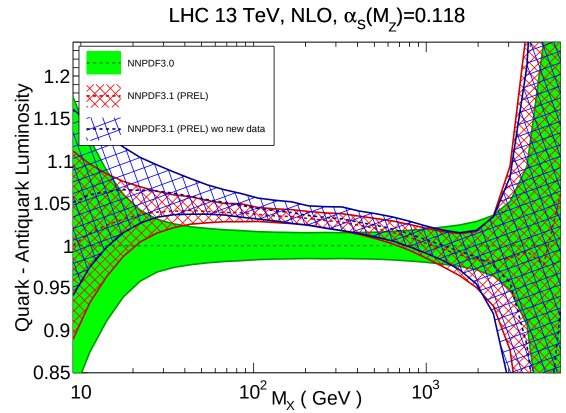
UP



ANTIDOWN



$Q\bar{q}$ LUMI



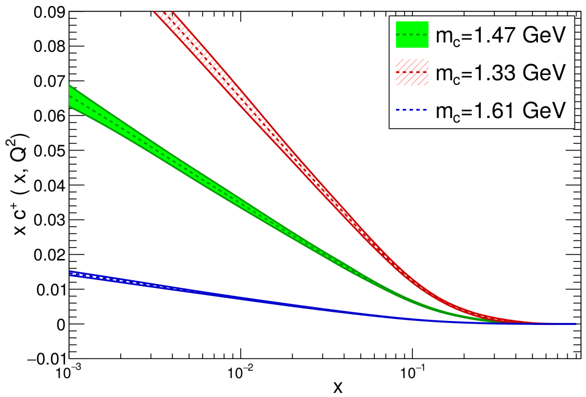
- DIFFERENCE BETWEEN 3.0 AND 3.1 W/. 3.0 DATASET: CHANGE OF THE VALUE OF THE CHARM MASS (POLE) FROM 1.275 GeV (NNPDF3.0) TO 1.47 (NNPDF3.1: PDG)
  $\Rightarrow$ SHIFT LIGHT QUARK PDFs BY ONE SIGMA I.E. $\sim 3\%$
- CURRENT GLOBAL PDF SETS HAVE $m_c = 1.275$ GeV (NNPDF3.0) 1.3 GeV (CT14) $m_c = 1.4$ GeV (MMHT), MOSTLY ON HISTORICAL GROUND
- NO CURRENT PDF SET HAS A MASS UNCERTAINTY (PAST SETS AVAILABLE FOR DIFFERENT m_c VALUES)
- WHY SUCH A BIG EFFECT?

FITTED CHARM vs PERTURBATIVE CHARM

LOW SCALE

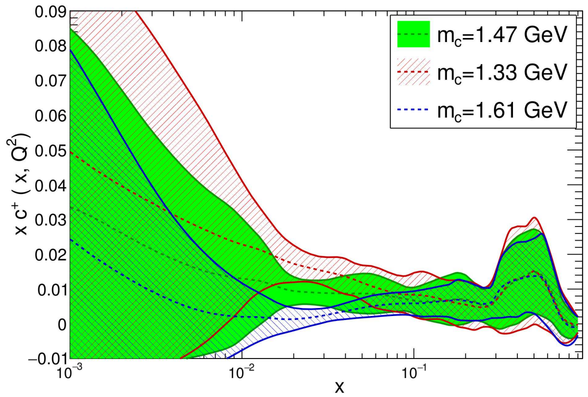
PERTURBATIVE

NNPDF3 NLO Perturbative Charm, Q=1.65 GeV



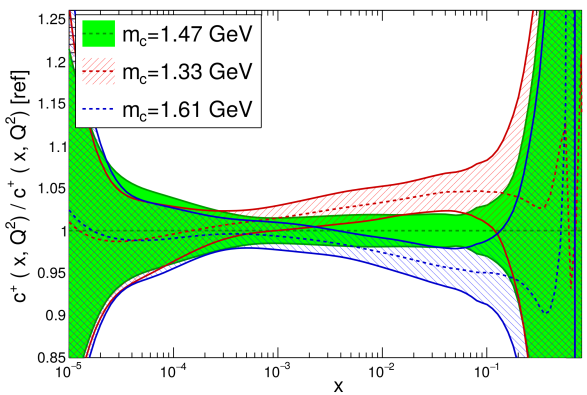
FITTED

NNPDF3 NLO Fitted Charm, Q=1.65 GeV

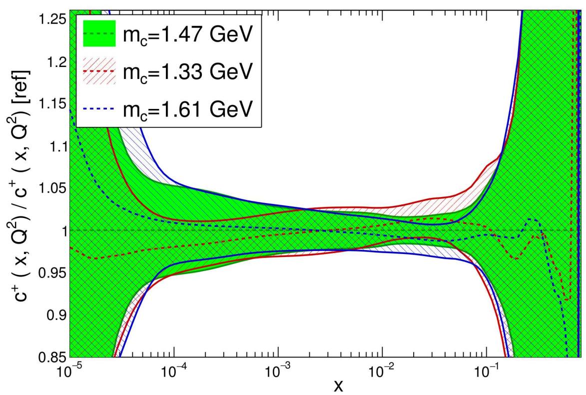


HIGH SCALE

NNPDF3 NLO Perturbative Charm, Q=100 GeV



NNPDF3 NLO Fitted Charm, Q=100 GeV



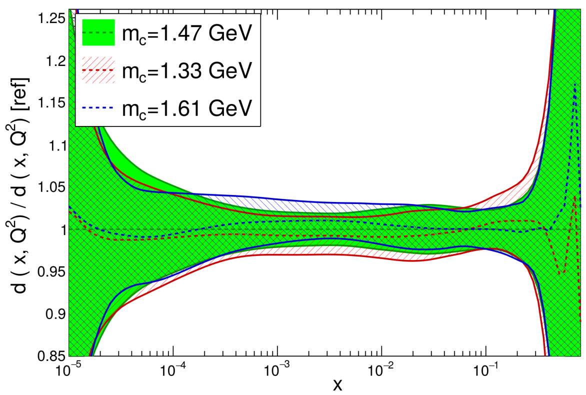
- PERTURBATIVE CHARM: DEPENDS SIGNIFICANTLY ON THE MASS WHICH SETS THE PHYSICAL THRESHOLD; DEPENDENCE SEEN BOTH AT LOW AND HIGH SCALE;
- FITTED CHARM: QUITE STABLE AT ALL SCALES

THE LIGHT QUARKS

DOWN

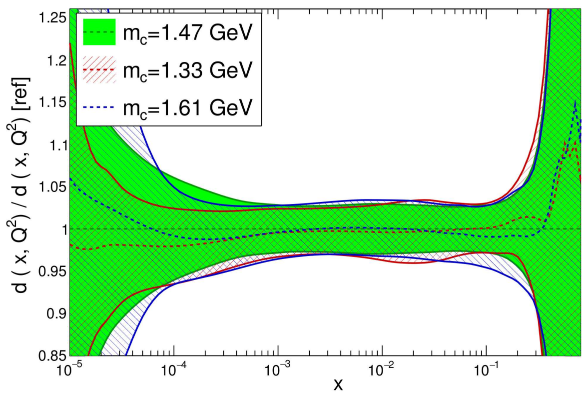
PERTURBATIVE

NNPDF3 NLO Perturbative Charm, Q=100 GeV



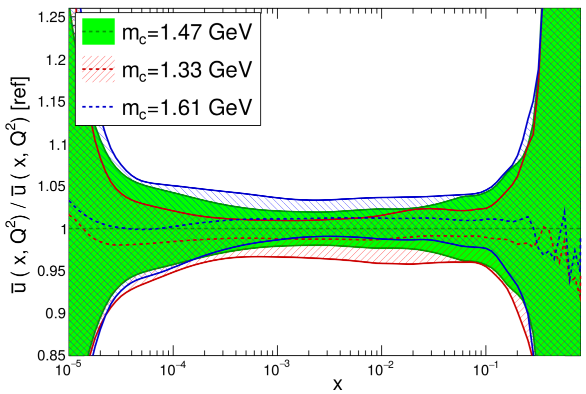
FITTED

NNPDF3 NLO Fitted Charm, Q=100 GeV

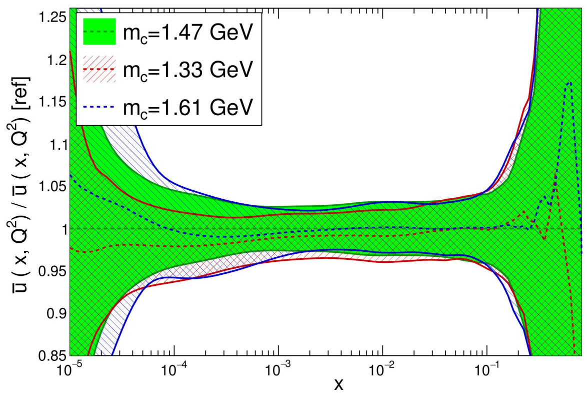


ANTIUP

NNPDF3 NLO Perturbative Charm, Q=100 GeV



NNPDF3 NLO Fitted Charm, Q=100 GeV

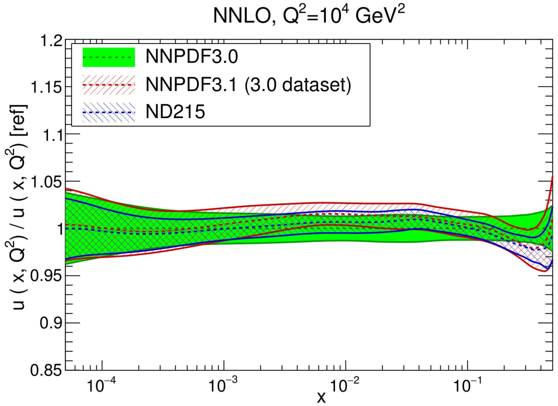


- PERTURBATIVE CHARM: LIGHT QUARKS DEPEND SIGNIFICANTLY ON THE MASS WHICH SETS THE PHYSICAL CHARM THRESHOLD; DEPENDENCE SEEN BOTH AT LOW AND HIGH SCALE;
- FITTED CHARM: LIGHT QUARKS INDEPENDENT OF CHARM MASS
- m_c UNCERTAINTY REABSORBED IN PDF UNCERTAINTY

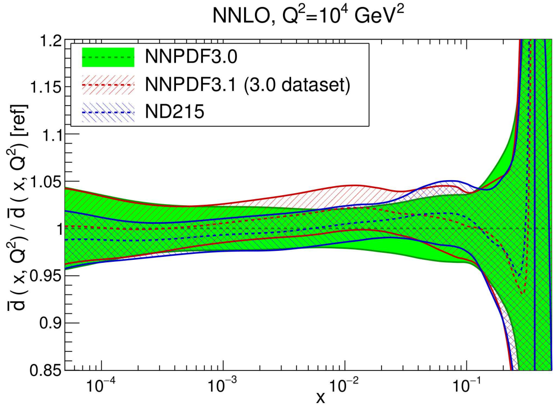
THE IMPACT OF THE CHARM MASS

FITTED CHARM

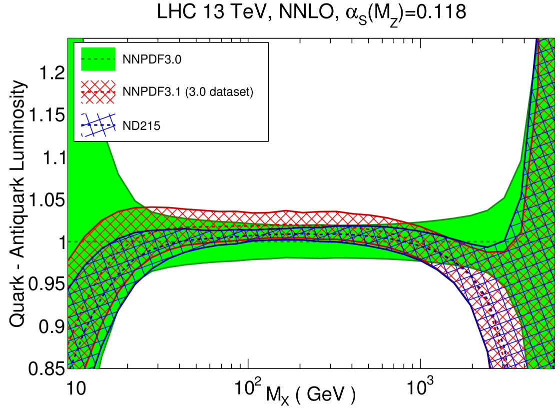
UP



ANTIDOWN



$Q\bar{q}$ LUMI



- DEPENDENCE ON MASS ALMOST DISAPPEARED!
- FIT MUCH CLOSER TO PREVIOUS NNPDF3.0

THEORETICAL UNCERTAINTIES:

LESSONS LEARNT

- THEORETICAL UNCERTAINTIES ARE NO LONGER SUBDOMINANT
- PDF UNCERTAINTIES SHOULD INCLUDE SOME ESTIMATE OF MHOU
- THE CHOICE OF HEAVY QUARK MASS VALUES MAY HAVE SIGNIFICANT EFFECTS
- HEAVY QUARK MASS VARIATION SHOULD BE PERFORMED
- FITTED CHARM PDF LEADS TO INCLUSION OF MASS UNCERTAINTY INTO PDF UNCERTAINTY

NO EFFECT THAT REQUIRES MORE THAN 10% ACCURACY IN MEASUREMENT IS WORTH INVESTIGATING

Walther Nernst

~~NO EFFECT THAT REQUIRES MORE THAN 10% ACCURACY IN MEASUREMENT IS WORTH INVESTIGATING~~

Walther Nernst

ACCURACY OF OBSERVATION IS THE EQUIVALENT OF ACCURACY OF THINKING

Wallace Stevens