DETERMINATION OF α_s

& THE CORRELATED REPLICA METHOD

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN

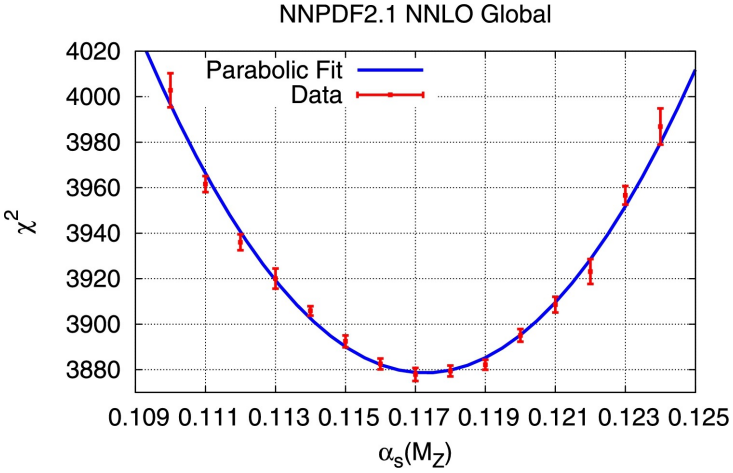
UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

PDF4LHC

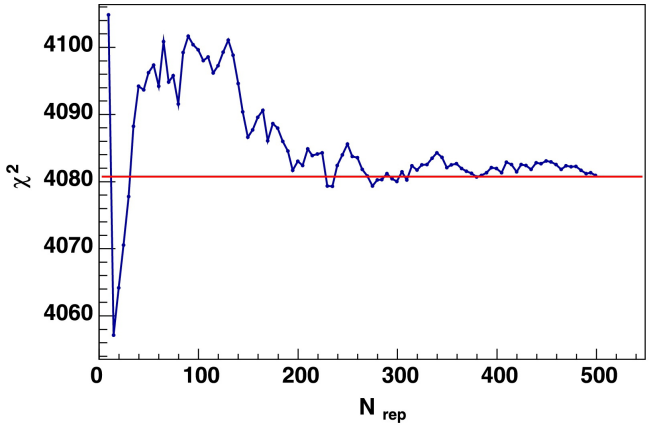
CERN, MARCH 28, 2018

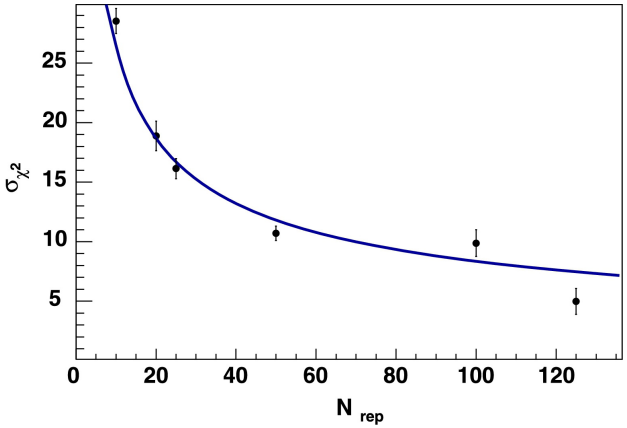
α_s FROM REPLICAS
NNPDF2.1 (2012)

- PERFORM GLOBAL FIT WITH DIFFERENT FIXED VALUES OF α_s
- DETERMINE χ^2 & FIT PARABOLA TO PROFILE
- BEST-FIT FROM MINIMUM UNCERTAINTY FROM $\Delta\chi^2 = 1$



PROBLEM (TRIVIAL)



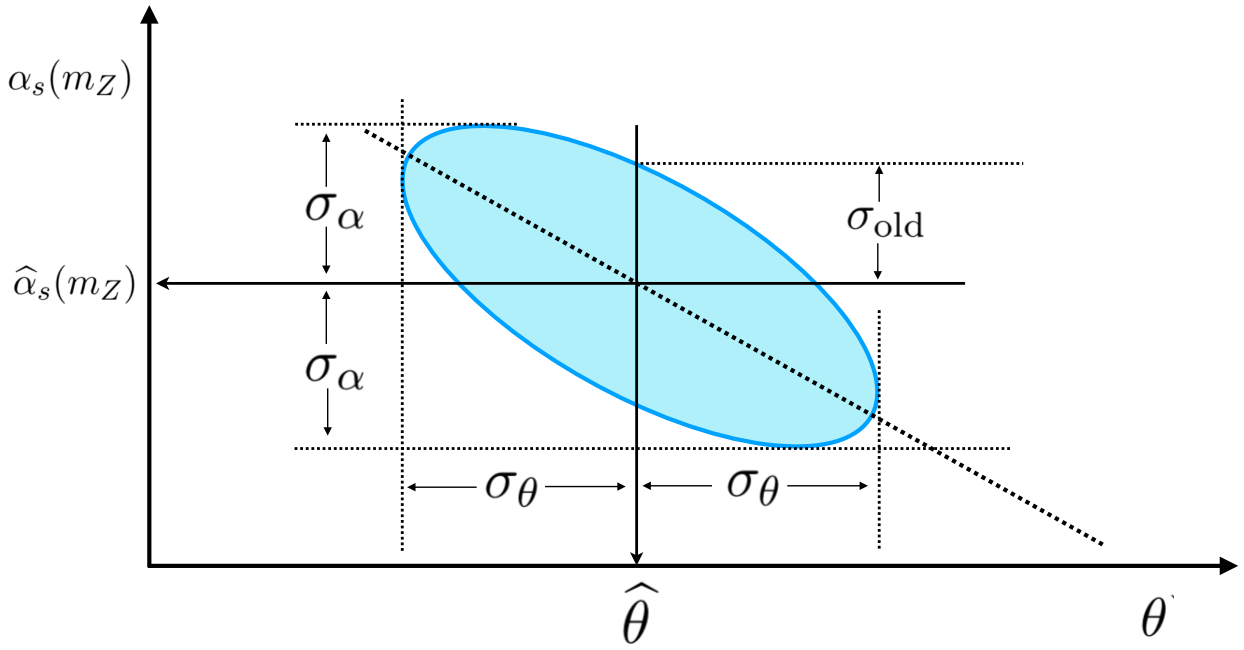


- $\sigma_{\chi^2} = \sqrt{2N_{\rm dat}} \Rightarrow$ NEED LARGE $N_{\rm rep}$ TO SUPPRESS χ^2 FLUCTUATIONS

α_s FROM REPLICAS

NNPDF2.1 (2012)

PROBLEMS (LESS TRIVIAL)



- MINIMUM DETERMINED ALONG THE "BEST PDF" LINE $\Rightarrow \sigma_{\rm old}$
  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)

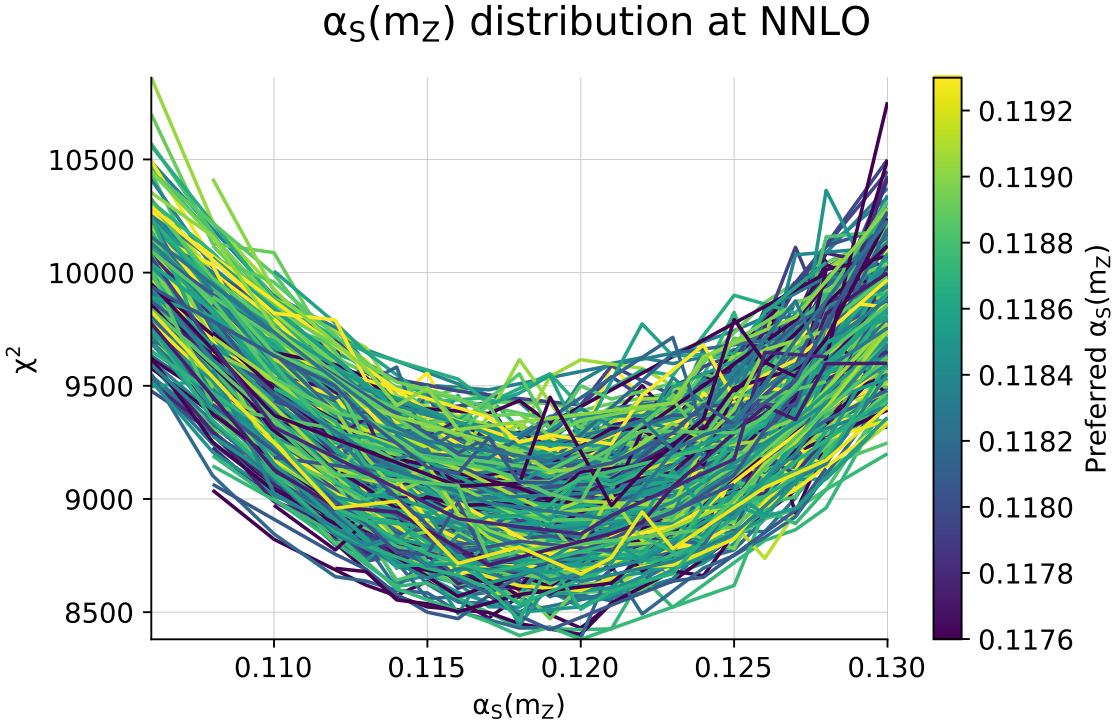
- NNPDF METHOD $\Rightarrow$ EACH PDF REPLICA FITTED BY GA TO DATA REPLICA
- IDEALLY PERFORM GENETIC MINIMIZATION IN (PDF, α_s) SPACE
- PROBLEM THEORY PREDICTION $\Leftrightarrow$ PRECOMPUTED GRIDS
  DEPEND ON α_s $\Rightarrow$ DIFFICULT TO TREAT AS CONTINUOUS PARAMETER

THE CORRELATED REPLICA METHOD

NNPDF3.1 (2018)

- NNPDF METHOD $\Rightarrow$ 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: C-REPLICA
  - EACH C-REPLICA $\Rightarrow$ χ^2 PROFILE $\Rightarrow$ α_s VALUE
  -

ENSEMBLE OF PARABOLAS

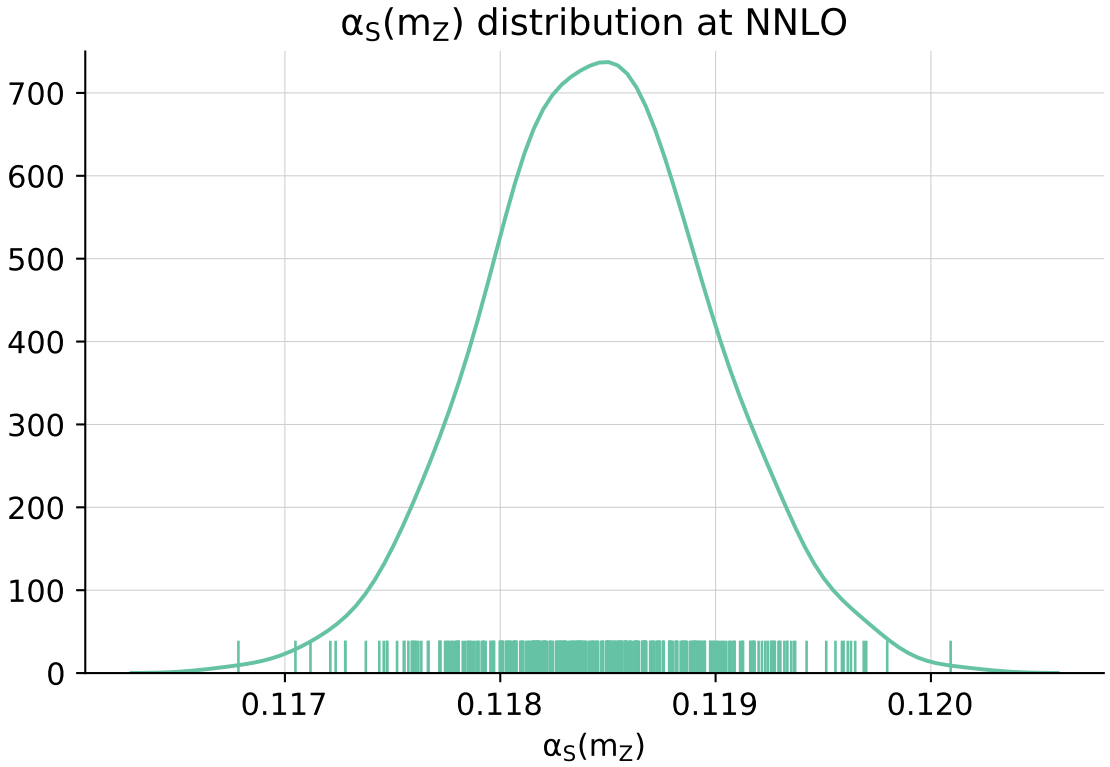


THE CORRELATED REPLICA METHOD

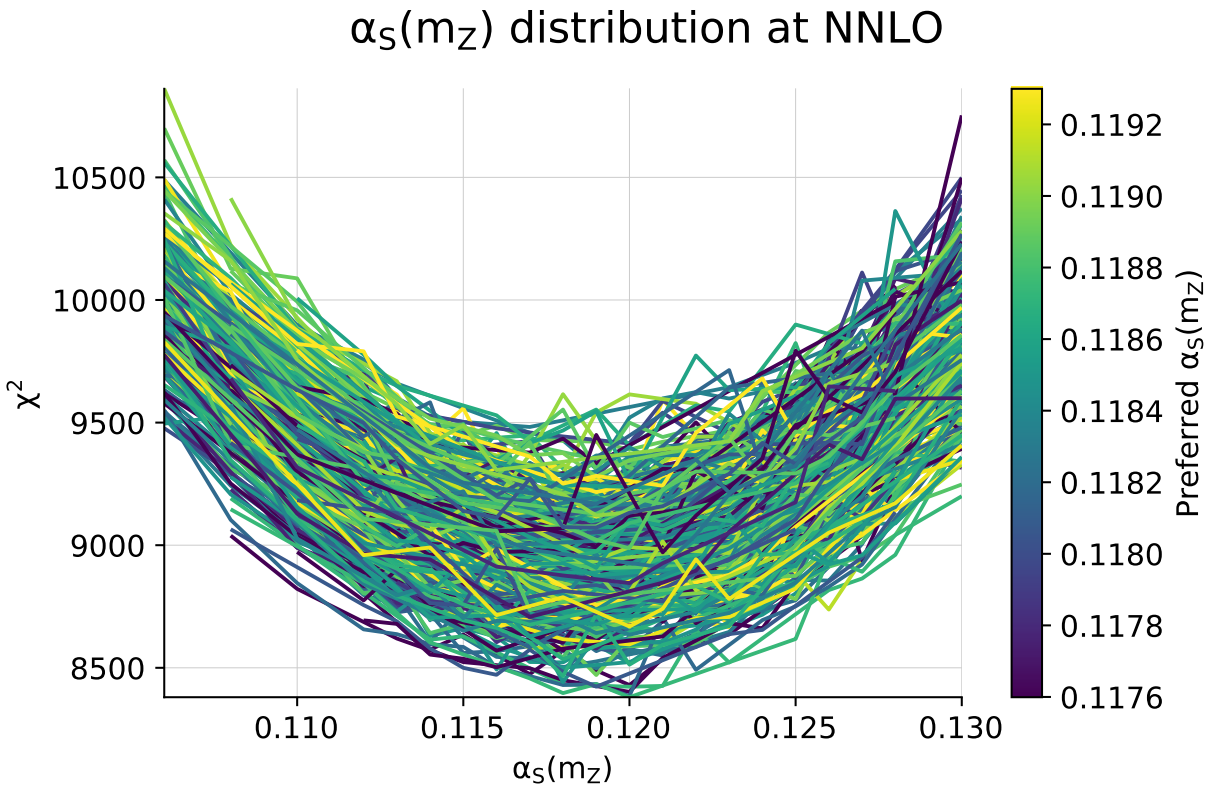
NNPDF3.1 (2018)

- 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

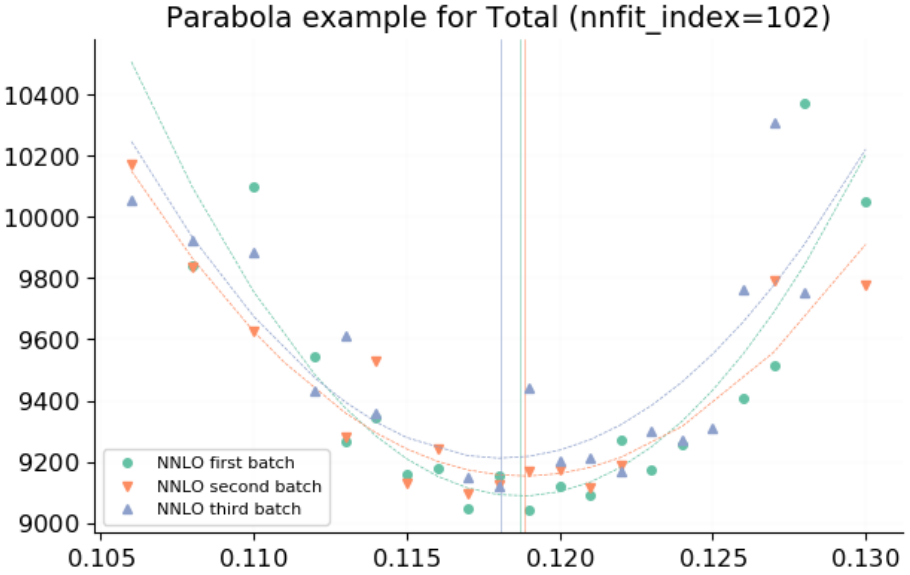
$$\alpha_s^{\rm NNLO}(M_Z) = 0.11845 \pm 0.00052^{\rm exp} \ (0.4\%)$$

THE METHODOLOGY AND ITS UNCERTAINTIES

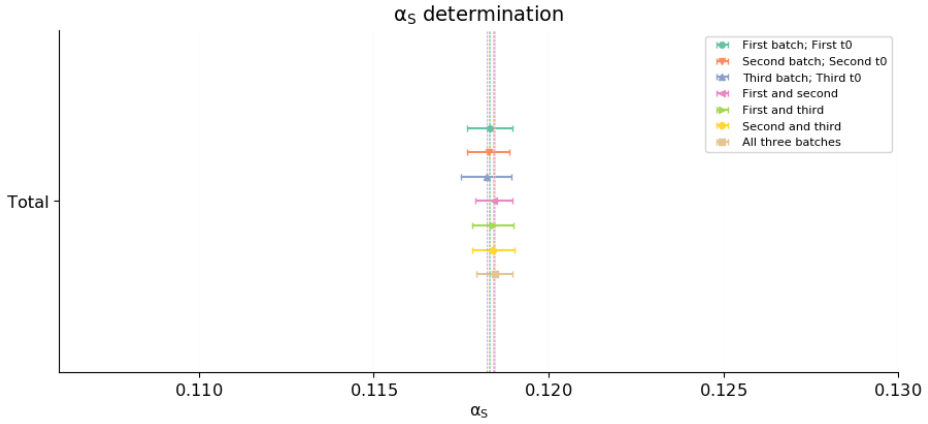
- WHY IS THE EXPERIMENTAL UNCERTAINTY SO SMALL?
- BECAUSE THERE IS A LOT OF INFORMATION IN C-REPLICAS

BATCH MINIMIZATION

DIFFERENT BATCHES FOR ONE REPLICA



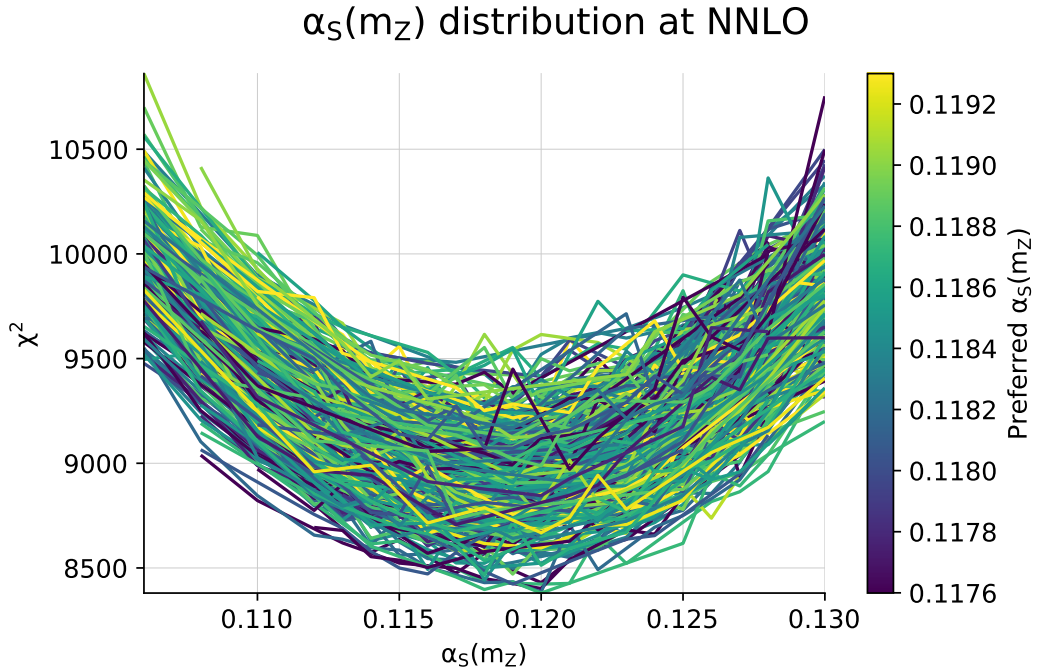
DEPENDENCE OF α_s ON BATCH CHOICE



- 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 $\sigma^{\rm exp} = 0.0006 - 0.0007$; THREE BATCHES $\sigma^{\rm exp} = 0.0005$

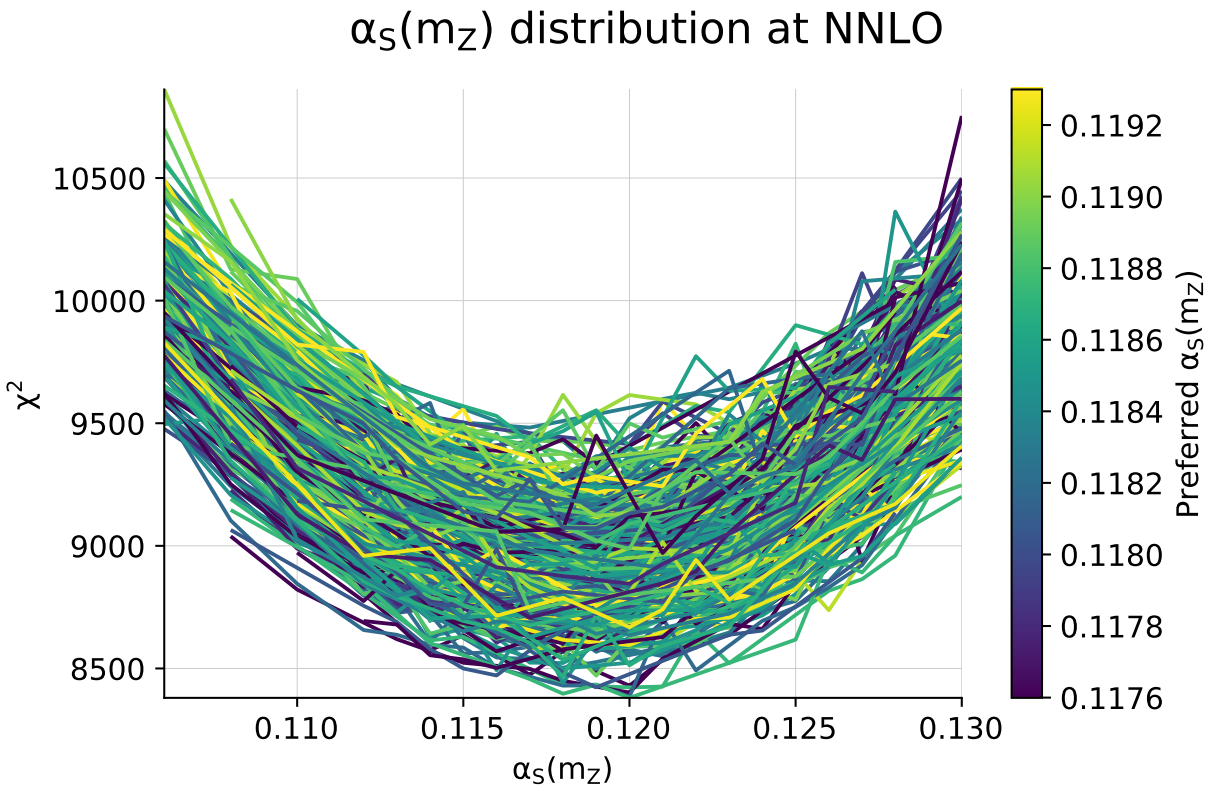
METHODOLOGICAL UNCERTAINTIES
FINITE-SIZE OF REPLICA SAMPLE

- 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



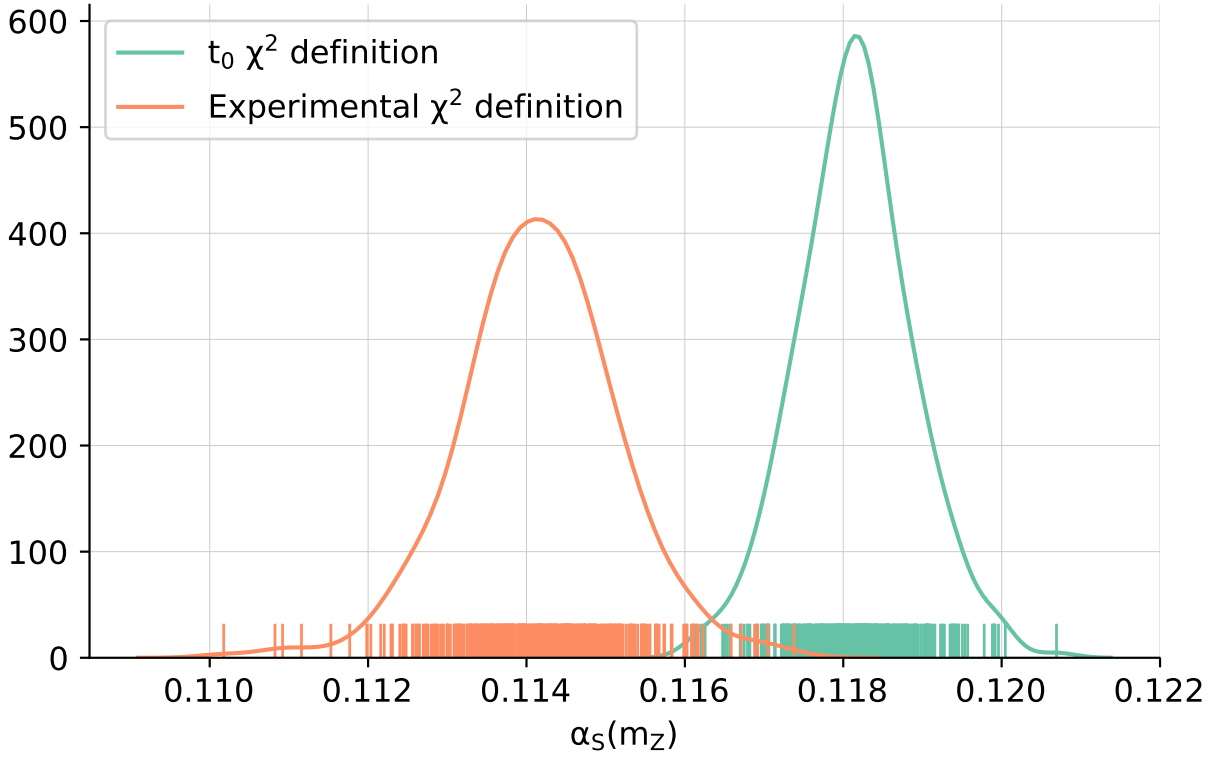
METHODOLOGICAL UNCERTAINTIES
PARABOLIC FITTING

- 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

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

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!

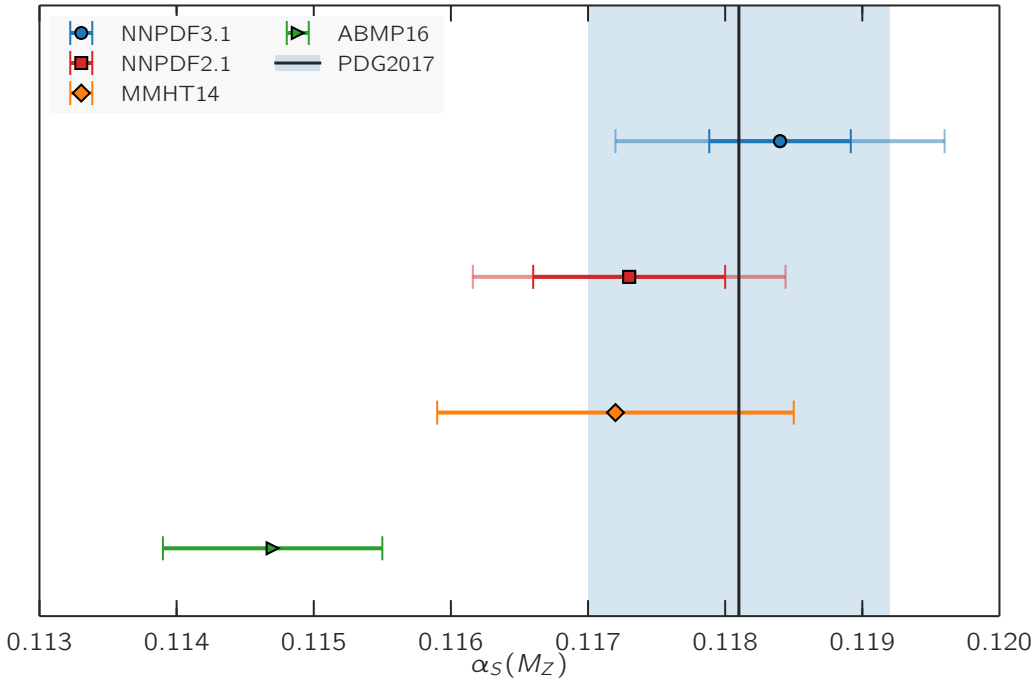
TOTAL THEORY UNCERTAINTY
OUR VERY CONSERVATIVE ESTIMATE:

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

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

FINAL RESULT & COMPARISON

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

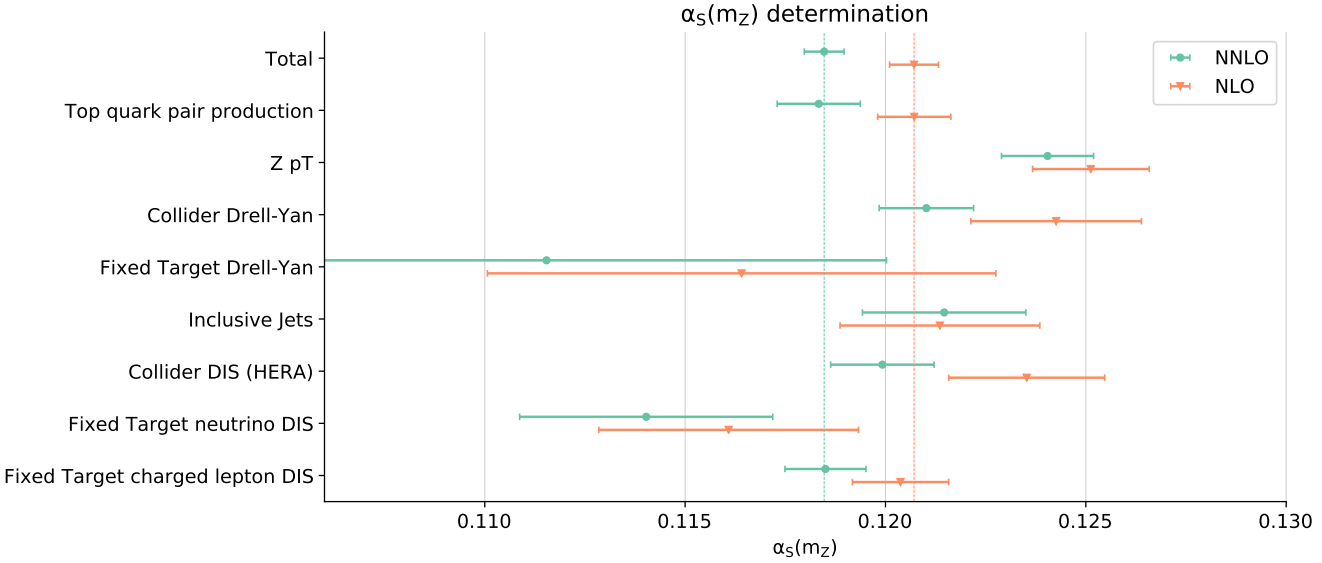


- SIGNIFICANTLY SMALLER EXP. UNCERTAINTY IN COMPARISON TO PREVIOUS NNPDF2.1 DETERMINATION (DESPITE MORE CONSERVATIVE ESTIMATE)
- SOMEWHAT LARGER CENTRAL VALUE THAN MMHT

THE IMPACT OF LHC DATA

PULLS FROM DATA SUBSETS

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



LESSONS

- NO SIGNS OF TENSION OR INCOMPATIBILITY
- HADRON COLLIDER DATA PULL TOWARDS LARGER α_s

CAVEATS

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

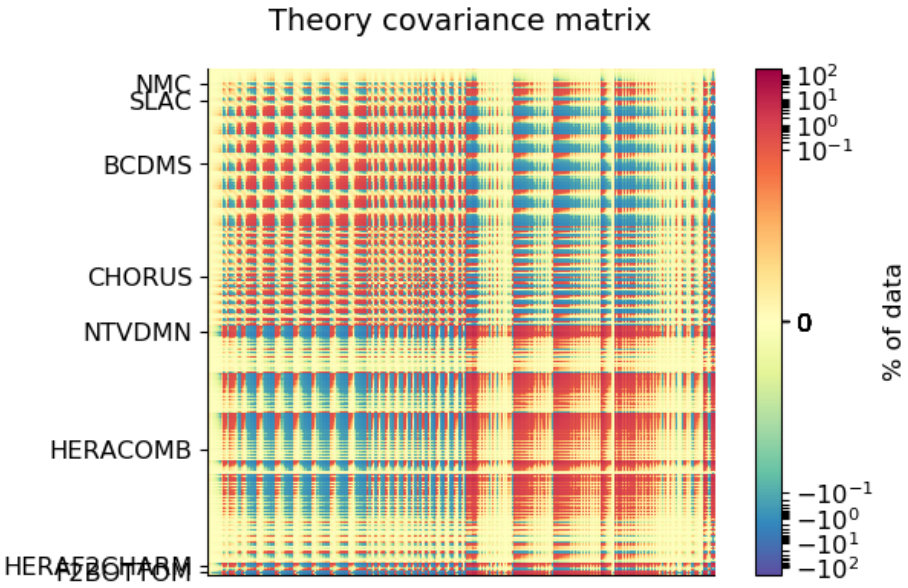
LESSONS LEARNT

- LHC DATA $\Rightarrow$ PDF + α_s UNCERTAINTIES AT THE SUBPERCENT LEVEL
- LHC DATA $\Rightarrow$ LARGER α_s VALUE
- PDF-SENSITIVE α_s DETERMINATION WITHOUT SIMULTANEOUS PDF FIT IS DANGEROUS AND LESS ACCURATE
- MORE GLOBAL DATASET $\Rightarrow$ GREATER PERTURBATIVE STABILITY

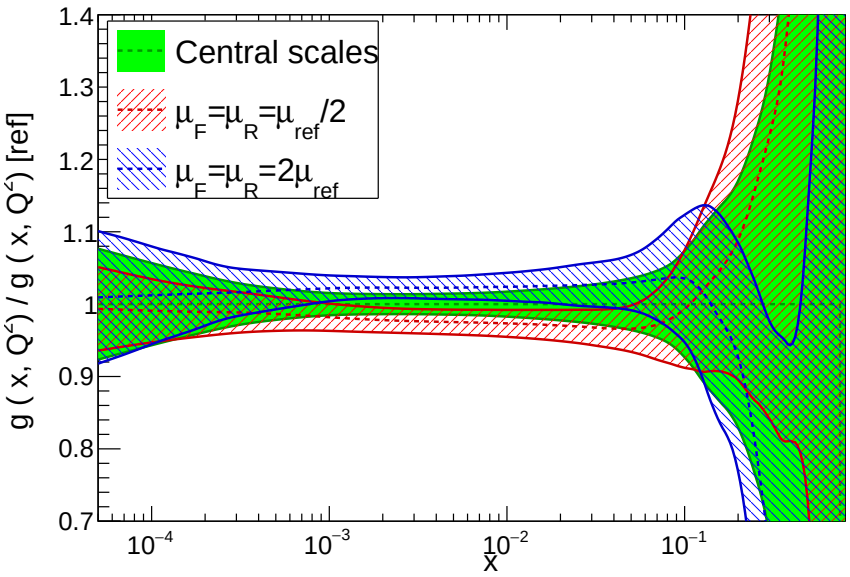
CAN WE DO BETTER?

THEORETICAL UNCERTAINTIES: MHO

THEORY UNCERTAINTIES: COVARIANCE MATRIX



GLUON PDF: SCALE VARIATION



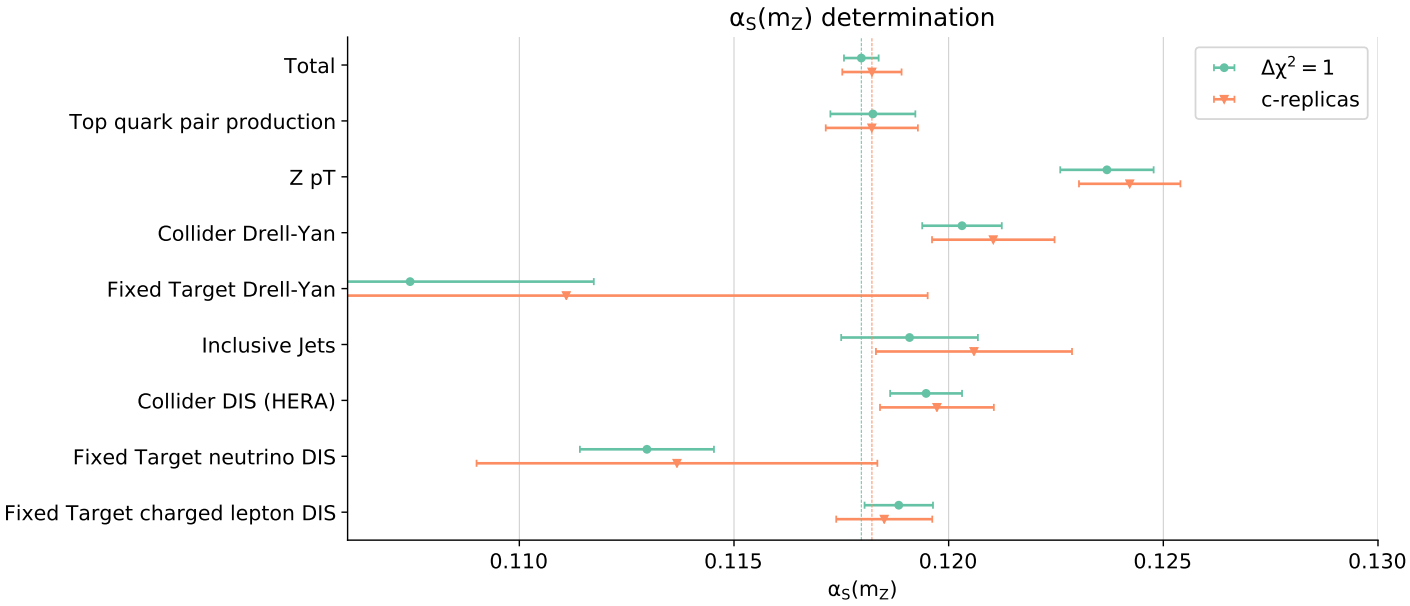
- PDF FITS WITH SCALE VARIED THEORY
- THEORY UNCERTAINTIES (FROM SCALE OR OTHERWISE)

  IN COVARIANCE MATRIX

  GENERALIZABLE TO OTHER THEORY UNCERTAINTIES
- THEORY UNCERTAINTIES INCLUDED IN PDF UNCERTAINTIES

EXTRAS

COMPARISON WITH THE OLD METHOD

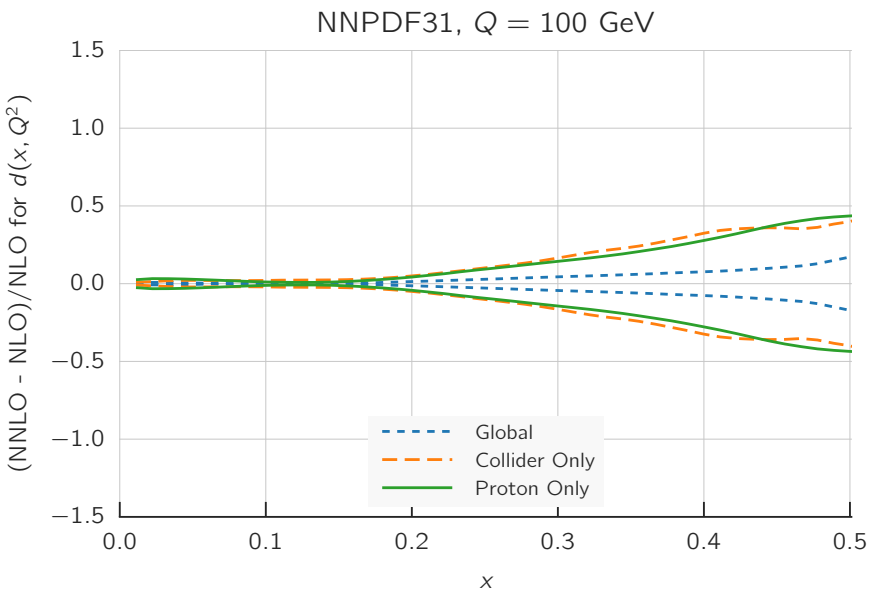


- GOOD AGREEMENT OF CENTRAL VALUES
- OLD METHOD $\Rightarrow$ SOMEWHAT UNDERESTIMATED UNCERTAINTIES

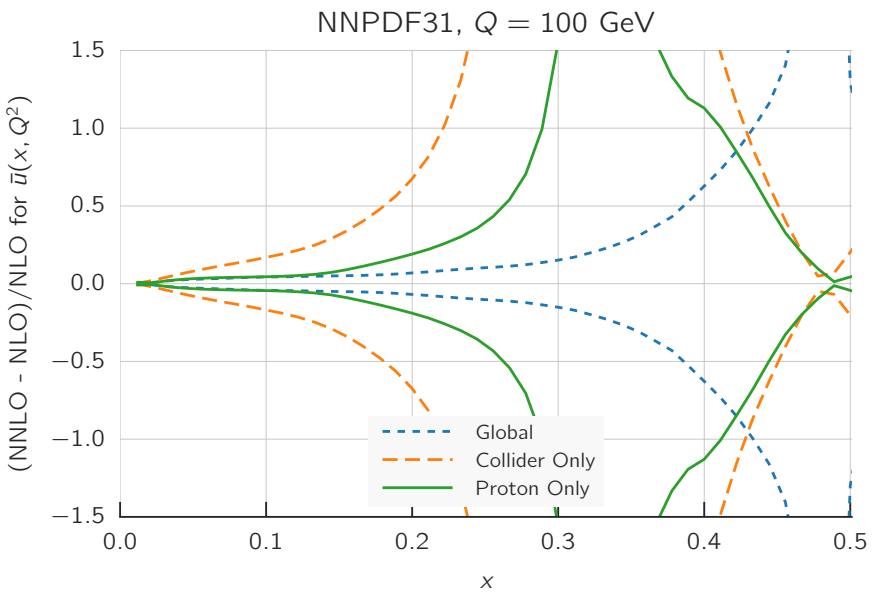
PERTURBATIVE STABILITY

GLOBAL VS RESTRICTED DATASETS

DOWN



ANTIUP



- NLO-NNLO SHIFTS SMALLER WITH LARGER DATASET
- GREATER STABILITY OF α_s ALSO OBSERVED