PARTON DISTRIBUTIONS

YESTERDAY, TODAY, AND TOMORROW

STEFANO FORTE

UNIVERSITÀ DI MILANO & INFN

UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

Istituto Nazionale di Fisica Nucleare

RWTH AACHEN

NOVEMBER 8, 2018

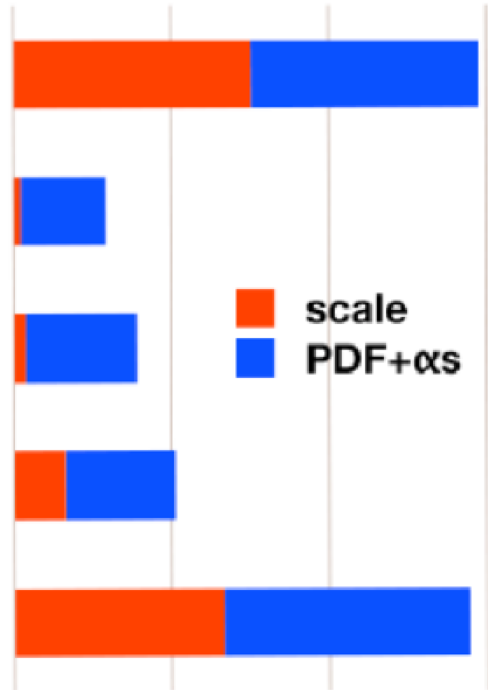
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 740006

PROLOGUE

PAST (DISTANT)

HIGGS PRODUCTION

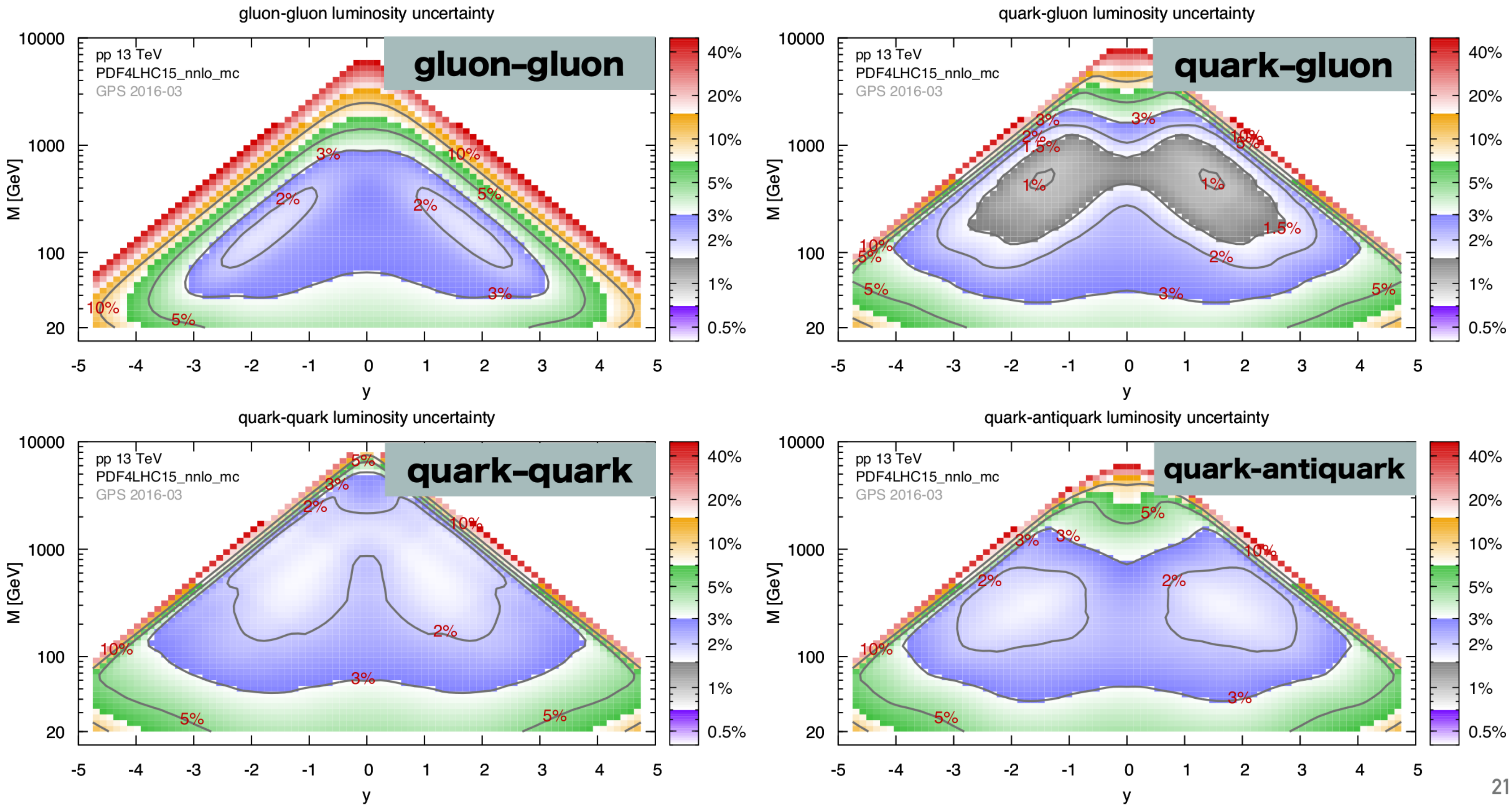
		σ (8 TeV)	uncertainty
NNLL QCD +NLO EW	gg→H	19.5 pb	14.7%
NNLO QCD +NLO EW	VBF	1.56 pb	2.9%
	WH	0.70 pb	3.9%
	ZH	0.39 pb	5.1%
NLO QCD	ttH	0.13 pb	14.4%



(J. Campbell, 2012)

PDF UNCERTAINTY EITHER DOMINANT, OR VERY LARGE, OR BOTH
TYPICAL PDF UNCERTAINTY $\sim 5 - 10\%$

PAST (NOT SO LONG AGO) THE PDF4LHC SET
LUMINOSITY UNCERTAINTIES VS RAPIDTY & MASS



G.P. Salam, 2016

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

TOWARDS 1% PDF UNCERTAINTIES?

PRESENT: PDF CHALLENGES

AN EXAMPLE: HIGGS IN GLUON FUSION

~~YR4~~ VS. NOW

~~(APPR)~~ N^{3}LO+N^{3}LL QCD (EFT); NLO ~~PURE~~ EW; NLO EXACT HQ;
NNLO APPROX TOP; NNLO PDFS

$$\sigma(\text{LHC13},\ m_H = 125\ \text{GeV}) = 48.58\ \text{pb} \pm \cancel{2.2^{\text{TH}}\ (4.5\%) \pm 1.6^{\text{PDF}+\alpha_s}\ (3.2\%)}$$
$$\sigma(\text{LHC13},\ m_H = 125\ \text{GeV}) = \qquad \text{pb} \pm 1.6^{\text{TH}}\ (3.3\%) \pm 1.4^{\text{PDF}+\alpha_s}\ (2.8\%)$$

PDF+α_s UNCERTAINTY

PDF: ~~± 0.9 pb (1.9%)~~ ± 0.5 pb (1%)
α_s: ± 1.3 pb (2.6%)

- BOTH PDF & THEORY UNCERTAINTY RAPIDLY DECREASING
- TOWARD 1% PDF UNCERTAINTIES!

SUMMARY

PAST: WHAT EVERYBODY DOES

- FROM DATA TO PDF UNCERTAINTIES
- PDF COMBINATION
- ANALYSIS TOOLS

PRESENT: WHAT WE DO

- THE IMPACT OF LHC DATA
- HEAVY QUARK TREATMENT
- THE PHOTON PDF
- α_s FROM LHC DATA

FUTURE: WHAT WE'LL SOON DO

- RESUMMED PDFS
- THEORY ERRORS
- EW CORRECTIONS

PAST

CONTEMPORARY PDF TIMELINE (ONLY PUBLISHED GLOBAL)

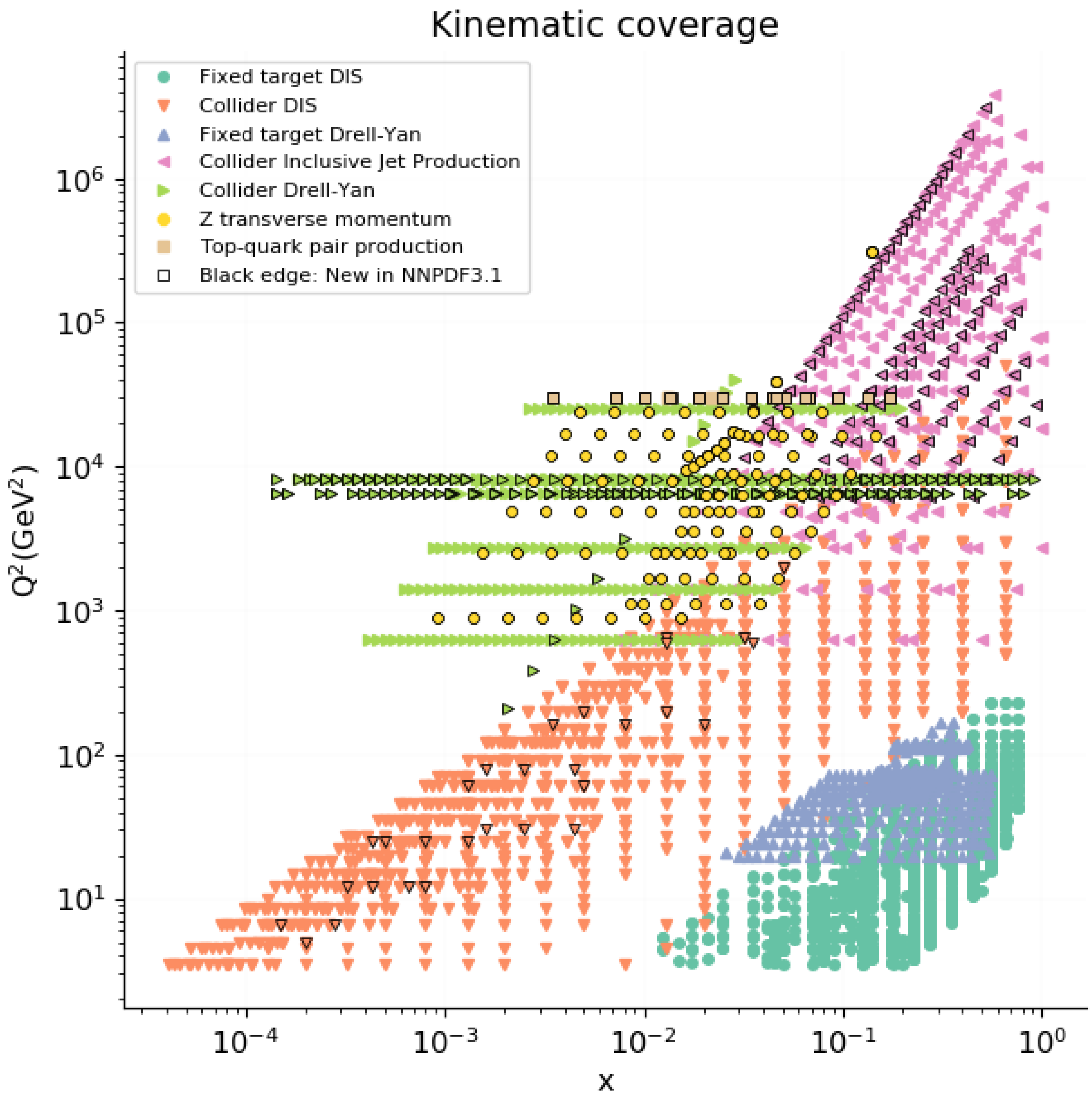
	2008		2009		2010		2011	2012		2013		2014		2015	2017	
SET / MONTH	CTEQ6.6 (02)	NNPDF1.0 (08)	MSTW (01)	ABKM09 (08)	NNPDF2.0 (02)	CT10 (NLO) (07)	NNPDF2.1 (NNLO) (07)	ABM11 (02)	NNPDF2.3 (07)	CT10 (NNLO) (02)	ABM12 (10)	NNPDF3.0 (10)	MMHT (12)	CT14 (06)	ABMP16 (01)	NNPDF3.1 (06)
F. T. DIS	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ZEUS+H1-HI	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
COMB. HI	✗	✗	✗	✗	✔	✗	✔	✗	✔	✗	✔	✔	✗	✗	✔	✔
ZEUS+H1-HII	✗	✗	✗	✗	✗	✗	some	✗	✗	some	✗	✔	✗	✗	✔	✔
HERA JETS	✗	✗	✔	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗	✗
F. T. DY	✔	✗	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
TEV W+Z	✔	✗	✔	✗	✔	✔	✔	✗	✔	✔	✗	✔	✔	✔	✗	✔
LHC W+Z	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	some	✔	✔	✔	some	✔
TEV JETS	✔	✗	✔	✗	✔	✔	✗	✔	✔	✔	✗	✔	✔	✔	✗	✔
LHC JETS	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗	✔	✔	✔	✗	✔
TOP TOTAL	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✔	✗	✗	✔	✔
SINGLE TOP TOTAL	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗
TOP DIFFERENTIAL	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔
W p_T	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗	✗	✗
W+C	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔	✗	✗	✗	✗
Z p_T	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✔

THEORY PROGRESS:

- MSTW, ABKM: all NNLO; NNPDF NNLO since 07/11 (2.1), CT since 02/13 (CT10); NNPDF THRESHOLD RESUMMATION (3.0RESUM, 07/15), SMALL x RESUMMATION (3.1SX, 10/17)
- MSTW, CT, NNPDF all GM-VFN; NNPDF since 01/11 (2.1); ABM FFN+ZM-VFN since 01/17 (ABMP16)
- NNPDF FITTED CHARM since 05/16 (NNPDF3IC)
- PHOTON PDF: (mrst2004qed), NNPDF2.3QED (08/13), NNPDF3.0QED (06/16), NNPDF3.1LUXQED (12/17)

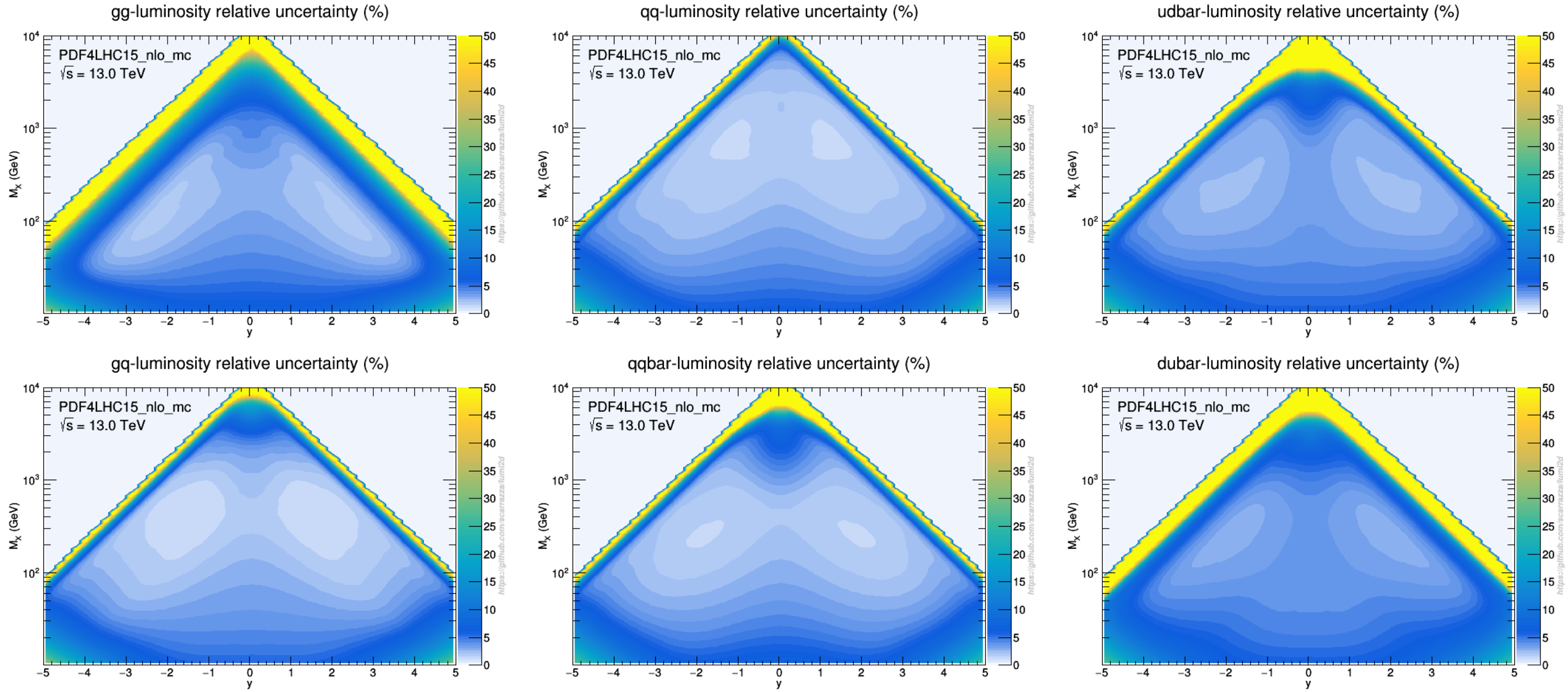
TYPICAL GLOBAL DATASET

NNPDF3.0 & NNPDF3.1



PDF4LHC15: PDF UNCERTAINTIES (NLO)

GLUON | SINGLET | FLAVORS



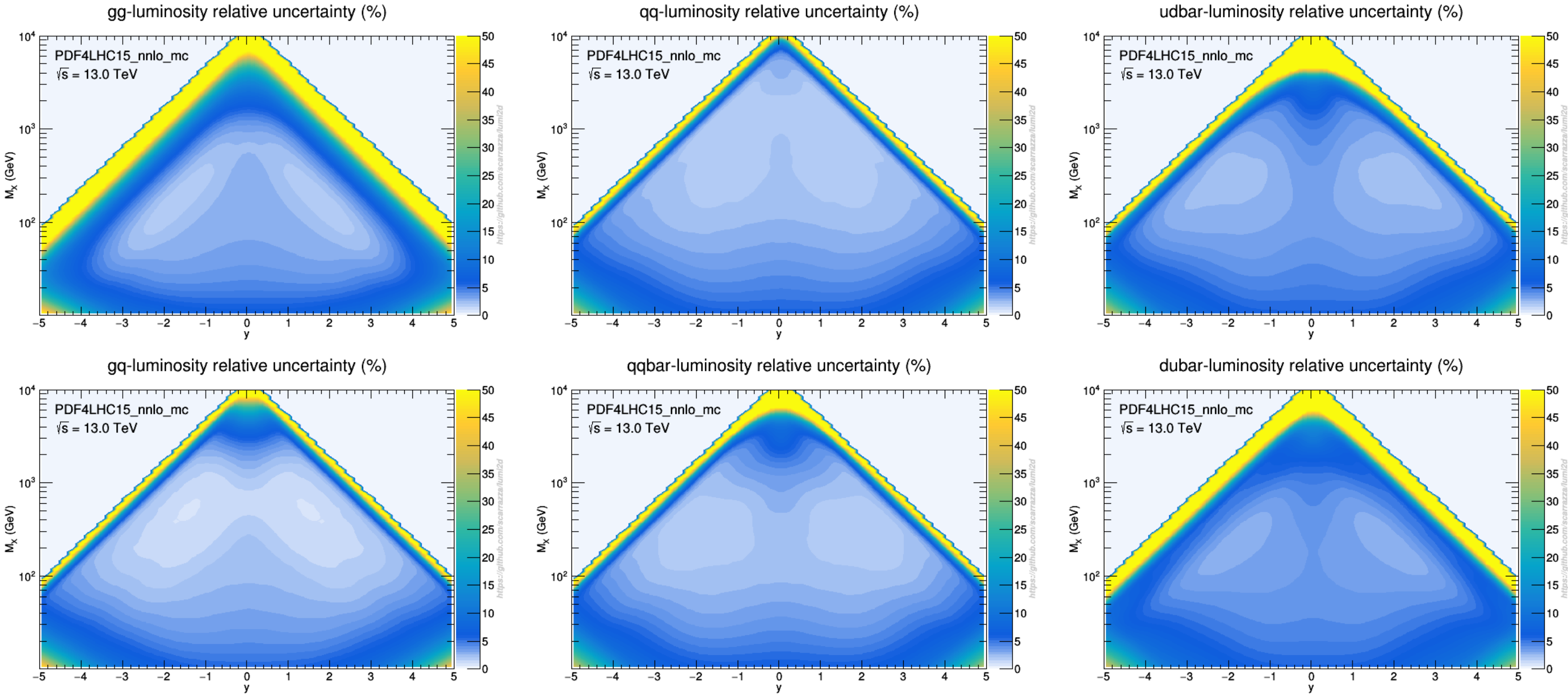
- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- TYPICAL UNCERTAINTIES IN DATA REGION $\sim 3-5\%$
- SWEET SPOT: VALENCE Q - G; DOWN TO 1%
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS
- .

PDF UNCERTAINTIES

THE KARLSRUHE PLOTS

PDF4LHC15: PDF UNCERTAINTIES (NNLO)

GLUON SINGLET FLAVORS



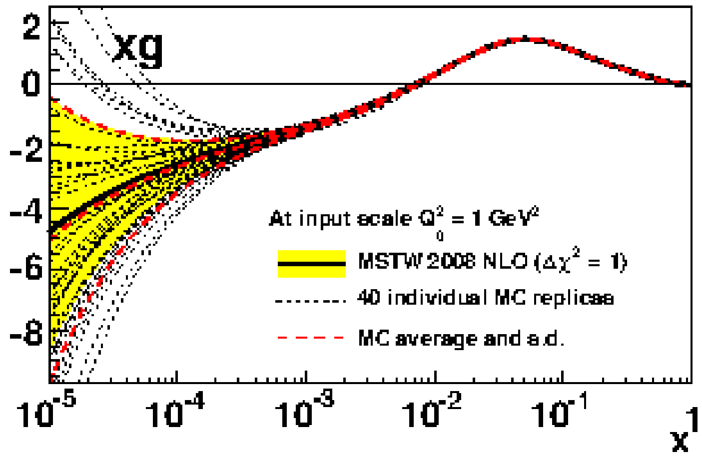
- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- TYPICAL UNCERTAINTIES IN DATA REGION $\sim 3-5\%$
- SWEET SPOT: VALENCE Q - G; DOWN TO 1%
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS
- NO QUALITATIVE DIFFERENCE BETWEEN NLO AND NNLO

PDF parametrization & delivery

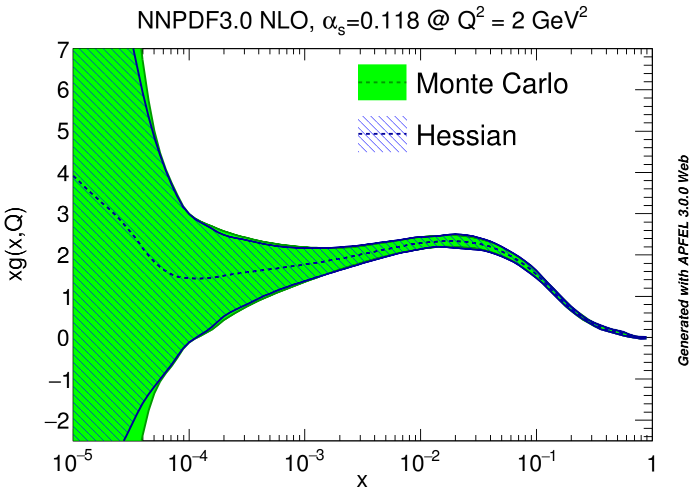
- traditionally, two delivery methods for PDFs
- HESSIAN a central PDF set, & error sets corresponding to eigenvectors of the covariance matrix in parameter space
  ADVANTAGE: efficient representation of uncertaity
  DISADVANTAGES: assumes gaussianity
- MONTECARLO a set of PDF replicas which represents the probability in PDF space (so the mean unbiasedly estimates the central value &c)
  ADVANTAGE: faithful representation of probability
  DISADVANTAGES: may need large number of replicas
- traditionally, delivery $\Leftrightarrow$ parametrization/minimization
  hessian used with relatively simple functional forms (small numbers of parameters) $\Leftrightarrow$ hessian minimization

TOOLS I
MC ⇔ HESSIAN

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



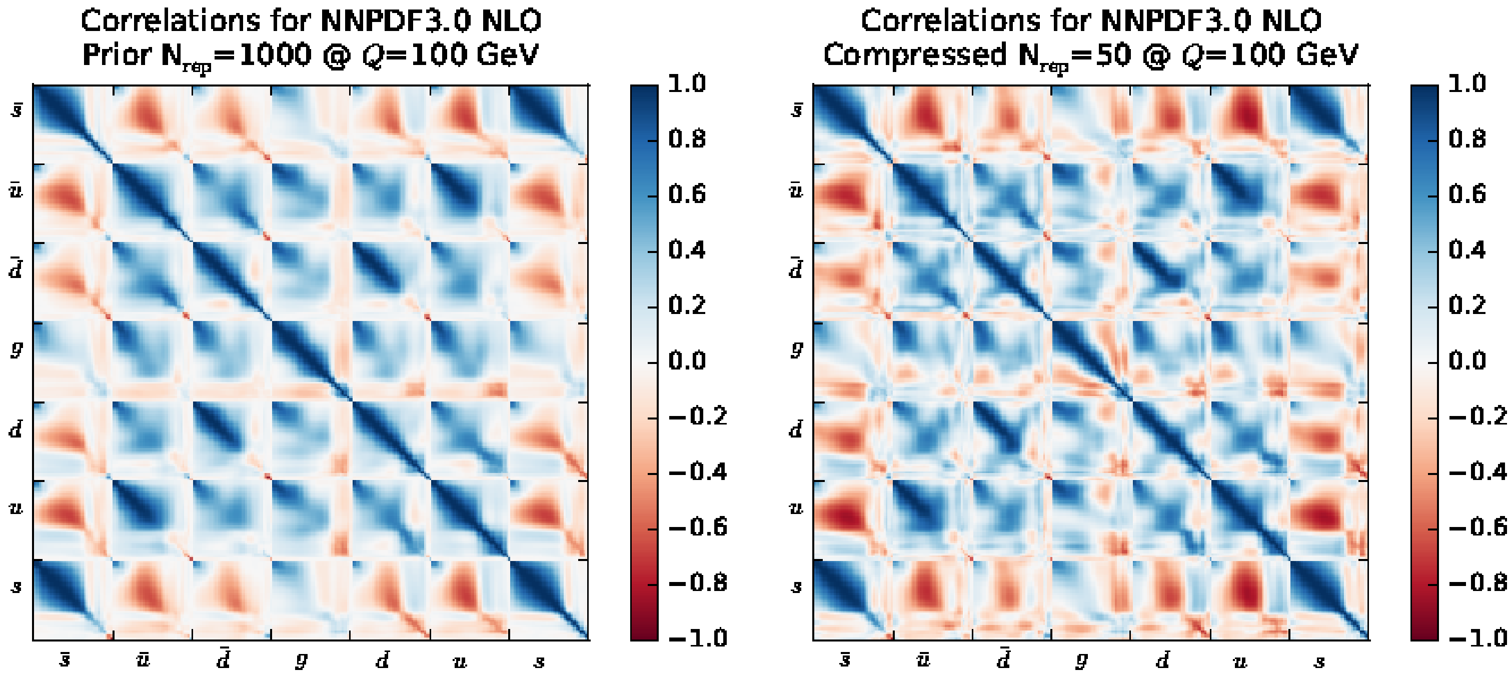
(Thorne, Watt, 2012)



(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

TOOLS II
MONTECARLO COMPRESSION



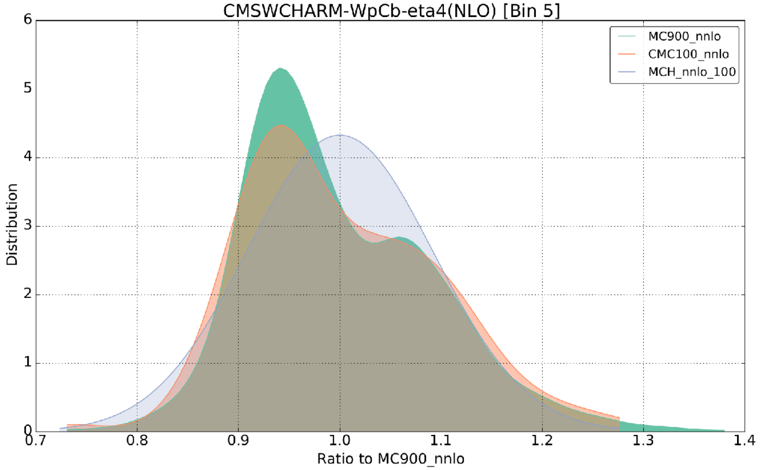
(Carrazza, Latorre, Kassabov, Rojo, 2015)

- 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

TOOLS III
NONGAUSSIAN BEHAVIOUR

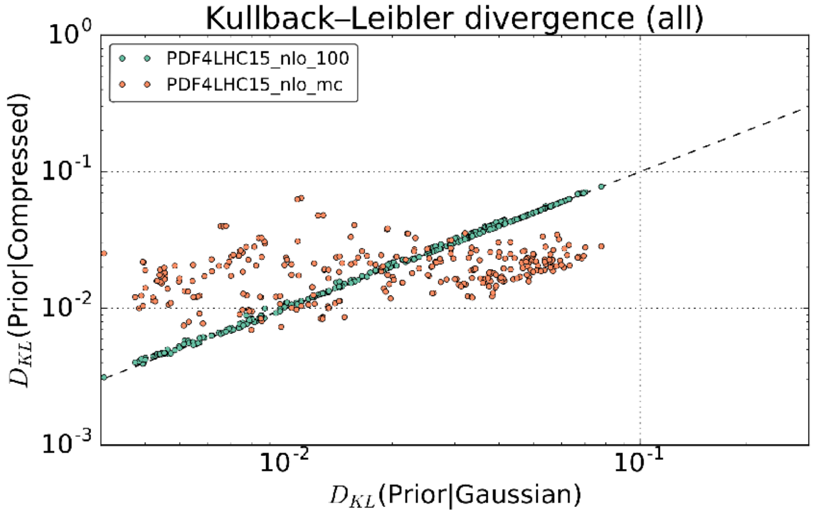
MONTE CARLO COMPARED TO HESSIAN

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

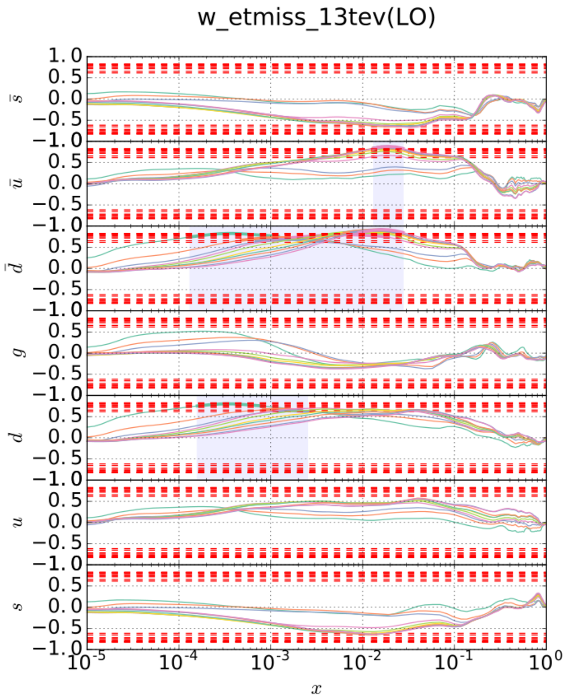
- DEFINE KULLBACK-LEIBLER DIVERGENCE $D_{\mathrm{KL}} = \int_{-\infty}^{\infty} P(x) \frac{\ln P(x)}{\ln Q(x)}\, dx$ BETWEEN A PRIOR P AND ITS REPRESENTATION Q
- D_{KL} BETWEEN PRIOR AND HESSIAN DEPENDS ON DEGREE OF GAUSSIANITY
- D_{KL} BETWEEN PRIOR AND COMPRESSED MC DOES NOT

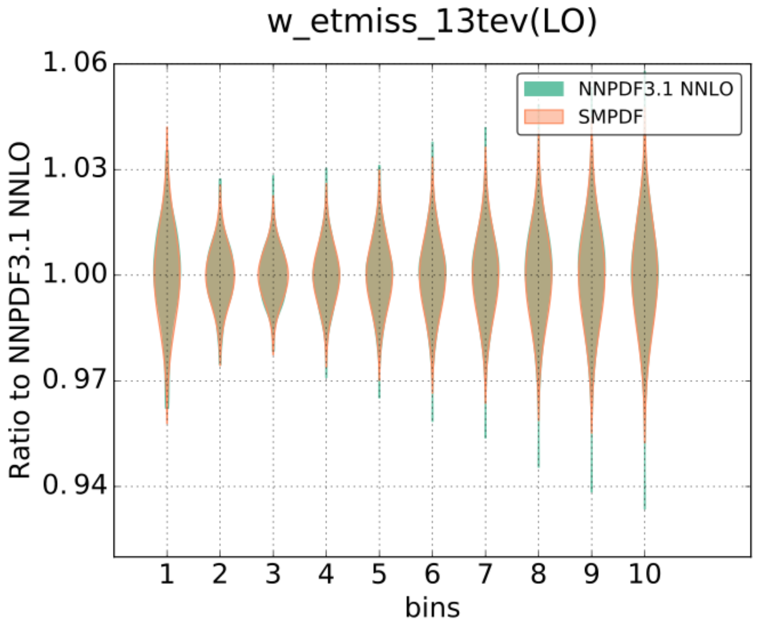


CAN (A) GAUGE WHEN MC IS MORE ADVANTAGEOUS THAN HESSIAN; (B) ASSESS THE ACCURACY OF COMPRESSION

TOOLS IV
OPTIMIZED PDFS: SMPDF

- OLD ASPIRATION: PDFS OPTIMIZED TO PROCESSES (Pumplin 2009)
- 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
- WEB INTERFACE AVAILABLE





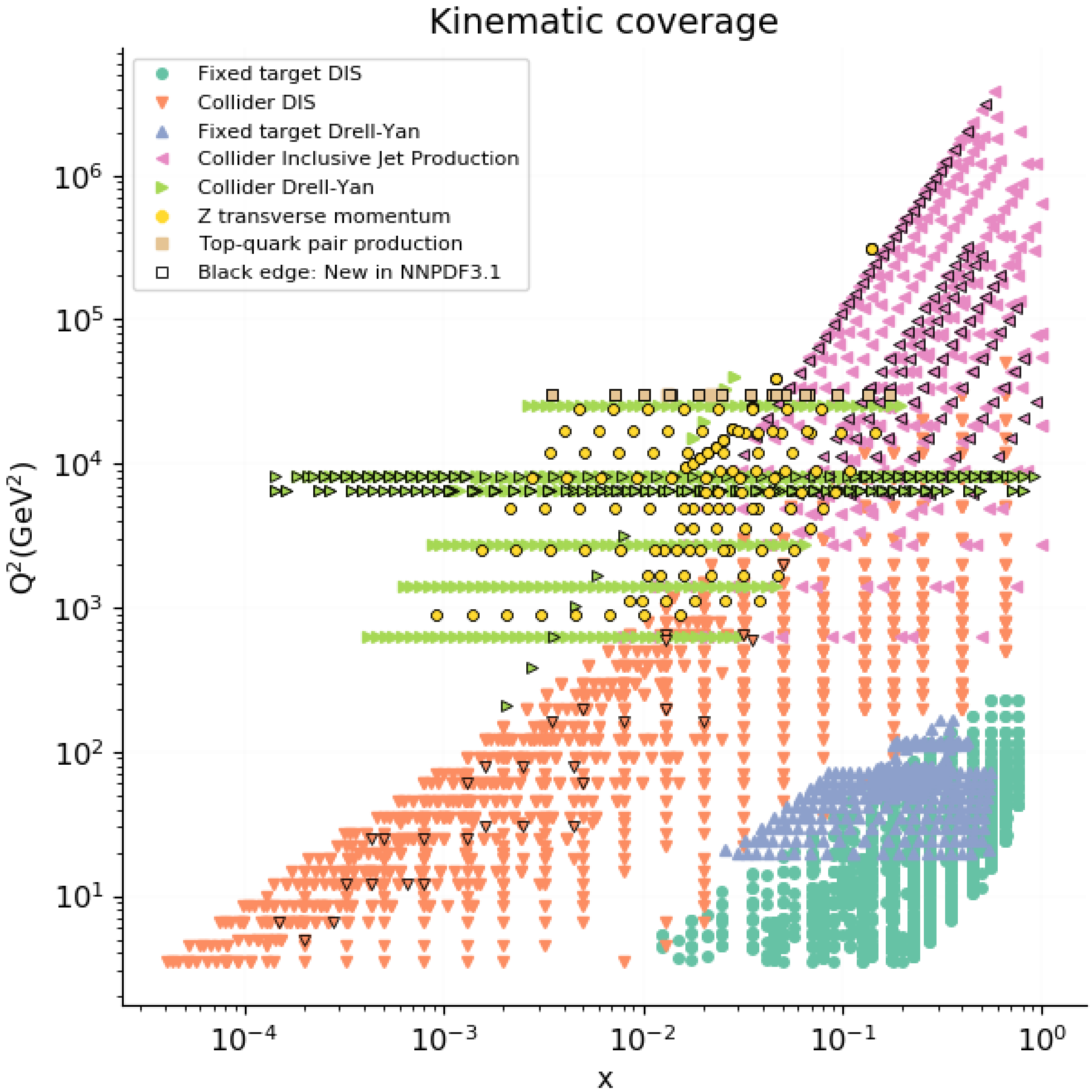
(Carrazza, SF, Kassabov, Rojo, 2016)

- EG ggH, $Hb\bar{b}$, W $E_T^{\rm miss}$ $\Rightarrow$ 11 EIGENVECTORS
- STUDY CORRELATIONS OF PDFS TO DATA AND AMONG THEMSELVES!

PRESENT

DATASET WIDENING

NNPDF3.0 vs NNPDF3.1

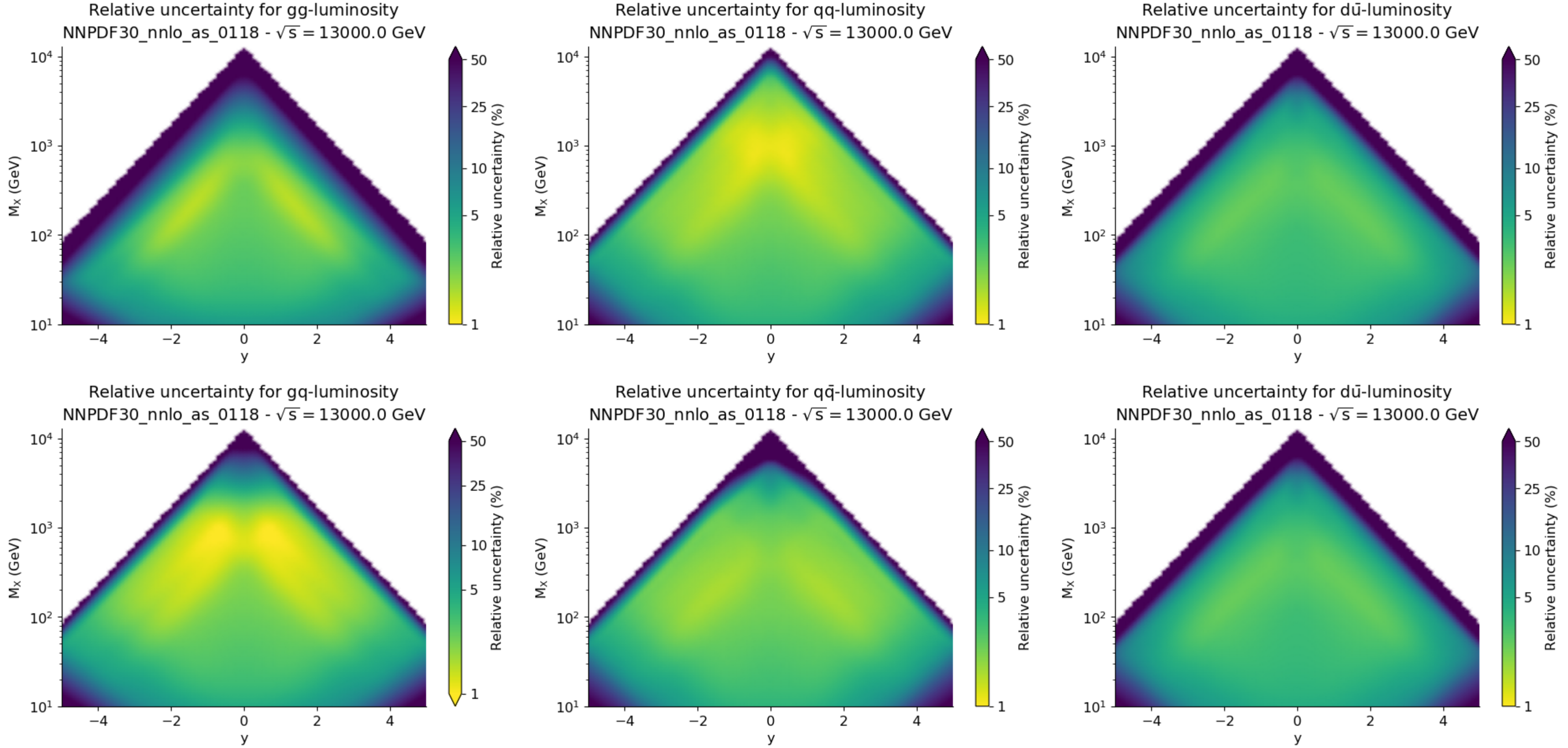


NEW DATA: (BLACK EDGE)

- HERA COMBINED F_2^b
- D0 W LEPTON ASYMMETRY
- ATLAS W, Z 2011, HIGH & LOW MASS DY 2011; CMS $W^\pm$ RAPIDITY 8TEV LHCB W, Z 7TEV & 8TEV
- ATLAS 7TEV JETS 2011, CMS 2.76TEV JETS
- ATLAS & CMS TOP DIFFERENTIAL RAPIDITY
- ATLAS Z p_T DIFFERENTIAL RAPIDITY & INVARIANT MASS 8TEV, CMS Z p_T DIFFERENTIAL RAPIDITY 8TEV

THE IMPACT OF LHC DATA
PDF UNCERTAINTIES IN DETAIL: NNPDF3.0 (NNLO)

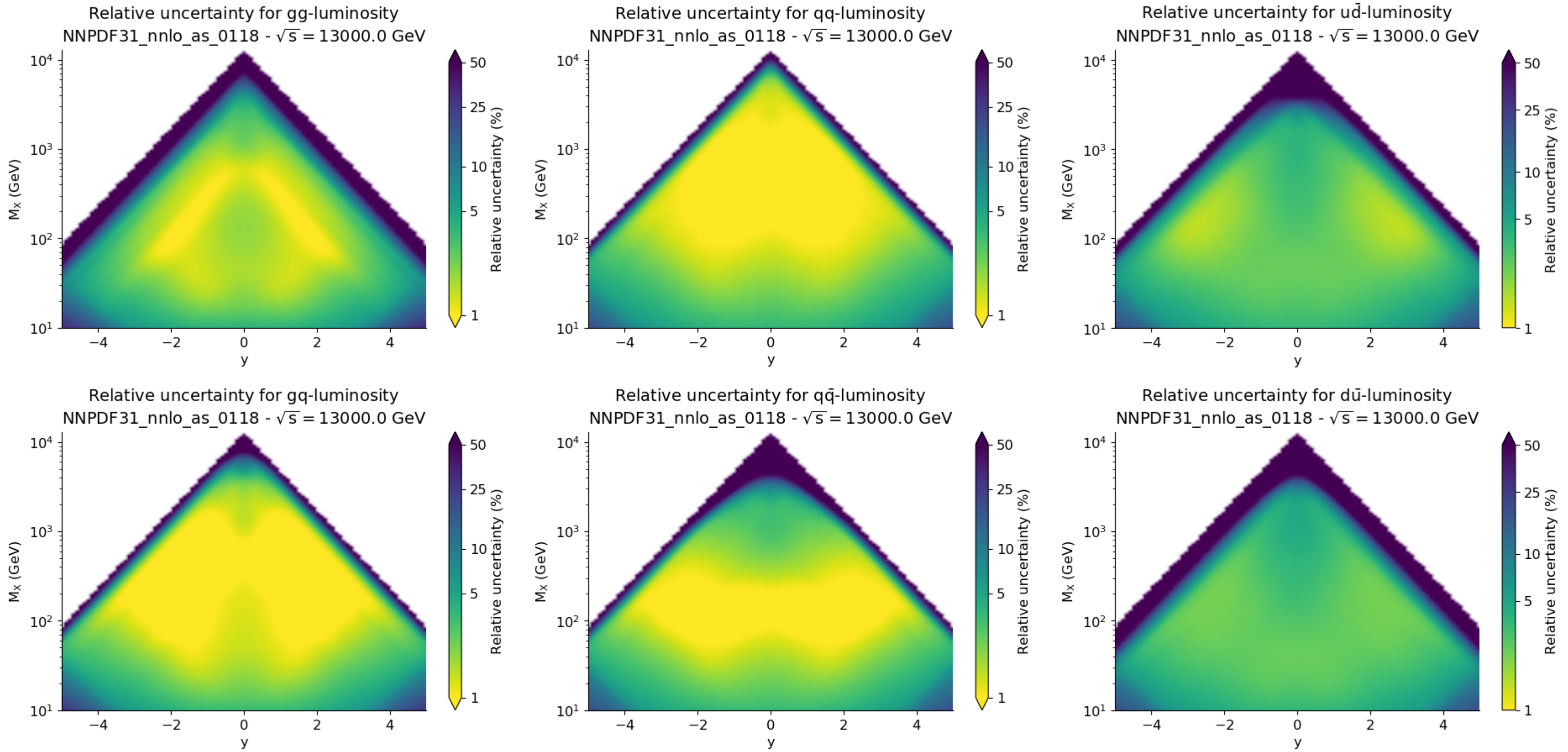
GLUON SINGLET FLAVORS



- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- TYPICAL UNCERTAINTIES IN DATA REGION $\sim 3-5\%$
- SWEET SPOT: VALENCE Q - G; DOWN TO 1%
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS

THE IMPACT OF LHC DATA
PDF UNCERTAINTIES IN DETAIL: NNPDF3.1 (NNLO)

GLUON SINGLET FLAVORS



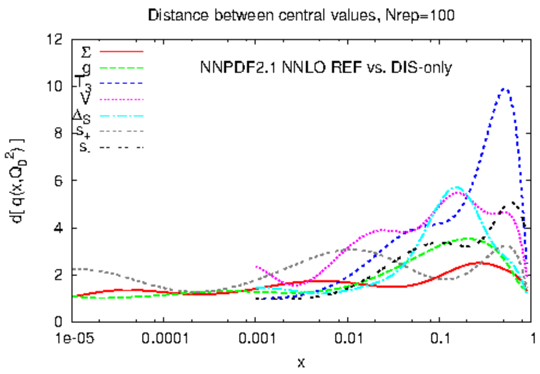
- GLUON BETTER KNOWN AT SMALL x, VALENCE QUARKS AT LARGE x, SEA QUARKS IN BETWEEN
- TYPICAL UNCERTAINTIES IN DATA REGION $\sim 1 - 3\%$
- SWEET SPOT: VALENCE Q - G; 1% OR BELOW
- UP BETTER KNOWN THAN DOWN; FLAVOR SINGLET BETTER THAN INDIVIDUAL FLAVORS
- NEW LHC DATA $\Rightarrow$ SIZABLE REDUCTION IN UNCERTAINTIES

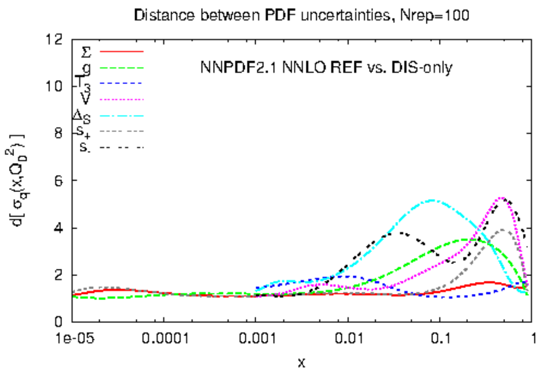
THE IMPACT OF LHC DATA

BEFORE LHC: PDFS MOSTLY DETERMINED BY DIS DATA

NNPDF2.1 VS NNPDF2.1 DIS ONLY

DISTANCES (difference in units of st. dev.)

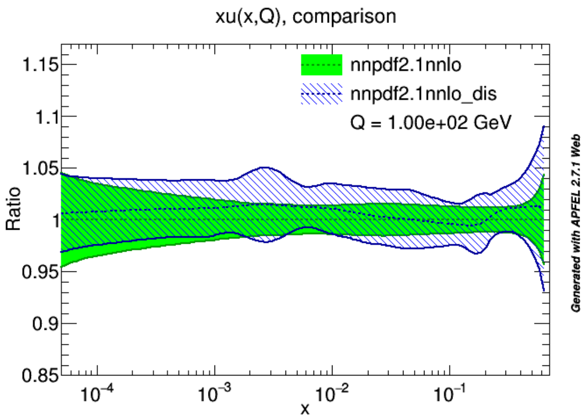




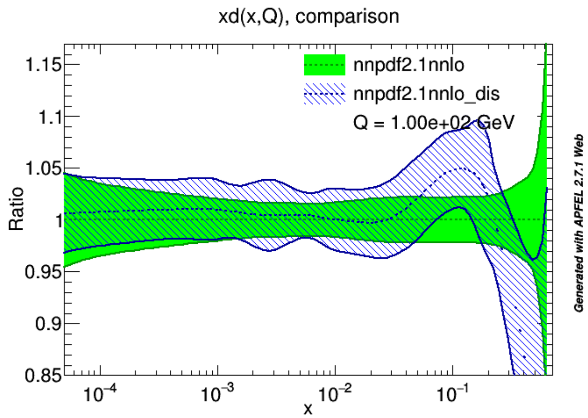
$d = 10 \Leftrightarrow$ one sigma difference

PDF COMPARISON

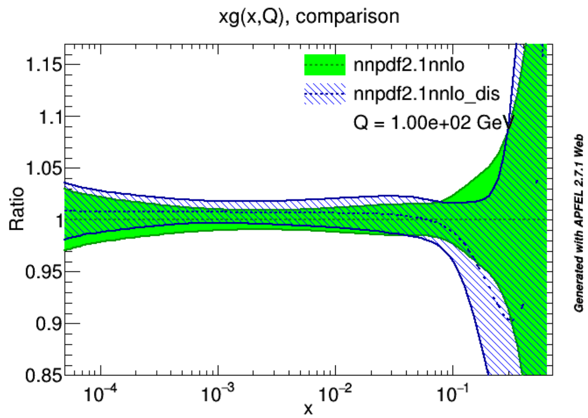
UP



DOWN



GLUON

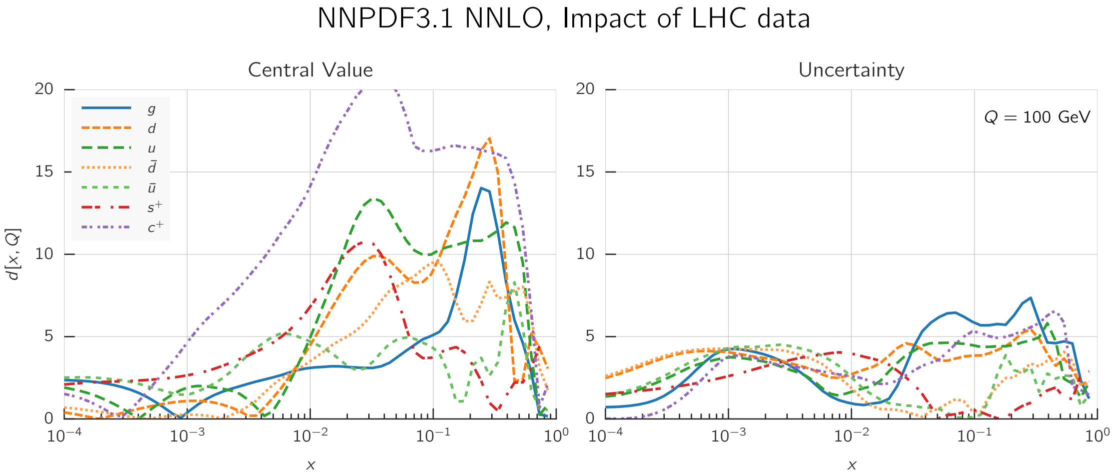


- ALL DIFFERENCES BELOW ONE SIGMA
- ONLY UP-DOWN SEPARATION SIGNIFICANTLY AFFECTED

THE IMPACT OF LHC DATA

NOW: PDFs LARGELY DETERMINED BY LHC DATA

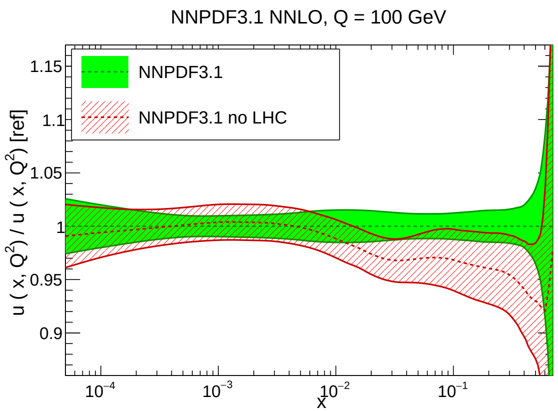
NNPDF3.1 VS NNPDF3.1 NO LHC
DISTANCES (difference in units of st. dev.)



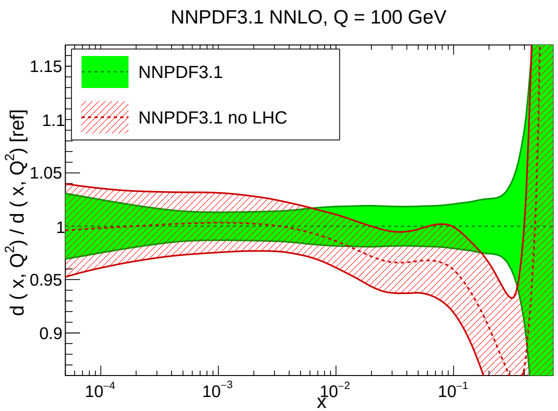
$d = 10 \Leftrightarrow$ one sigma difference

PDF COMPARISON

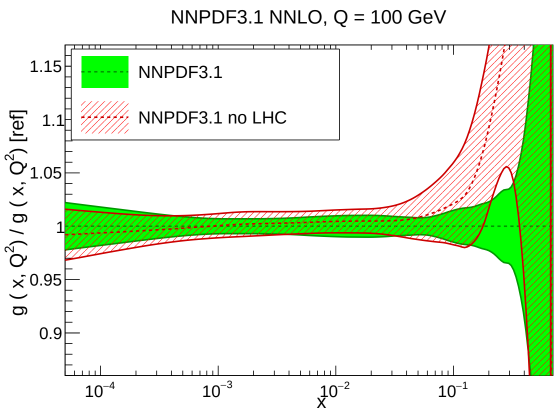
UP



DOWN



GLUON

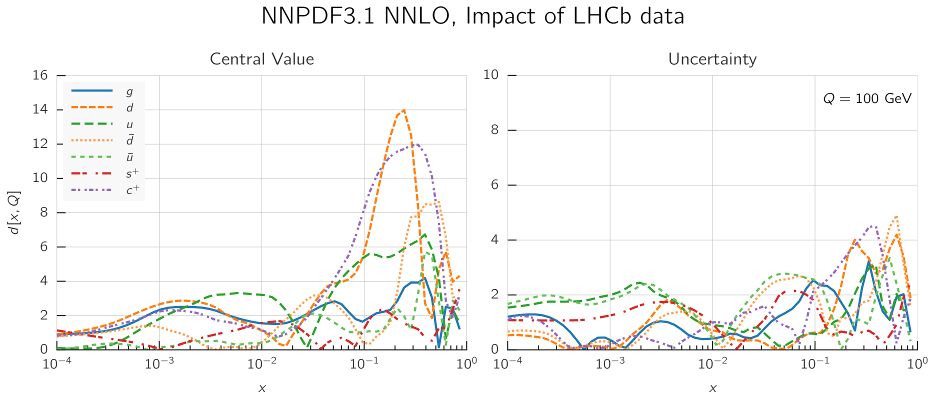


- MANY PDFs CHANGE BY MORE THAN ONE SIGMA
- BOTH FLAVOR SEPARATION & GLUON SIGNIFICANTLY AFFECTED

THE IMPACT OF LHC DATA
FLAVOR SEPARATION

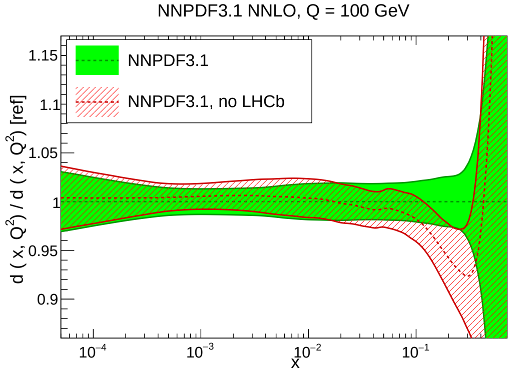
- BEFORE LHC $\Rightarrow$ CC DIS, TeV FIXED-TARGET DY, W ASYM.
- AFTER LHC $\Rightarrow$ WIDE RANGE OF W, Z PRODUCTION DATA

IMPACT OF LHCb
DISTANCES (difference in units of st. dev.)

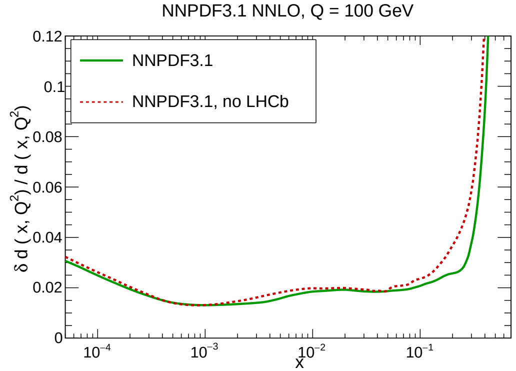


PDF COMPARISON: DOWN

CENTRAL VALUE



UNCERTAINTY

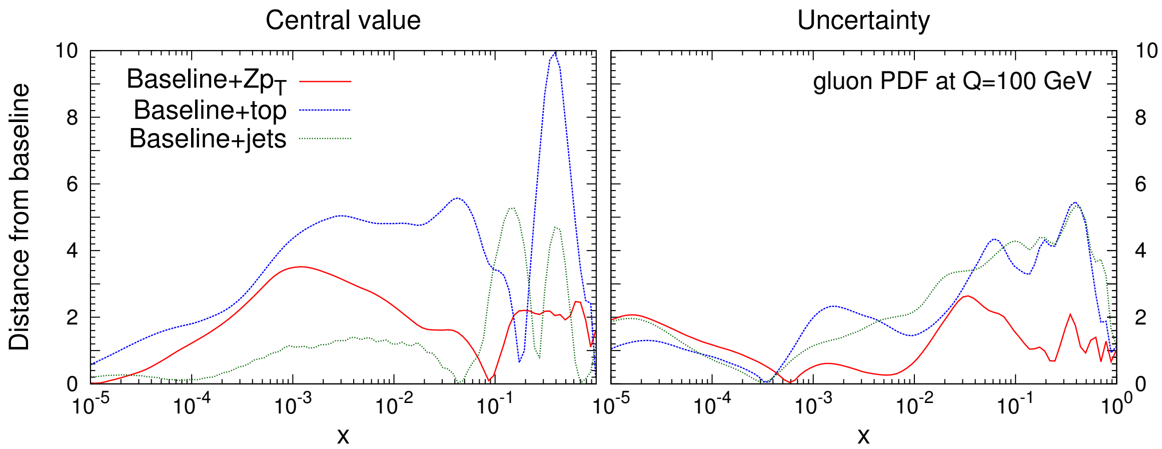


- SIZABLE SHIFT OF CENTRAL VALUE BY ALMOST ONE SIGMA
- LARGE x UNCERTAINTY DOWN BY LARGE FACTOR!

THE IMPACT OF LHC DATA
THE GLUON

- BEFORE LHC ⇒ DIS SCALING VIOLATIONS, TEV JETS AT LARGE X
- AFTER LHC ⇒ JETS; $Z\ p_t$, TOP

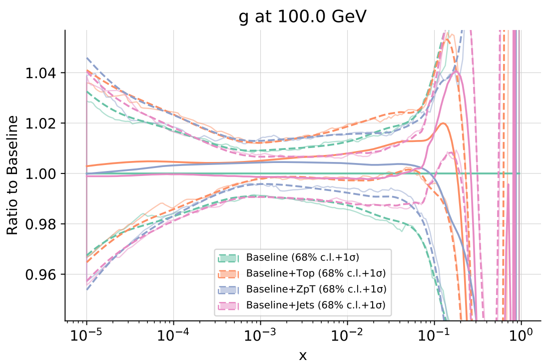
DISTANCES (difference in units of st. dev.)



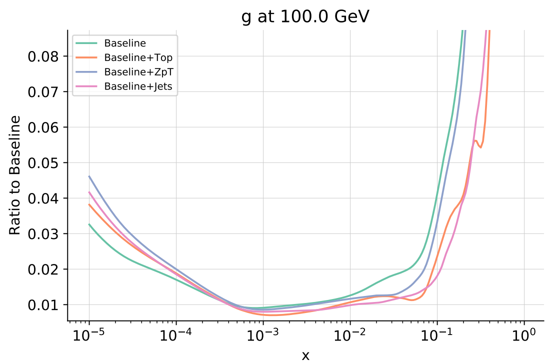
(Nocera, Ubiali, 2017)

PDF COMPARISON: GLUON

CENTRAL VALUE



UNCERTAINTY

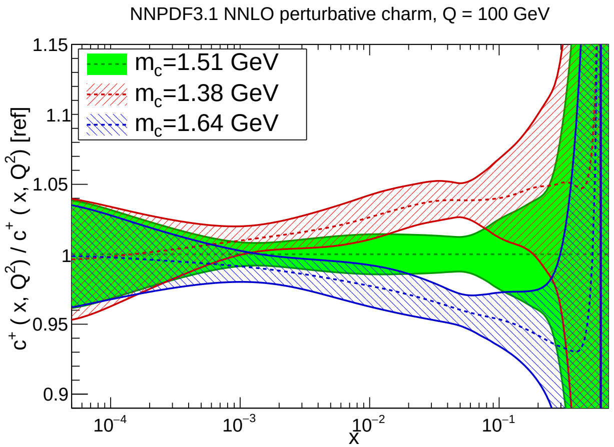


- TOP HAS LARGEST IMPACT, FOLLOWED BY JETS
- ALL LHC DATA PULL CENTRAL VALUE IN SAME DIRECTION!

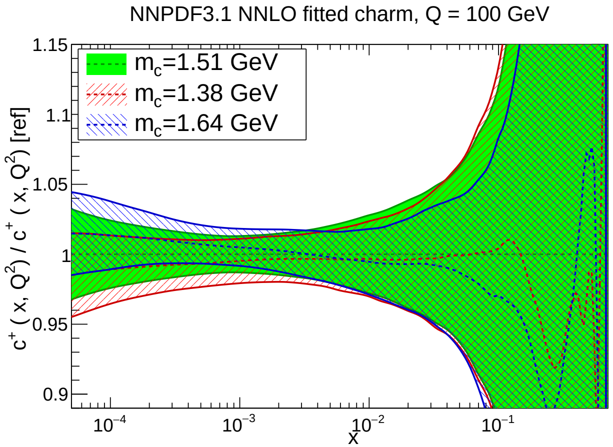
HEAVY QUARK PDFs
CHARM FROM DATA

- CHARM SHOULD NOT DEPEND STRONGLY ON CHARM MASS

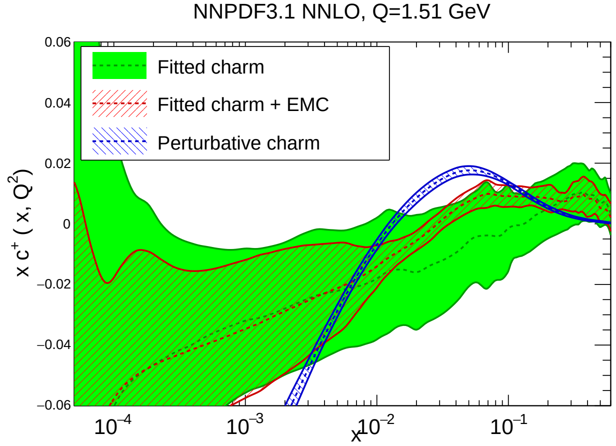
PERTURBATIVE CHARM VS m_c



FITTED CHARM VS m_c



FITTED VS. PERTURBATIVE CHARM



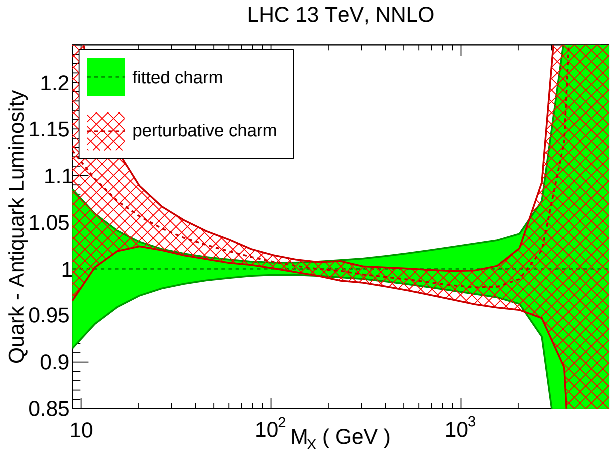
- ITS SHAPE SHOULD NOT BE DETERMINED BY FIRST-ORDER MATCHING (NO HIGHER NONTRIVIAL ORDERS KNOWN)

- MIGHT EVEN HAVE A NONPERTURBATIVE COMPONENT

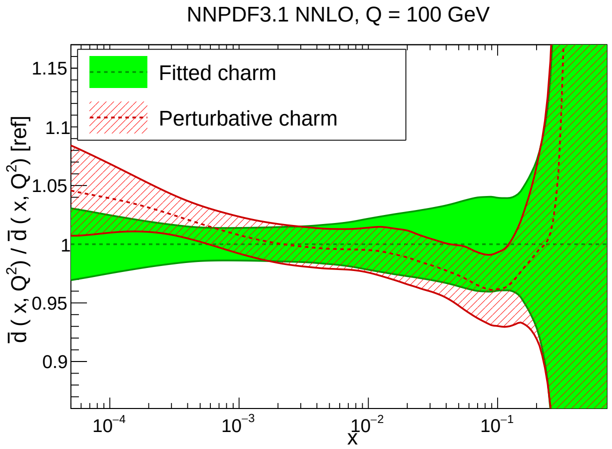
FITTED VS. LO PERTURBATIVE:
SUPPRESSED AT MEDIUM-SMALL x, ENHANCED AT VERY SMALL, VERY LARGE x

THE CHARM PDF FROM DATA
IMPACT ON LIGHT QUARK PDFS
FITTED VS. PERTURBATIVE CHARM

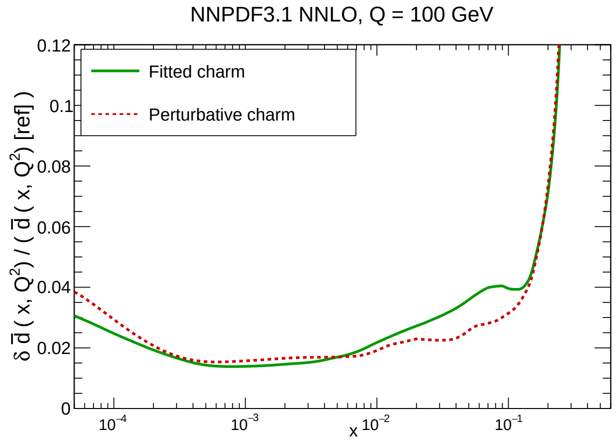
QQBAR LUMI



ANTIDOWN PDF



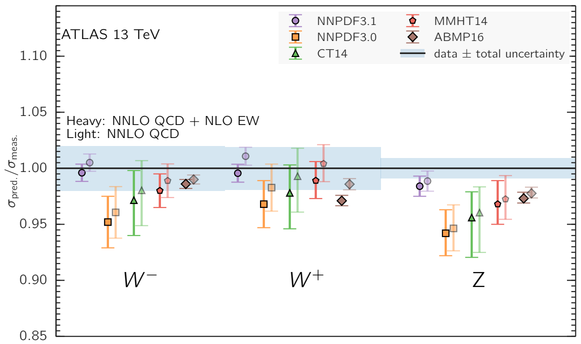
ANTIDOWN PDF UNCERTAINTY



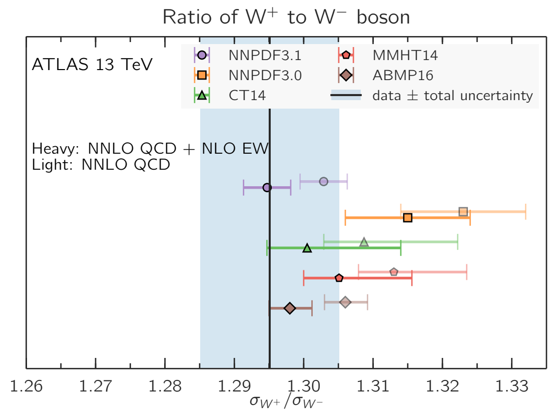
- QUARK (ESPECIALLY QUARK-ANTIQUARK) LUMI AFFECTED BECAUSE OF CHARM SUPPRESSION AT MEDIUM-x
- FLAVOR DECOMPOSITION ALTERED
- UNCERTAINTIES ON LIGHT QUARKS NOT SIGNIFICANTLY INCREASED
- AGREEMENT OF 13TeV W,Z PREDICTED CROSS-SECTIONS IMPROVES!

CHARM FROM DATA
IMPACT ON PHENOMENOLOGY

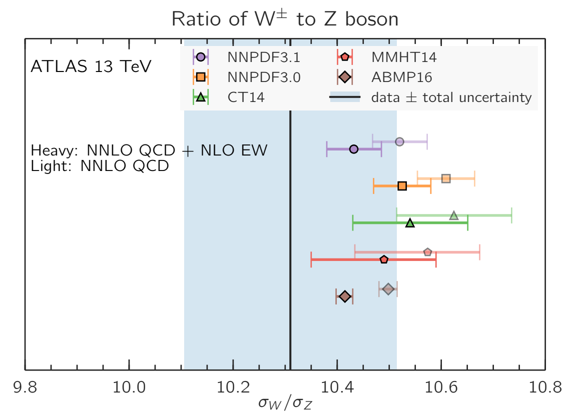
DRELL-YAN XSECTS



W^+/W^- XSECT RATIO



W/Z XSECT RATIO



- W, Z CROSS-SECTIONS AT 13 TEV IN PERFECT AGREEMENT WITH DATA THANKS TO FITTED CHARM!
- ELECTROWEAK CORRECTIONS IMPORTANT

THE PHOTON PDF BREAKTHROUGH

(Manohar, Nason, Salam, Zanderighi, 2016)

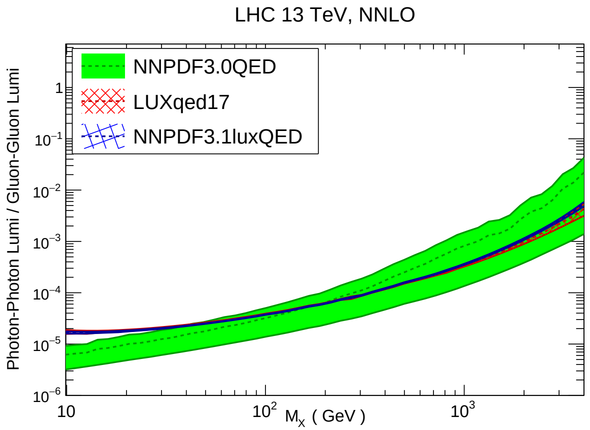
- QED IS PERTURBATIVE DOWN TO LOW SCALES ⇒ THE PHOTON PDF MUST BE COMPUTABLE IF THE INPUT QUARK SUBSTRUCTURE IS KNOWN
- ⇒ PDF EXPRESSED IN TERMS OF THE STRUCTURE FUNCTION INTEGRATED OVER ALL SCALES
- LUX16/LUX17 SETS CONSTRUCTED FROM PDF4LHC15 ⇒ AGREE WELL WITH NNPDF3.0 QED, MUCH SMALLER UNCERTAINTY

THE LUXQED PHOTON PDF

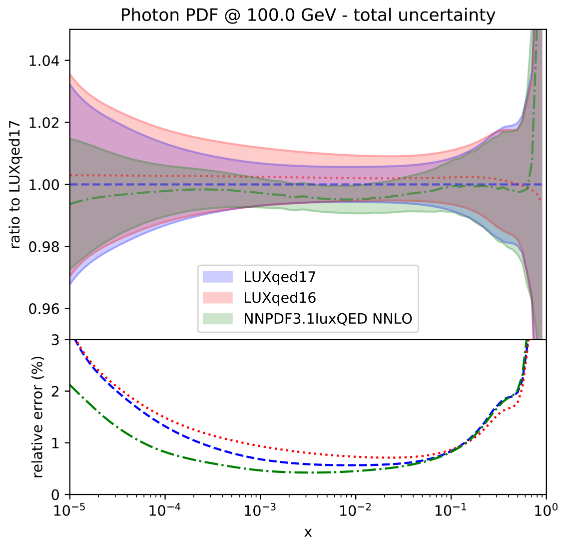
(Carrazza et al., 2017)

- FIRST PDF SET BASED ON CONSISTENT FIT WITH LUX CONSTRAINT: NNPDF3.1LUXQED
- NNPDF3.1LUXQED VS LUX17: GOOD AGREEMENT BUT SMALLER UNCERTAINTIES
- SIZABLE IMPACT ON PRECISION PHYSICS: EG ASSOCIATE HIGGS PROD. WITH W

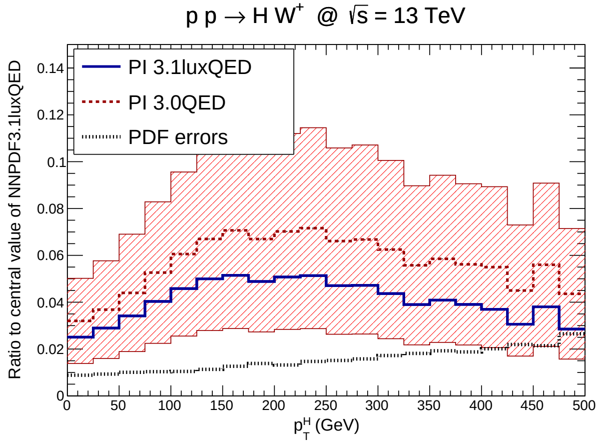
$\gamma\gamma$ LUMI COMPARISON



γ PDF COMPARISON

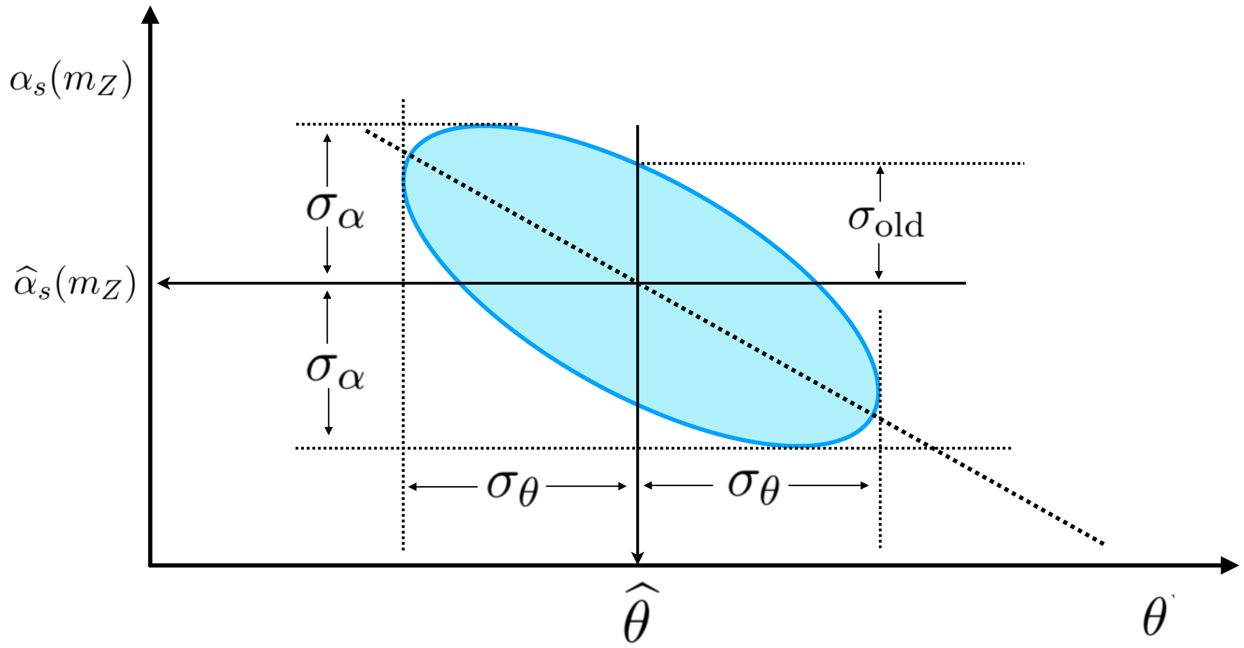


γ-INDUCED VS QCD: HW



α_s DETERMINATION

WHAT'S THE PROBLEM?



- MINIMUM DETERMINED ALONG THE "BEST PDF" LINE $\Rightarrow \sigma_{\text{old}}$
  FOR HIGHLY CORRELATED VARIABLES & UNEQUAL SEMIAXES,
  MAY UNDERESTIMATE ONE-σ ERROR $\Rightarrow \sigma_\alpha$
- IN NNPDF METHODOLOGY,
  PDF UNCERTAINTY $\Leftrightarrow$ DETERMINED FROM REPLICA SAMPLE VARIANCE
  $\Rightarrow$ IS IT $\Delta\chi^2 = 1$? (TOLERANCE?)

NEED SIMULTANEOUS MINIMZATION IN (PDF, α_s) SPACE!

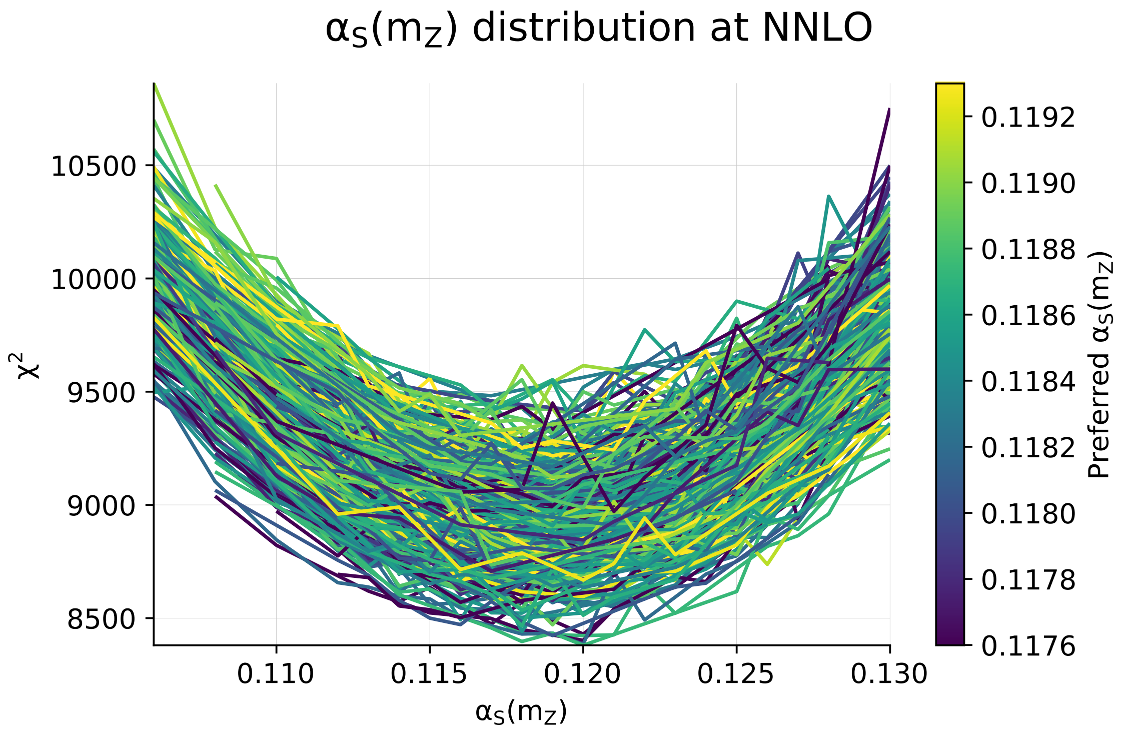
THE CORRELATED REPLICA METHOD
NNPDF3.1 (2018)

- NNPDF METHOD $\Rightarrow$ EACH PDF REPLICA FITTED BY GA TO DATA REPLICA
- IDEALLY PERFORM GENETIC MINIMIZATION IN (PDF, α_s) SPACE
- PROBLEM THEORY PREDICTION $\Leftrightarrow$ PRECOMPUTED GRIDS
  DEPEND ON α_s $\Rightarrow$ DIFFICULT TO TREAT AS CONTINUOUS PARAMETER

THE CORRELATED REPLICA METHOD
NNPDF3.1 (2018)

- NNPDF METHOD ⇒ EACH PDF REPLICA FITTED BY GA TO DATA REPLICA
- IDEALLY PERFORM GENETIC MINIMIZATION IN (PDF, α_s) SPACE
- PROBLEM THEORY PREDICTION ⇔ PRECOMPUTED GRIDS DEPEND ON α_s ⇒ DIFFICULT TO TREAT AS CONTINUOUS PARAMETER
- SOLUTION DETERMINE BEST-FIT PDF REPLICA TO EACH DATA REPLICA FOR SEVERAL (DISCRETE) α_s VALUES: C-REPLICA
  - EACH C-REPLICA ⇒ χ^2 PROFILE ⇒ α_s VALUE
  -

ENSEMBLE OF PARABOLAS

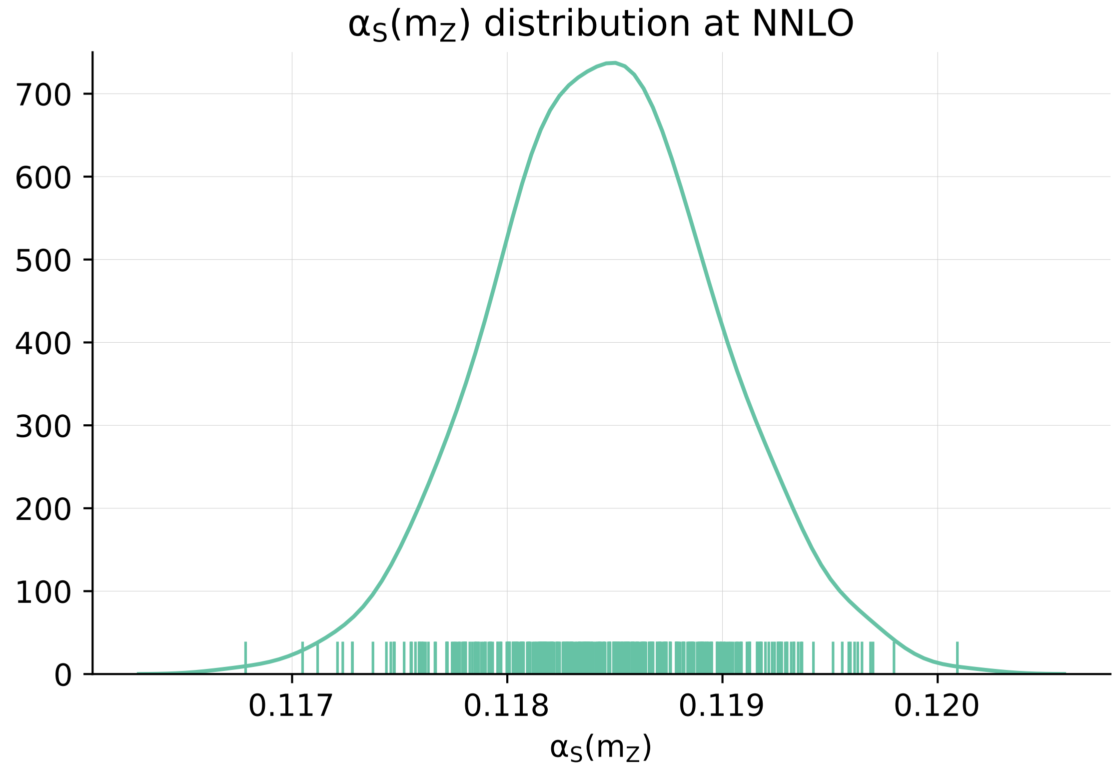


THE CORRELATED REPLICA METHOD

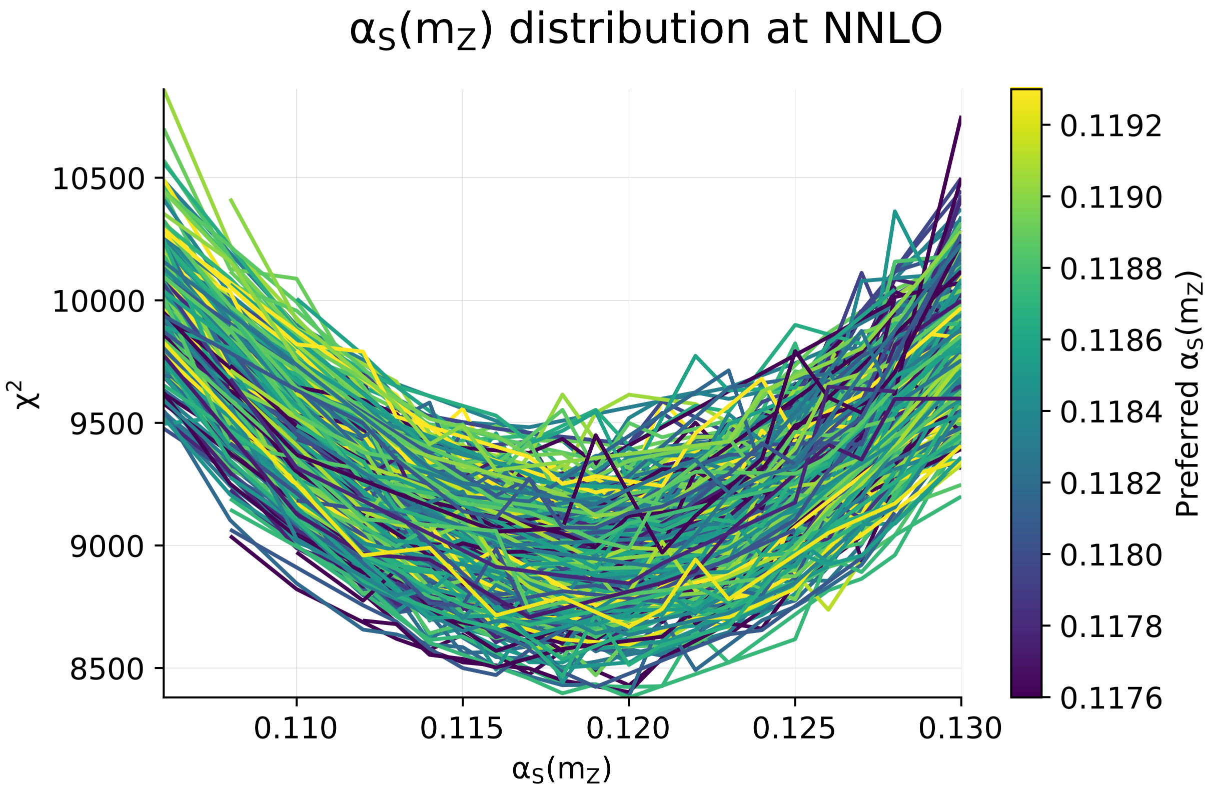
NNPDF3.1 (2018)

- NNPDF METHOD $\Rightarrow$ EACH PDF REPLICA FITTED BY GA TO DATA REPLICA
- IDEALLY PERFORM GENETIC MINIMIZATION IN (PDF, α_s) SPACE
- PROBLEM THEORY PREDICTION $\Leftrightarrow$ PRECOMPUTED GRIDS DEPEND ON α_s $\Rightarrow$ DIFFICULT TO TREAT AS CONTINUOUS PARAMETER
- SOLUTION DETERMINE BEST-FIT PDF REPLICA TO EACH DATA REPLICA FOR SEVERAL (DISCRETE) α_s VALUES:
  - EACH C-REPLICA $\Rightarrow$ χ^2 PROFILE $\Rightarrow$ α_s VALUE
  - EACH C-REPLICA $\Rightarrow$ BEST-FIT α_sREPLICA

ENSEMBLE OF α_s VALUES



α_s FROM CORRELATED REPLICAS



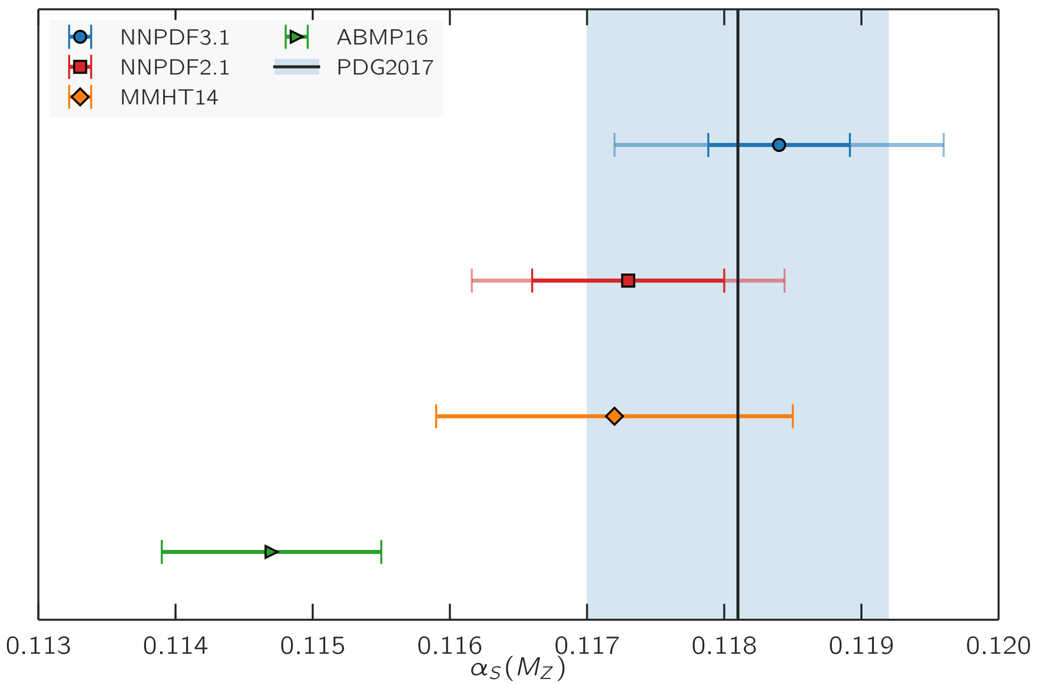
- NNPDF3.1 DATASET (ONLY NNLO JET DATA) $\Rightarrow$ 3979 DATAPOINTS
- 400 C-REPLICAS FOR 21 α_s VALUES: $\alpha_s(M_z) =$ 0.106, 0.108, 0.102, 0.112, 0.113, 0.114, 0.115, 0.116, 0.117, 0.118, 0.119, 0.120, 0.121, 0.122, 0.123, 0.124, 0.125, 0.126, 0.127, 0.128, 0.130
- EXPERIMENTAL UNCERTAINTY $\Leftrightarrow$ STANDARD DEV. OVER REPLICA SAMPLE

THE RESULT

$$\alpha_s^{\rm NNLO}(M_Z) = 0.11845 \pm 0.00052^{\rm exp}\ (0.4\%)$$

FINAL RESULT & COMPARISON

$$\alpha_s^{\mathrm{NNLO}}(M_Z) = 0.1185 \pm 0.0005^{\mathrm{exp}} \pm 0.0001^{\mathrm{meth}} \pm 0.0011^{\mathrm{th}} = 0.1185 \pm 0.0012 \; (1\%)$$

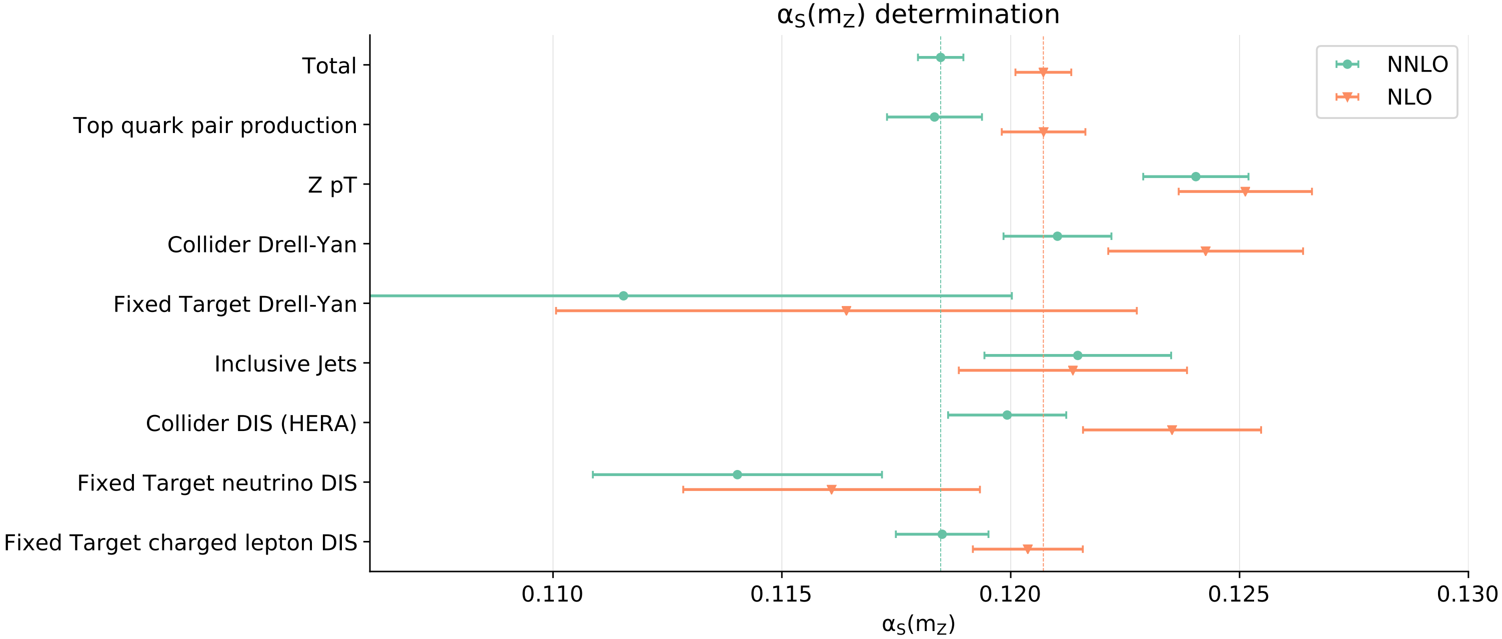


- SIGNIFICANTLY SMALLER EXP. UNCERTAINTY IN COMPARISON TO PREVIOUS NNPDF2.1 DETERMINATION (DESPITE MORE CONSERVATIVE ESTIMATE)
- SOMEWHAT LARGER CENTRAL VALUE THAN MMHT

α_s FROM A GLOBAL FIT

PULLS FROM DATA SUBSETS

α_s values from minimization of partial χ^2 for each c-replica



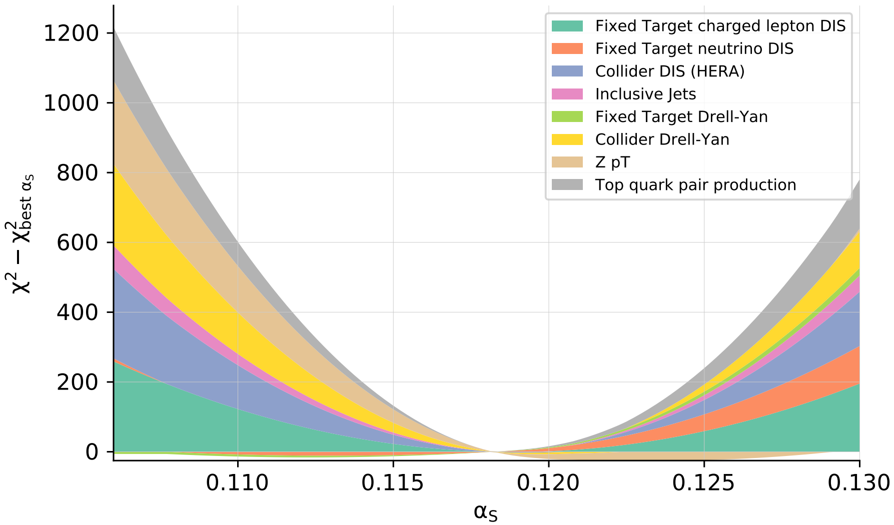
PULLS DON'T ADD TO ZERO?!

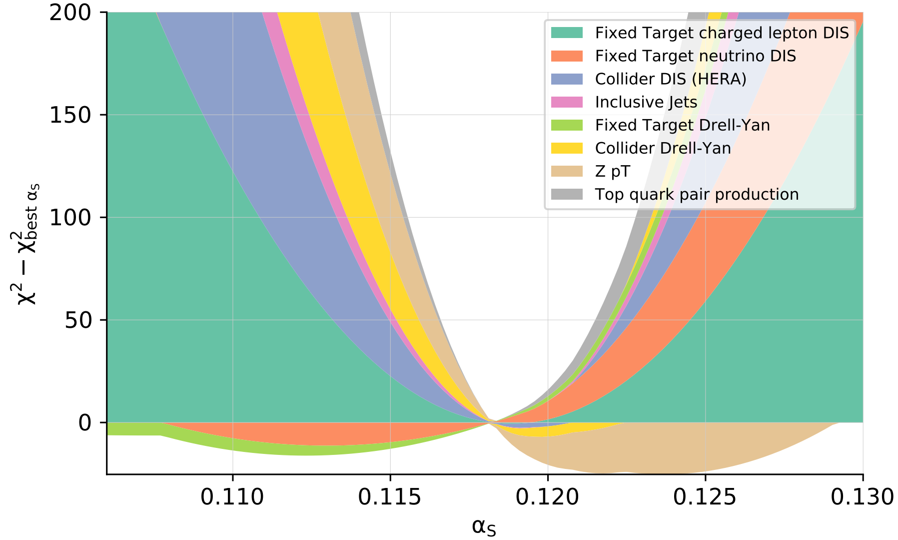
- PARTIAL VALUES ARE NOT PARTIAL BEST-FITS (PDFs COME FROM GLOBAL FIT)
- PULLS CORRELATED THROUGH UNDERLYING PDFs
- REPLICA SELECTION BASED ON GLOBAL SET (SOME PARABOLAE $\Rightarrow$ NO MINIMUM)
- MISSING SYSTEMATIC CORRELATION BETWEEN DATASETS

CANNOT DETERMINE α_s WITHOUT ALSO DETERMINING THE PDF

THE IMPACT OF LHC DATA

PARTIAL CONTRIBUTION TO χ^2 ALONG BEST-FIT LINE





LESSONS

- COLLIDER DATA HAVE A LARGE IMPACT
- NO SIGNS OF TENSION OR INCOMPATIBILITY
- LHC DATA PULL TOWARDS LARGER α_s

FUTURE

RESUMMED PDFs

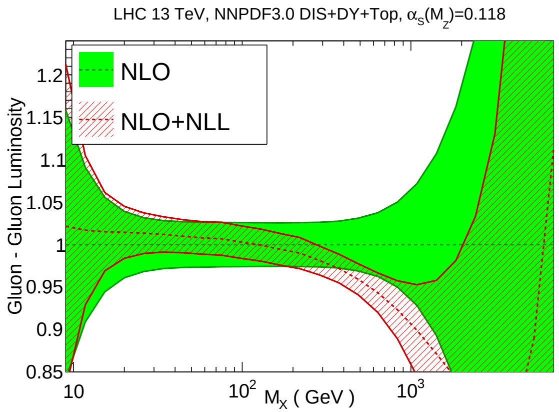
- RESUMMATION NOT INCLUDED IN DEFAULT PDF SETS
- RESUMMED CALCULATIONS MUST USE RESUMMED PDFs! (M. Spira)
- KEPT UNDER CONTROL IN FITS BY CHOICE OF CUTS

PDFs WITH THRESHOLD (LARGE x) RESUMMATION

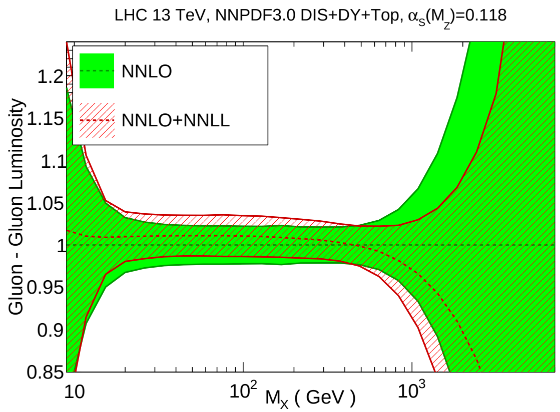
- FIRST SET: `NNPDF3.0resum`
- RESUMMATION INCLUDED IN FIT (DIS, DY, TOP DATA), EFFECTS NOT NEGLIGIGLE AT NLLO, LARGE x, MORE MODERATE AT NNLO
- EFFECT ON PDFs COMPARABLE TO EFFECT ON MATRIX ELEMENT, ANTICORRELATED TO IT
- RELEVANT FOR NEW PHYSICS SEARCHES

(Bonvini et al., 2015)

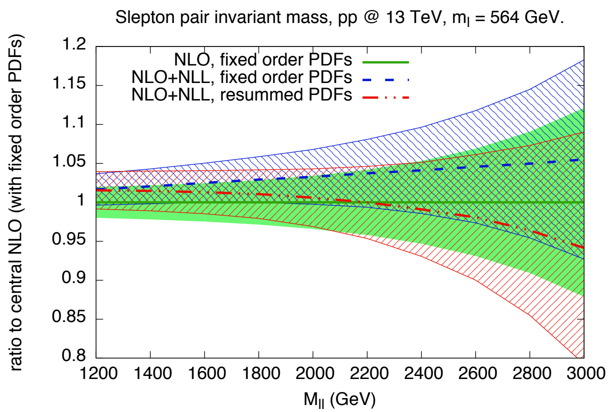
GLUON: NLO VS NLL



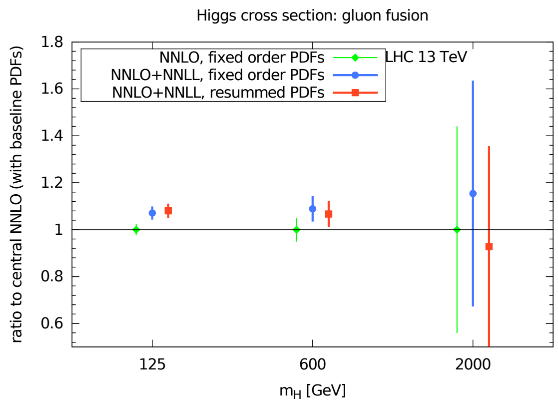
GLUON: NNLO VS NNLL



SLEPTON PAIR PRODUCTION

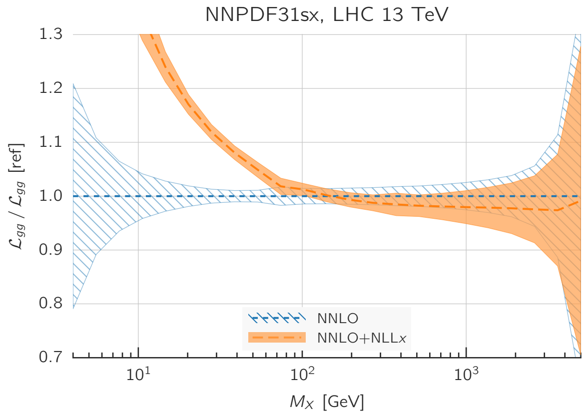


HIGGS IN GLUON FUSION VS m_H



PDFs with high energy (small x) resummation

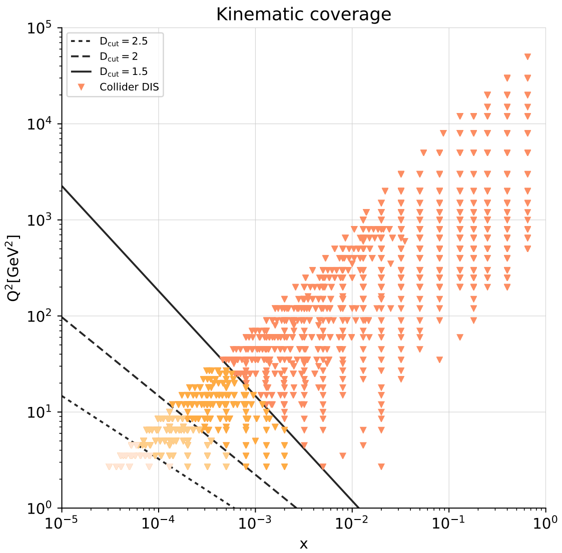
Gluon luminosity



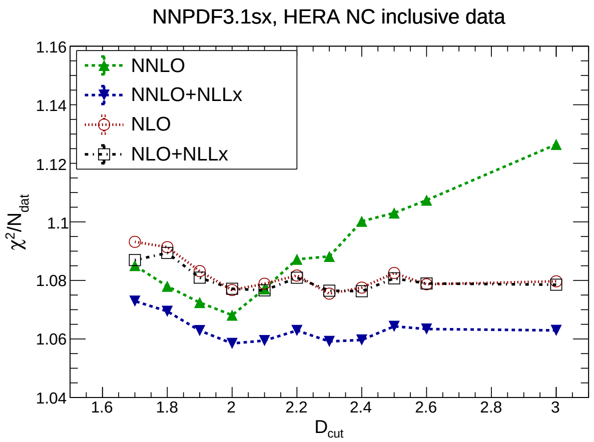
- First set: `NNPDF3.0sx`
- High energy resummation included in GLAP evolution& for DIS, effects
- Stabilizes perturbative expansion
- Large effects for future colliders, or light final states (b production at LHC)

(Ball et al., 2017)

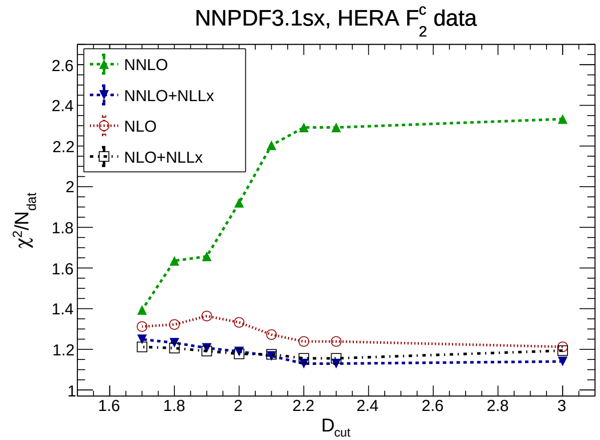
Kinematic cuts



Inclusive F_2 fit quality



Charm F_2^c fit quality

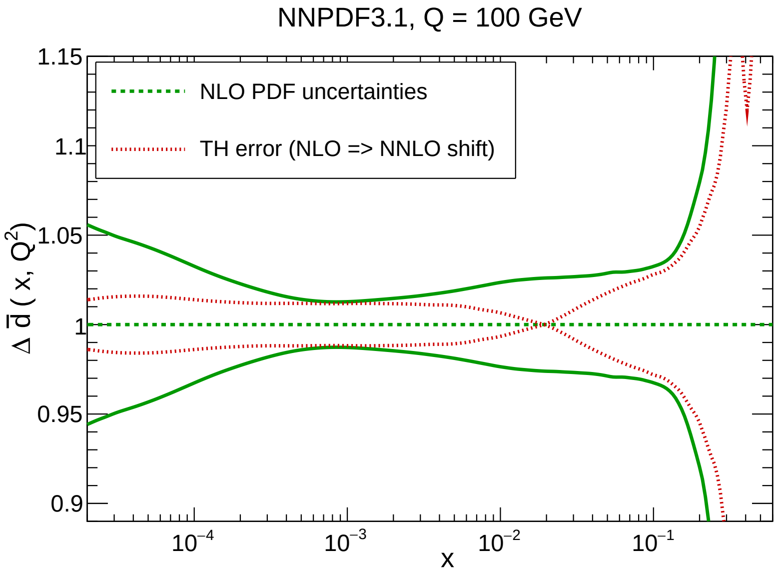


THEORY UNCERTAINTIES
THE MISSING HIGHER ORDER UNCERTAINTY

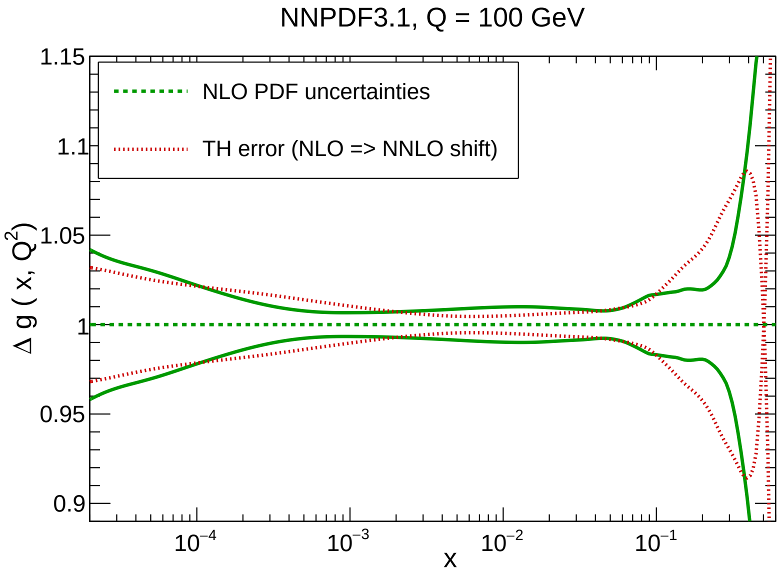
- DOMINANT THEORY UNCERTAINTY ON QCD PREDICTIONS $\Rightarrow$ MHOU (SCALE)
- NOT INCLUDED IN PDF UNCERTAINTY
- HOW LARGE IS IT?
  $\Rightarrow$ AT NLO, CAN CHECK NLO-NNLO PDF SHIFT

NLO-NNLO SHIFT VS. NLO PDF UNCERTAINTY (NNPDF3.1)

ANTIDOWN



GLUON



- TODAY: NLO PDF & MHOU UNCERTAINTIES COMPARABLE
- NEAR FUTURE: WORRY ABOUT NNLO MHOU!

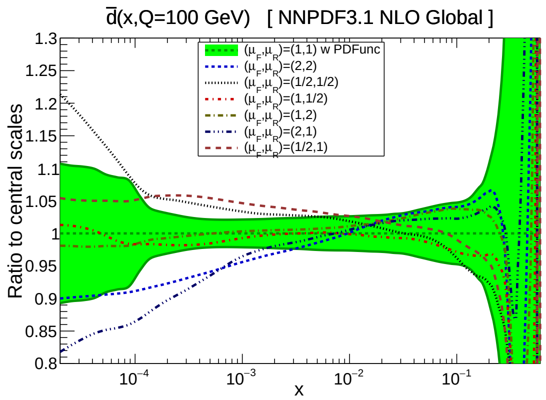
SCALE VARIATION IN PDF FITTING

SIMPLE ESTIMATE FOR PDF MHOU

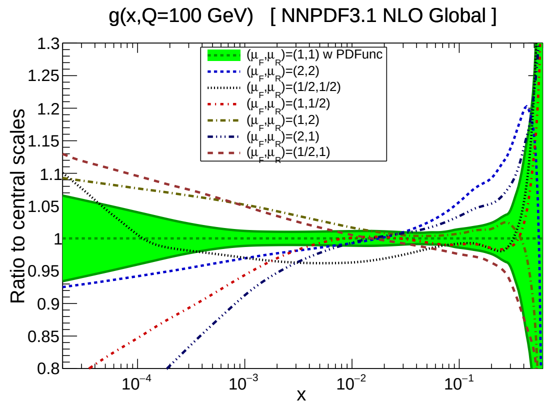
- PERFORM FIT WITH VARIOUS SCALE CHOICES
- TAKE ENVELOPE OF RESULTS
- COMPARE TO NLO-NNLO SHIFT

FITS WITH SCALE VARIATION

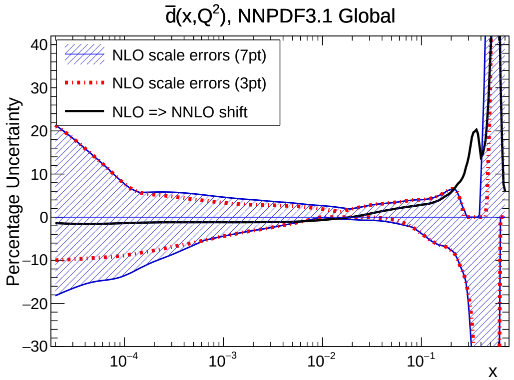
ANTIDOWN

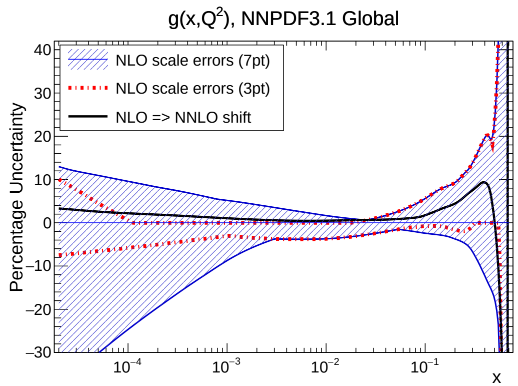


GLUON



SCALE UNCERTAINTY VS NLO-NNLO SHIFT





- REASONABLE ASSESSMENT OF FIRST MHO
- BUT ENVELOPE HAS NO STATISTICAL INTERPRETATION

CONSISTENT INCLUSION
MHOU IN PDF ERRORS

MHOU AS A FIT UNCERTAINTY

- PDFs ARE DETERMINED BY MAXIMIZING THE LIKELIHOOD

$$P = N \exp - \left(\frac{d-t}{2\sigma^2_{exp}} \right)$$

d, t ARE REALLY VECTORS AND $1/\sigma^2$ THE INVERSE COVARIANCE MATRIX

- EXPRESSION COMPLETELY SYMMETRIC IN d AND t
- CAN ENDOW THEORY PREDICTION WITH UNCERTAINTY, & FOLD INTO COVARIANCE MATRX

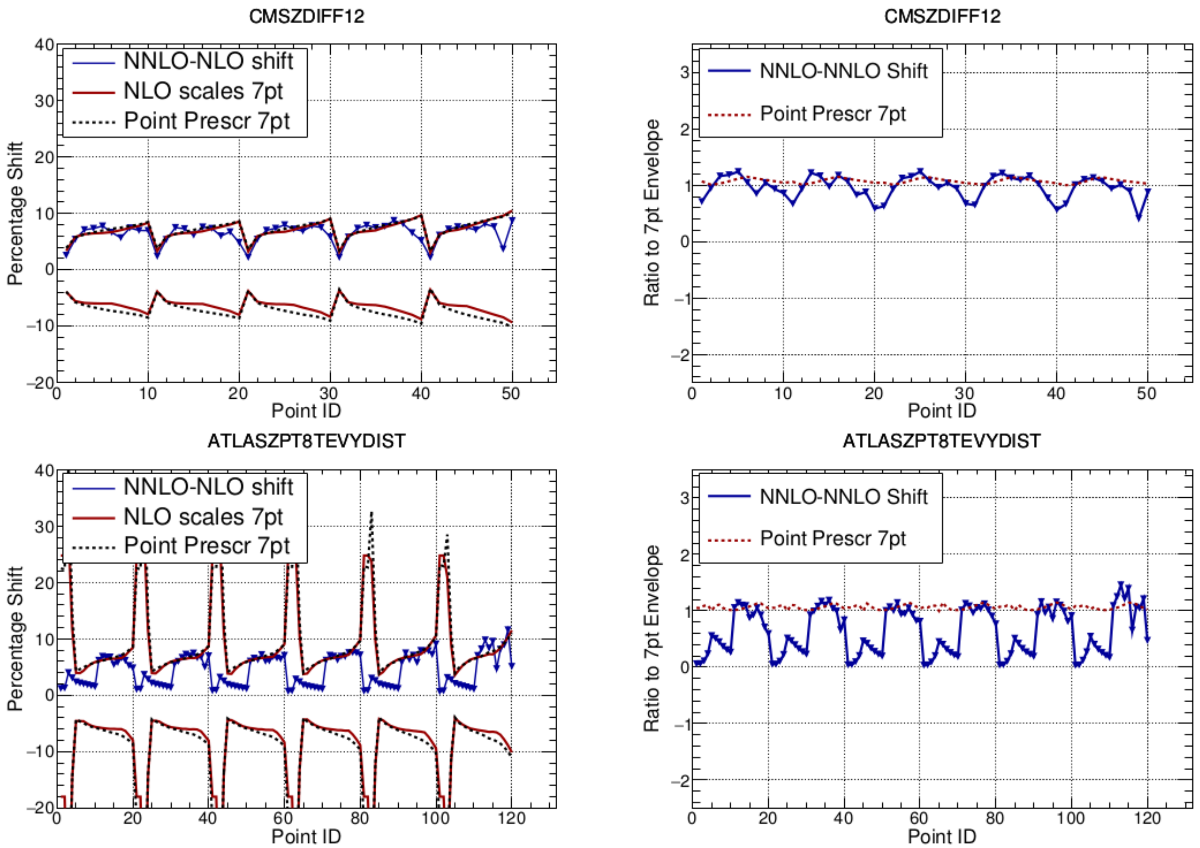
THE THEORY COVARIANCE MATRIX

- PERFORM CALCULATION WITH VARIOUS SCALE CHOICES
- USE DISTRIBUTION OF RESULTS TO COMPUTE THEORY COVARIANCE MATRIX
- PERFORM FIT WITH EXTRA CONTRIBUTION TO COVARIANCE MATRIX
  $\Rightarrow$ MHOU INCLUDED IN PDF UNCERTAINTY

PRESCRIPTIONS & VALIDATION

- DEFINE “POINT PRESCRIPTION:
  E.G. SUM OVER 6 SCALE CHOICES, WITH CORRELATED FACTORIZATION & UNCORRELATED RENORMALIZATION (“7 POINT”)
- CHOOSE CORRELATION PATTERN (“FAMILY OF EXPERIMENTS”)
- COMPARE TO SHIFTS

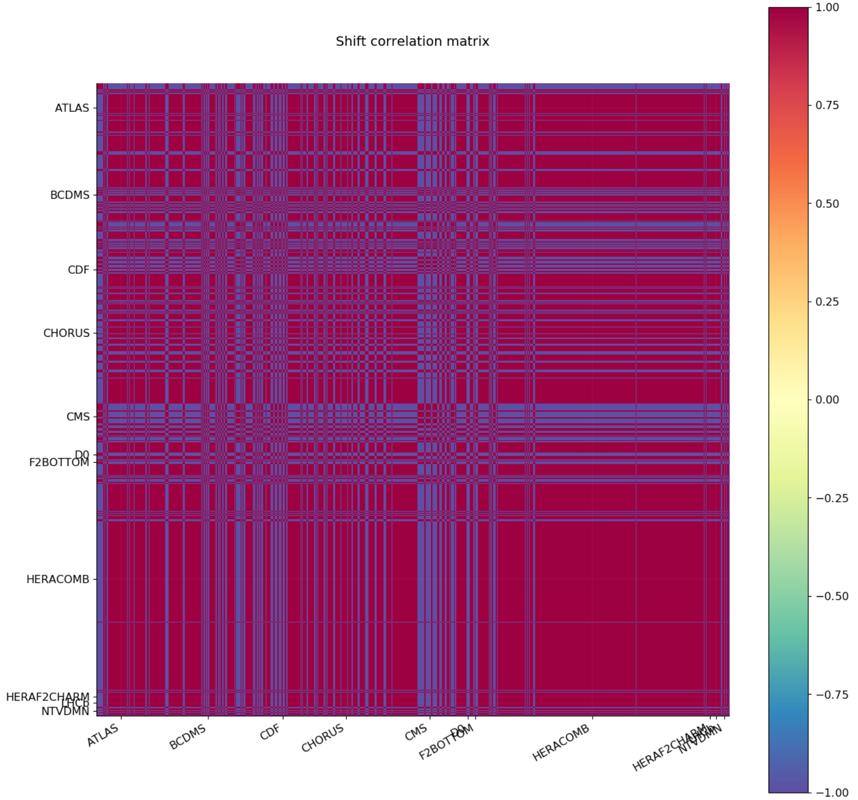
VALIDATION: DIAGONAL UNCERTAINTIES

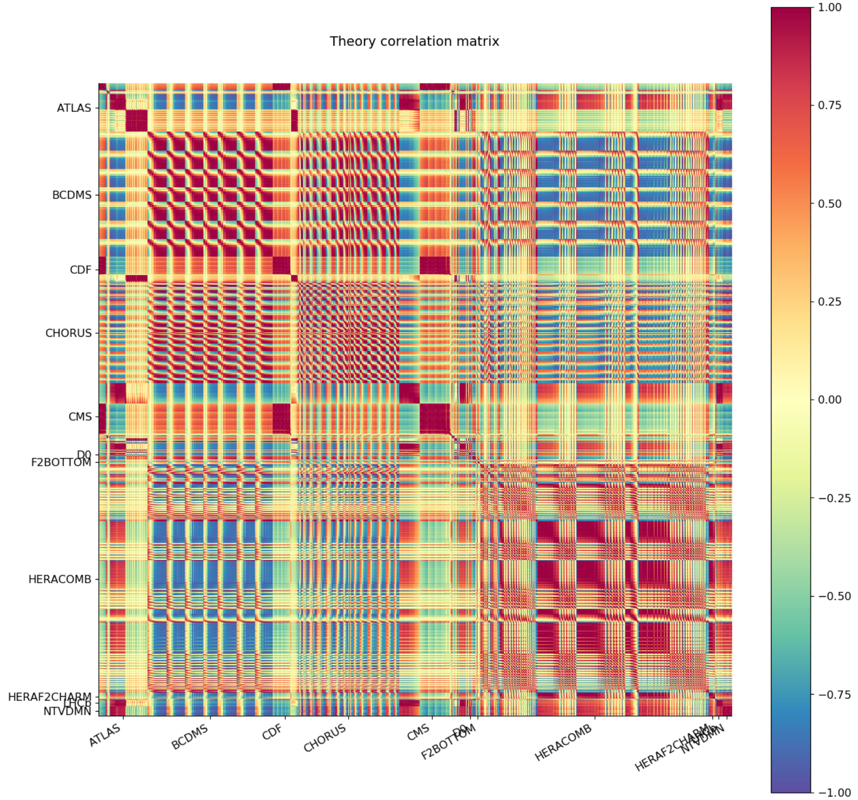


PRESCRIPTIONS & VALIDATION

- DEFINE "POINT PRESCRIPTION:
  E.G. SUM OVER 6 SCALE CHOICES, WITH CORRELATED FACTORIZATION & UNCORRELATED RENORMALIZATION ("7 POINT")
- CHOOSE CORRELATION PATTERN ("FAMILY OF EXPERIMENTS")
- COMPARE TO SHIFTS

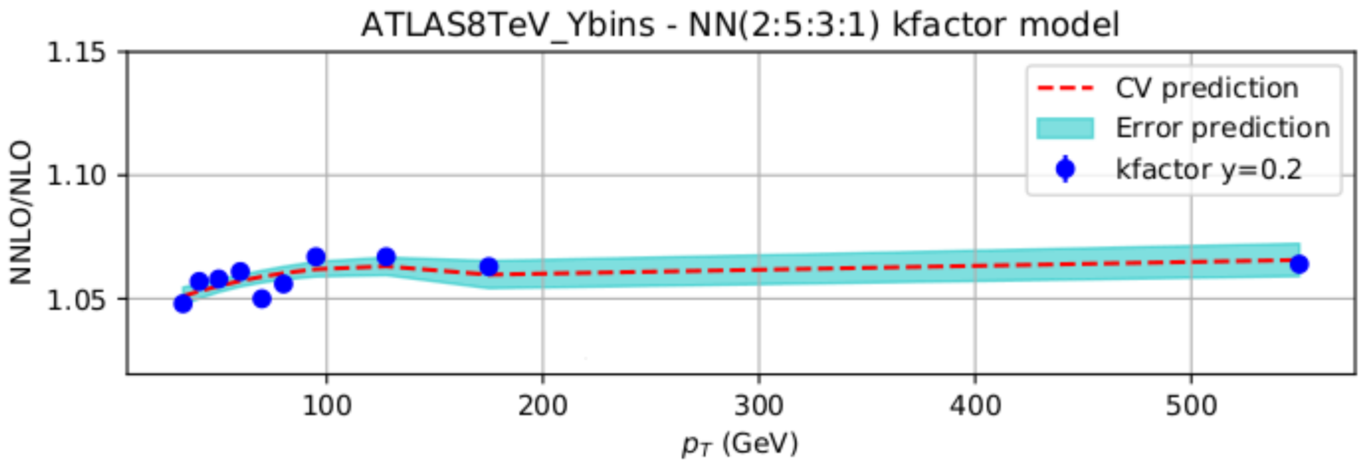
VALIDATION: CORRELATIONS





MORE THEORY UNCERTAINTIES

EXAMPLE: ATLAS 7 TEV p_T DISTRIBUTION
THE NNLO/NLO K-FACTOR



(Boughezal, Liu, Petriello, 2017)

- NUCLEAR CORRECTIONS
- NUMERICAL UNCERTAINTIES
- . . .

. . . CAN BE TREATED SIMILARLY

ELECTROWEAK CORRECTIONS

PAST (NNPDF3.0)

EW CORRECTIONS INCLUDED WHEN LARGE FOR SPECIFIC PROCESSES

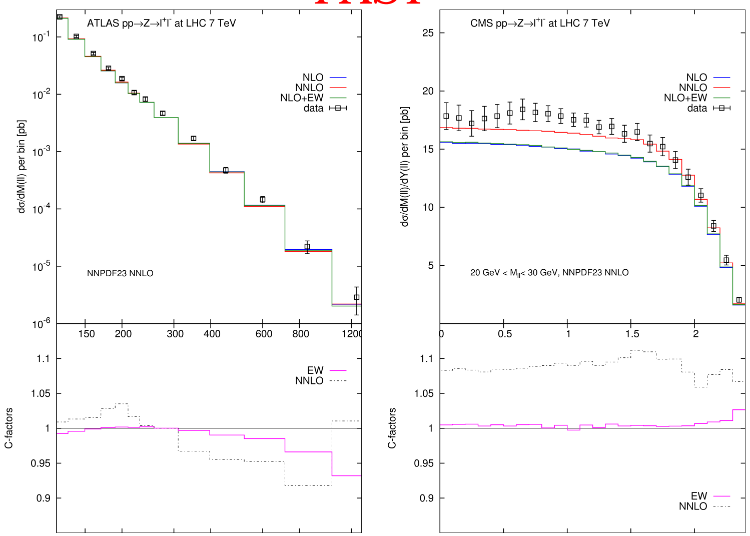
PRESENT (NNPDF3.1)

ALL DATA WITH LARGE EW CORRECTIONS REMOVED FROM DATASER

FUTURE (NNPDF4.0)

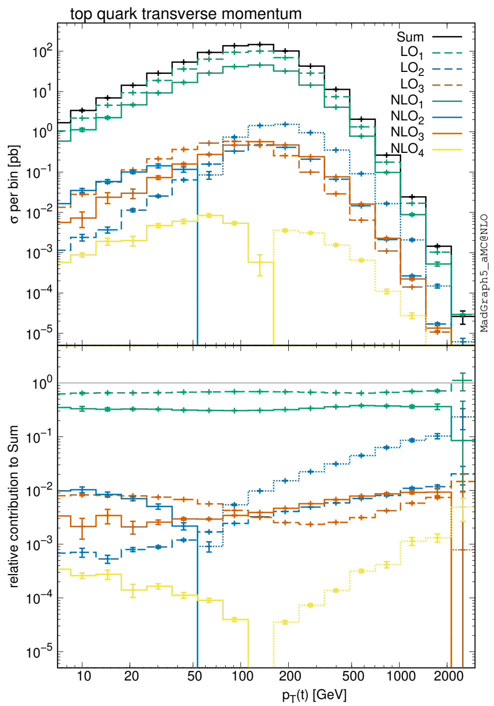
EW INCLUDED SYSTEMATICALLY ⇒ `MadGraph5_aMC@NLO` USED FOR ALL PROCESSES

PAST



NNPDF3.0

FUTURE



(Frederix, Frixione, Hirschi, Pagani, Shao, Zaro, 2017)

WHAT I TALKED ABOUT
AND WHEN

- NEW COMBINED PDF SETS PDF4LHC20?
  - RELIABLE EXPERIMENTAL COVARIANCE MATRICES ⇒ LHC RUN II
  - WIDEST DATASET ⇒ LHC RUN II-RUN III
- RESUMMED PDFS, N^{3}LO PDFS, EW CORRECTIONS ⇒ LHC RUN II-RUN III
  - AUTOMATIZATION OF HO CORRECTIONS
  - PUBLIC CODES
- THEORY UNCERTAINTIES
  - THEORY PREDICTIONS WITH SUB-PERCENT NUMERICAL ACCURACY ⇒ NOW – LHC RUN II
  - FITTED HEAVY QUARK PDFS ⇒ NNPDF3.1 (2018), PDF4LHC20
  - PDFS WITH THEORY UNCERTAINTY ⇒ NNPDF4.0 (2019)
- CONSENSUS α_s WITH SUB-PERCENT ACCURACY ⇒ HL-LHC

RUN II: 2015-2018; RUN III: 2020-2023; HL-LHC: 2026-. . .

WHAT I DID NOT TALK ABOUT
AND WHY

- HIGHER TWISTS, NUCLEAR CORRECTIONS $\Rightarrow$ HISTORY!
  - FIXED-TARGET DATA OBSOLETE
  - COLLIDER-ONLY PDFs
- SYSTEMATIC CLOSURE TESTING $\Rightarrow$ KNOWN SINCE NNPDF3.0
  - TUNING OF FITTING METHODOLOGY
  - VALIDATION OF PDF UNCERTAINTIES
  - STATISTICAL INTERPRETATION OF PDF CONFIDENCE LEVELS
- PDFs AT SUBPERCENT ACCURACY
  - OPTIMIZED MINIMIZATION $\Rightarrow$ N^{3}PDF
  - MACHINE LEARNING PDF MINIMIZATION $\Rightarrow$ N^{3}PDF
  - THEORY UNCERTAINTIES FROM RESUMMATION $\Rightarrow$ N^{3}PDF
  - PS-PDFs, MC-P DFs $\Rightarrow$ N^{3}PDF

EXTRAS

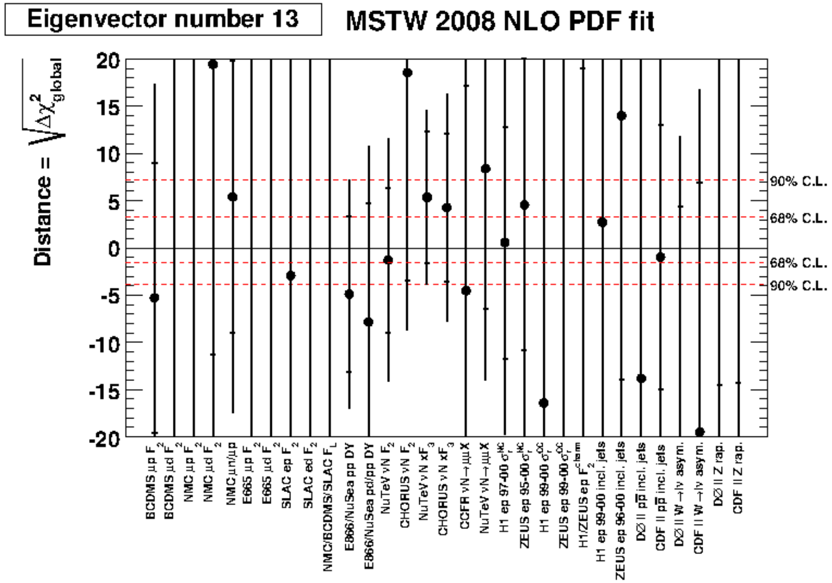
CAN WE TRUST PDF UNCERTAINTIES?

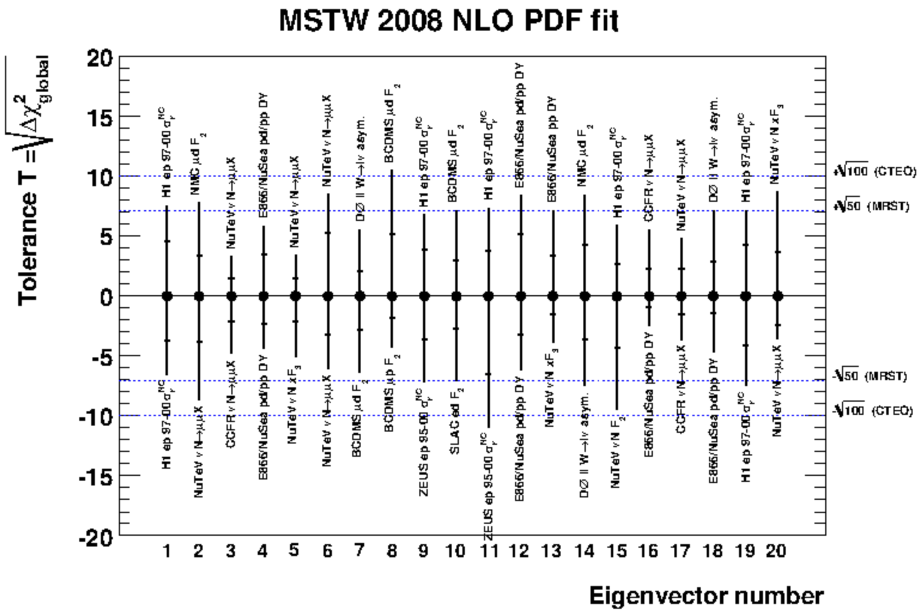
- "PDF" UNCERTAINTIES REFLECT UNCERTAINTY FROM THE DATA & METHODOLOGY (NOT THEORY)
- UNCERTAINTIES ON GLOBAL FITS ⇒ SIMILAR SIZE DESPITE DIFFERENT PROCEDURES
- DUE TO UNCERTAINTY TUNING

TOLERANCE (MMHT-CT)

GLOBAL MSTW TOLERANCE

MSTW TOLERANCE PLOT FOR 13TH EIGENVEC.



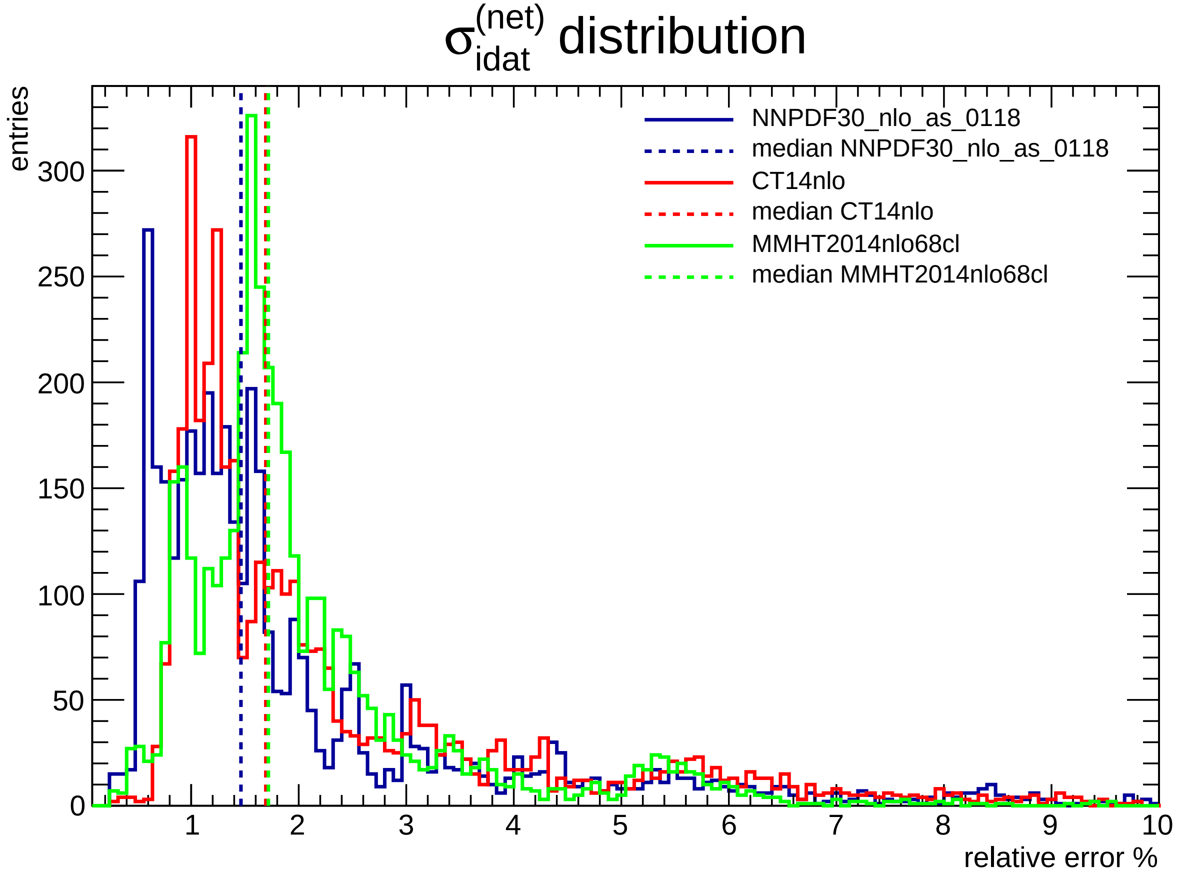


- (MSTW/MMHT) FOR EACH EIGENVECTOR IN PARAMETER SPACE DETERMINE CONFIDENCE LIMIT FOR THE DISTRIBUTION OF BEST-FITS OF EACH EXPERIMENT
- RESCALE $\Delta\chi^2 = T$ INTERVAL SUCH THAT CORRECT CONFIDENCE INTERVALS ARE REPRODUCED
- WHY DO WE NEED TOLERANCE?
- DO WE UNDERSTAND PDF UNCERTAINTIES?

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



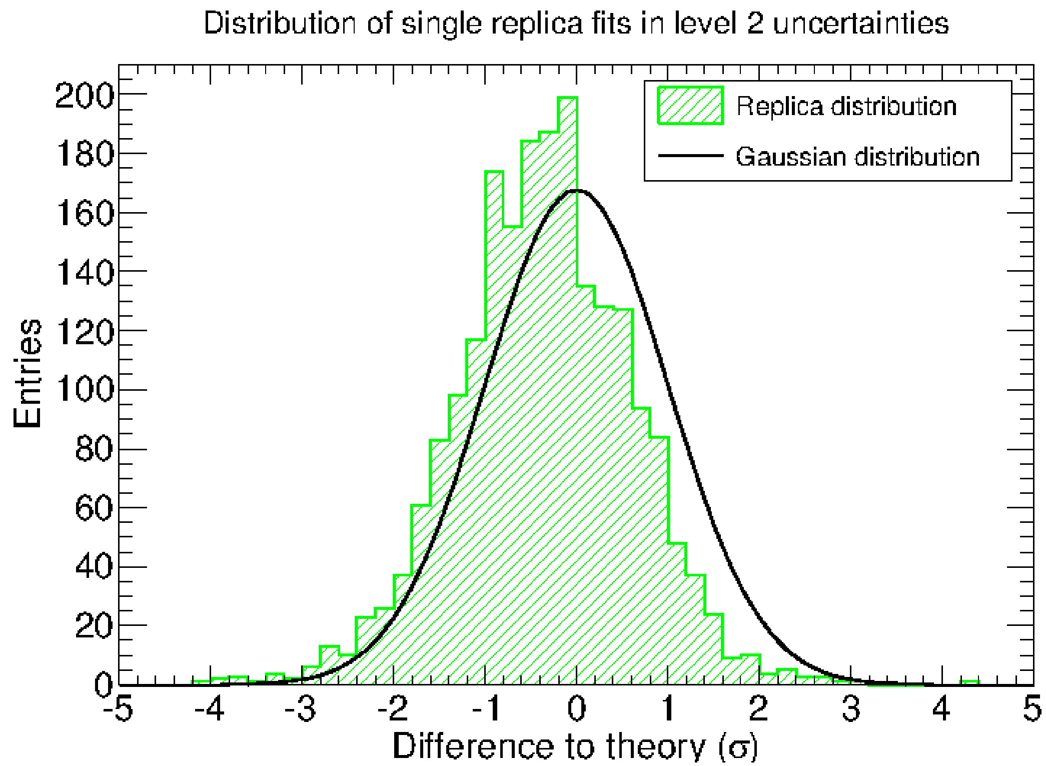
- MEDIAN SIMILAR
- DISTRIBUTION VERY DIFFERENT!
- NNPDF: SMALLER MODE, BUT FAT TAIL ⇔ GREATER FLEXIBILITY

ARE PDF UNCERTAINTIES FAITHFUL?
CLOSURE TESTING
BASIC IDEA

- ASSUME PDFs KNOWN: GENERATE FAKE EXPERIMENTAL DATA
- CAN DECIDE DATA UNCERTAINTY (ZERO, OR AS IN REAL DATA, OR . . .)
- FIT PDFs TO FAKE DATA:
- TEST WHETHER PDF UNCERTAINTY FAITHFULLY REFLECTS DATA UNCERTAINTY: E.G. TRUE VALUE WITHIN ONE-SIGMA 68% OF TIMES

CLOSURE-TESTING THE PDF UNCERTAINTIES
RESULTS

UNCERTAINTIES: DISTRIBUTION OF DEVIATIONS BETWEEN FITTED AND "TRUE" PDFs, SAMPLED AT 20 POINTS BETWEEN 10^{-5} AND 1



FIND 0.699% FOR ONE-SIGMA, 0.948% FOR TWO-SIGMA C.L.

ARE PDF UNCERTAINTIES OPTIMAL?
THE $\Delta\chi^2$ PROBLEM

- TOLERANCE MIGHT COMPENSATE FOR MISSING FUNCTIONAL UNCERTAINTY
- BUT WHAT IS $\Delta\chi^2$ FOR AN NNPDF FIT?
- CAN ANSWER USING HESSIAN CONVERSION! $\Delta\chi^2 = 16 \pm 15$
  - NON-PARABOLIC BEHAVIOUR NEAR MINIMUM ON SCALE OF UNCERTAINTIES?
  - INEFFICIENCY OF THE MINIMIZATION PROCEDURE?

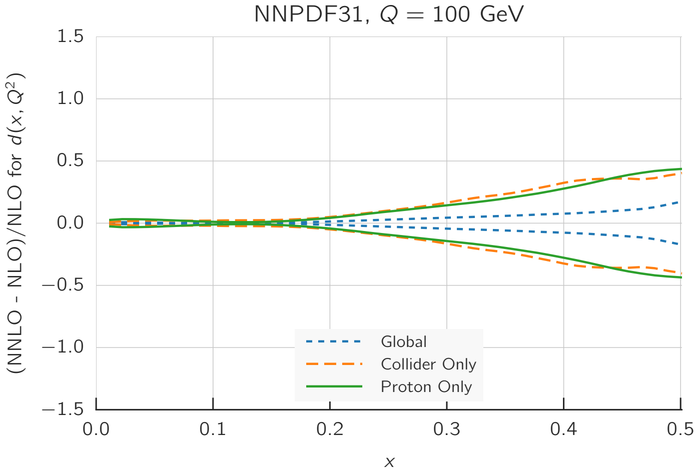
DEPENDENCE ON THE DATASET

PERTURBATIVE STABILITY

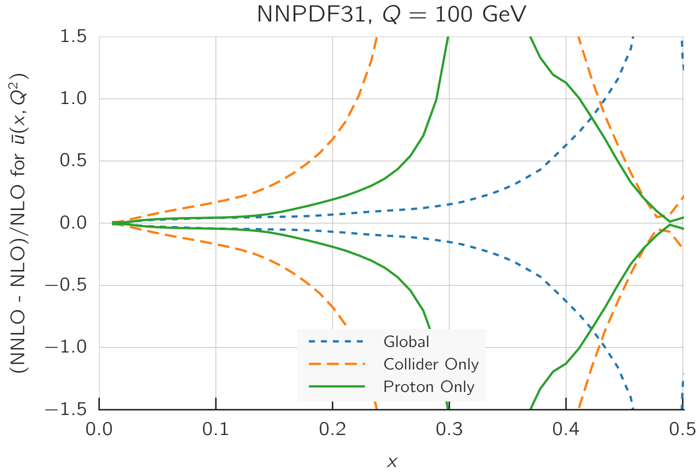
NNLO-NLO SHIFTS

GLOBAL VS. COLLIDER ONLY VS. PROTON ONLY PDFS

DOWN



ANTIUP



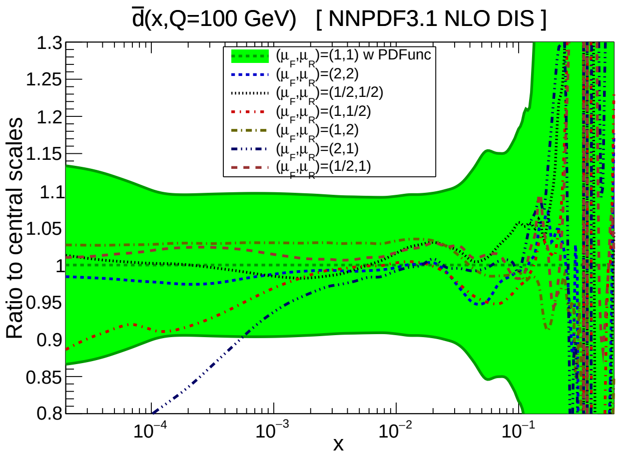
- MHOU DEPENDS ON DATASET
- NLO-NNLO SHIFTS GENERALLY SMALLER WITH LARGER DATASET
- $\Rightarrow$ CORRELATION PATTERN BETWEEN MHO TERMS

DEPENDENCE ON THE DATASET

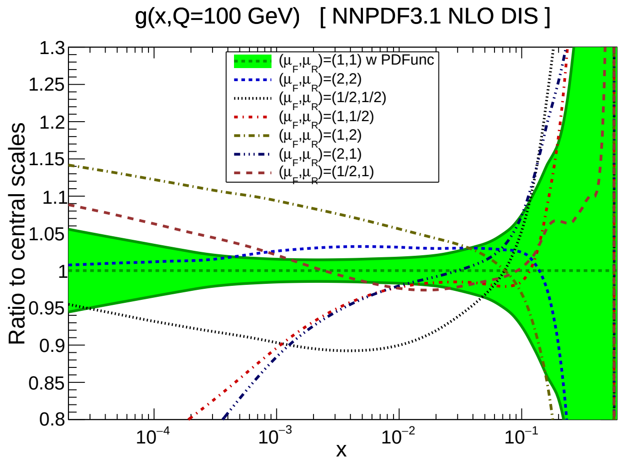
SCALE VARIATION

DIS ONLY

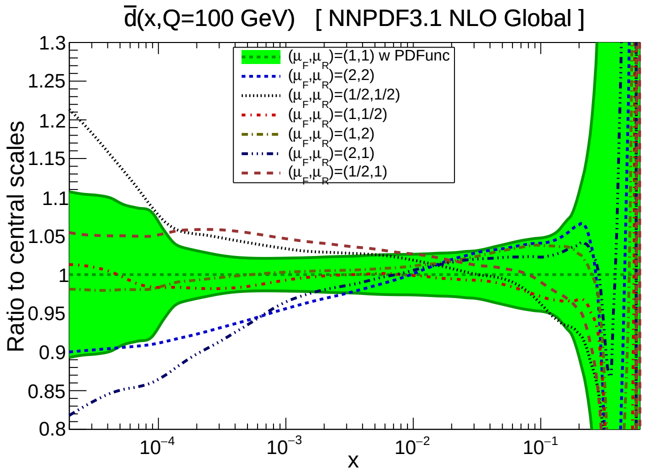
ANTIDOWN

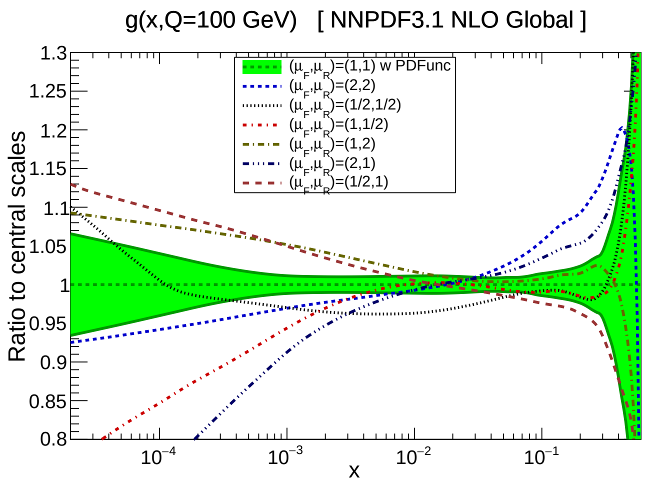


GLUON



GLOBAL





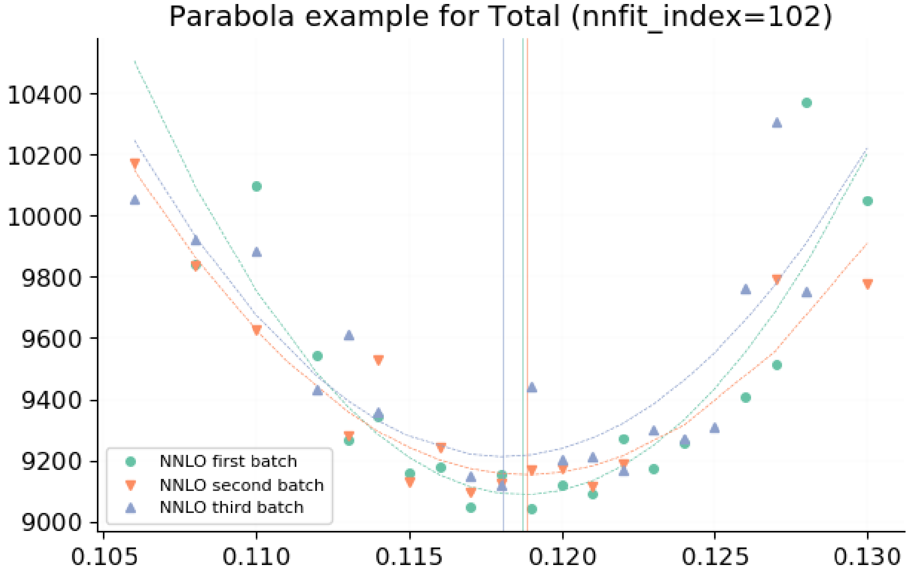
- SCALE DEPENDENCE IS (OF COURSE!) DATASET DEPENDENT
- BOTH SIZE AND PATTERN OF BEST-FIT RESULTS DEPEND ON PROCESS
- WHAT IS THE CORRELATION BETWEEN SCALE DEPENDENCES?
- ENVELOPE??

THE METHODOLOGY AND ITS UNCERTAINTIES

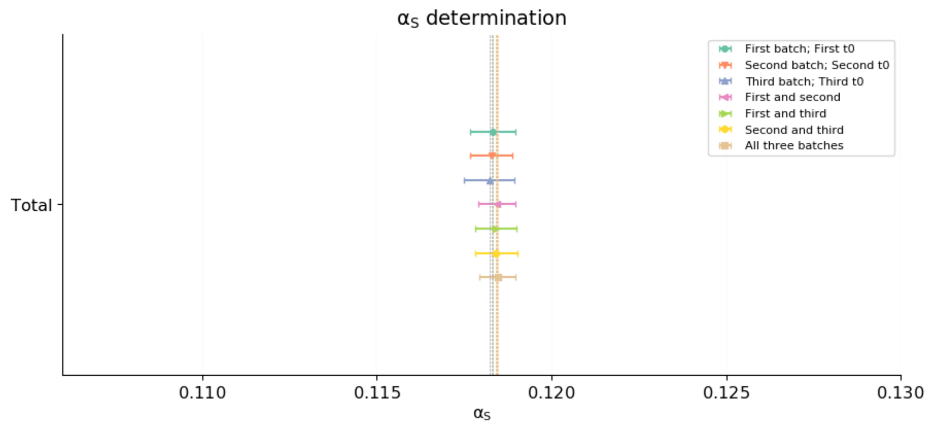
- WHY IS THE EXPERIMENTAL UNCERTAINTY SO SMALL?
- BECAUSE THERE IS A LOT OF INFORMATION IN C-REPLICAS

BATCH MIMIMIZATION

DIFFERENT BATCHES FOR ONE REPLICA



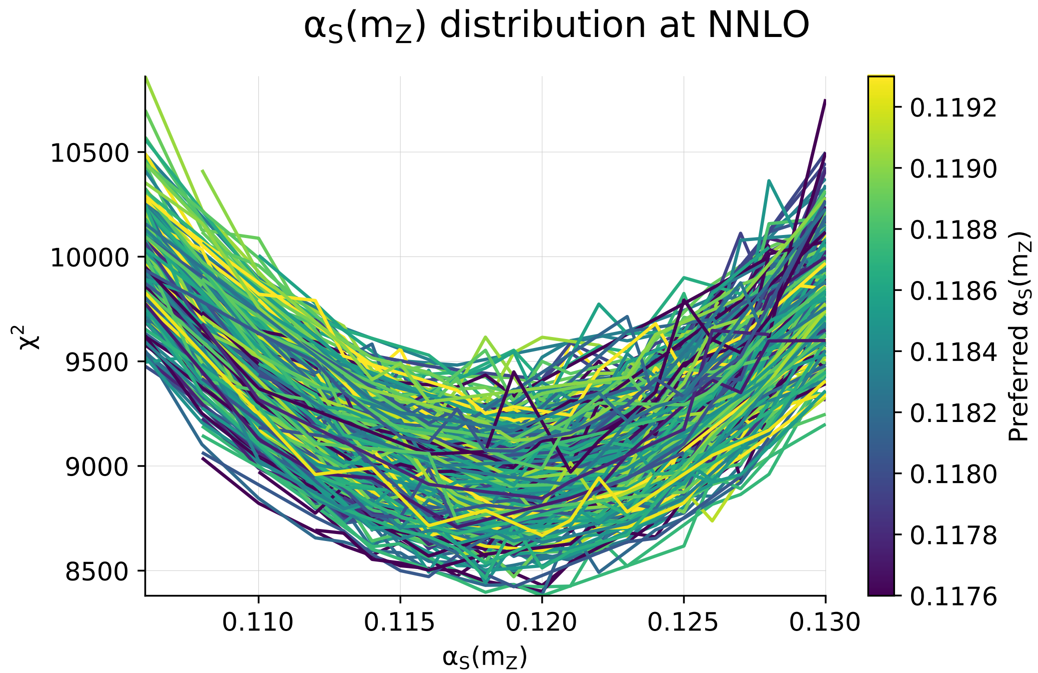
DEPENDENCE OF α_s ON BATCH CHOICE



- EACH REPLICA FOR EACH α_s FITTED THREE TIMES (BATCHES)
- BEST FIT CHOSEN
- SINGLE, DOUBLE & TRIPLE BATCH COMPARED $\Rightarrow$
  LITTLE DEPENDENCE OF CENTRAL VALUE, IMPROVEMENT OF UNCERTAINTIES
  SINGLE BATCH $\sigma^{\exp} = 0.0006 - 0.0007$; THREE BATCHES $\sigma^{\exp} = 0.0005$

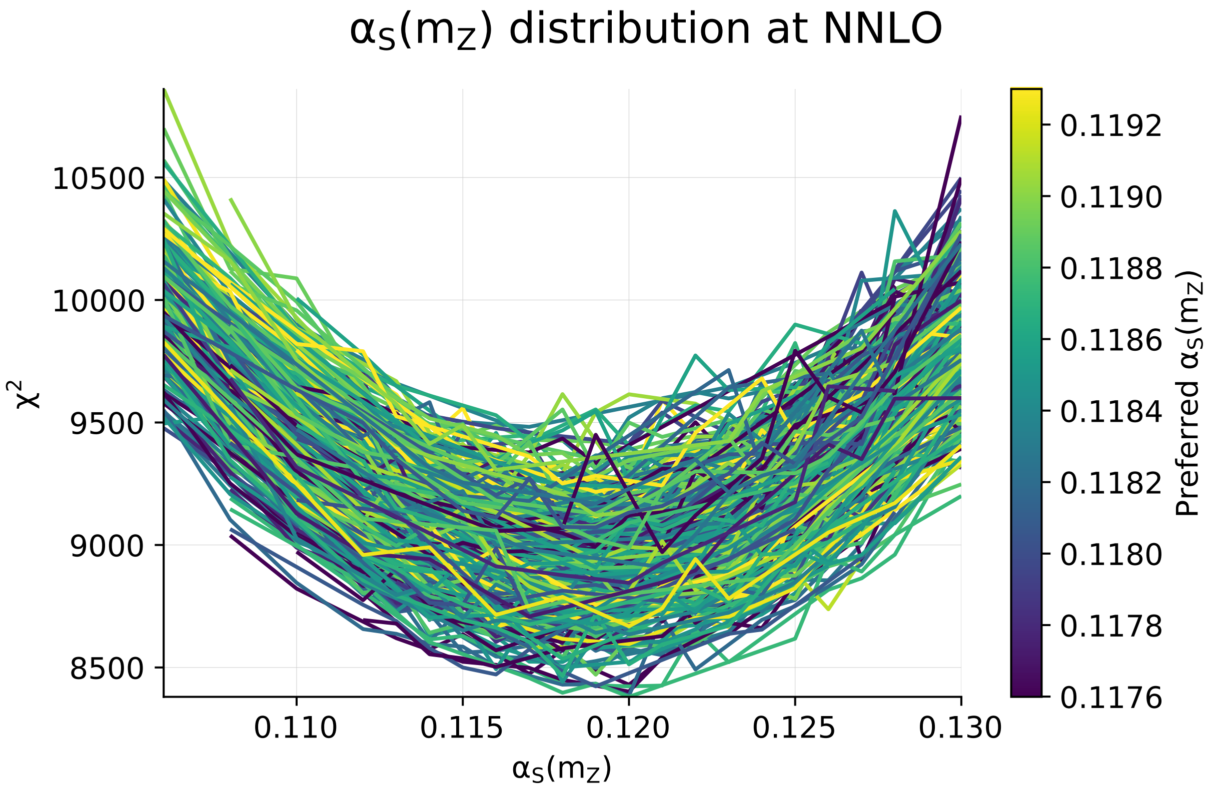
METHODOLOGICAL UNCERTAINTIES
FINITE-SIZE OF REPLICA SAMPLE

- ESTIMATED BY BOOTSTRAPPING: $\Delta^{\text{finite size}} = 0.00003$ (0.03%)
- DEPENDS ON REPLICA SELECTION $\Leftrightarrow$ ONLY
  INCLUDE A C-REPLICA IF FIT CONVERGED FOR AT LEAST $N_{\min}$ α_s VALUES
- MORE RESTRICTIVE SELECTION $\Rightarrow$ SMALLER σ^{exp}, LARGER $\Delta^{\text{finite size}} = 0.00009$
- WITH MOST RESTRICTIVE SELECTION ($N_{\min} = 18$, 12 SURVIVING REPLICAS)
  $\Delta^{\text{finite size}} = 0.00009$ (0.08%), BUT $\sigma^{\text{exp}} = 0.00031$ (0.3%)
- NO DEPENDENCE OF α_s VALUE ON CURVE SELECTION
- DEFAULT: CONSERVATIVE CHOICE $N_{\min} = 6$, 379 SURVIVING REPLICAS)
  $\Delta^{\text{finite size}} = 0.00003$ (0.03%), BUT $\sigma^{\text{exp}} = 0.00052$ (0.5%)
  MINIMIZES UNCERTAINTY ON UNCERTAINTY

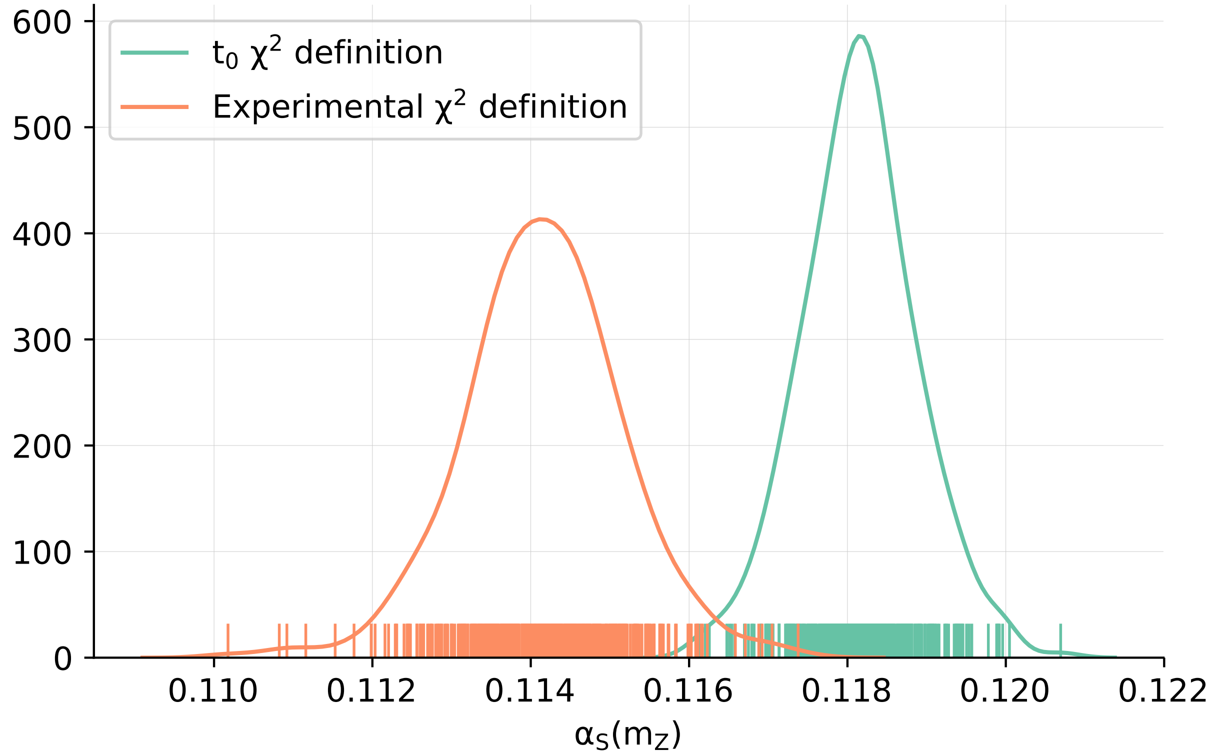


METHODOLOGICAL UNCERTAINTIES
PARABOLIC FITTING

- FIT $\exp \alpha_s$ OR $\ln(\alpha_s + 1) \Rightarrow$ NO CHANGE
- REMOVE OUTER VALUES OF α_s, SYMMETRICALLY OR ASYMMETRICALLY $\Rightarrow$ LARGEST SHIFT $\Delta^{\text{parabolic}} = 0.00010$ (0.08%) $\Rightarrow$ DOMINANT METH. UNCERTAINTY
- NO EVIDENCE THAT CUBIC FIT IS BETTER THAN QUADRATIC



METHODOLOGICAL UNCERTAINTIES
NORMALIZATION UNCERTAINTIES



- MULTIPLICATIVE UNCERTAINTIES (ALL HADRON COLLIDER SYST.)
  $\Rightarrow$ t_0 METHOD, REQUIRES PRIOR FIT
- GROSSLY BIASED IF EXPERIMENTAL COVARIANCE MATRIX USED
- PRIOR FIT TAKEN FROM THREE DIFFERENT BATCHES
  $\Rightarrow$ VERY SMALL SHIFT $\Delta^{t_0} = 0.00004$ (0.03%)

TOTAL METHODOLOGICAL UNCERTAINTY

$$\sigma^{\text{meth}} = 0.00011 \ (0.09\%)$$

MISSING HIGHER ORDER UNCERTAINTIES

$$\alpha_s^{\rm NLO}(M_Z) = 0.11845 \pm 0.00052^{\rm exp}\ (0.4\%)$$
$$\alpha_s^{\rm NNLO}(M_Z) = 0.12067 \pm 0.00065^{\rm exp}\ (0.5\%)$$
$$\Delta\alpha_s^{\rm NLO-NNLO} = 0.0022$$

- CACCIARI-HOUDEAU: $\alpha_s^{\rm N^kLO} = \sum_{n=0}^{k} c_n [\alpha_s^{\rm true}]^n$;
  BAYESIAN ESTIMATE FOR RANGE OF NEXT MISSING COEFFICIENTS
- DIFFICULTIES:
  - $\alpha_s^{\rm true}$ NOT KNOWN $\Rightarrow$ VARY IN WIDE RANGE
  - $\alpha_s^{\rm LO}$ EFFECTIVELY NOT KNOWN $\Rightarrow$ VARY IN WIDE RANGE
  - IS IT MEANINGFUL FOR A COMBINATION OF PROCESSES?

$$\Delta\alpha_s^{\rm NLO,\,CH} = 0.003$$
$$\Delta\alpha_s^{\rm NNLO,\,CH} = 0.0004$$

NOTE AT NLO AGREES WELL WITH KNOWN NNLO-NLO SHIFT!

TOTAL THEORY UNCERTAINTY

VERY CONSERVATIVE ESTIMATE:

$$\sigma^{\rm th} = 0.0011\ (0.9\%)$$

- HALF THE NLO-NNLO SHIFT
- THREE TIMES THE CH ESTIMATE