PDFs WITH 1% ACCURACY

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN

UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

Istituto Nazionale di Fisica Nucleare

CMS SMP-COM MEETING

JULY 6, 2021

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

PDFs WITH 1% ACCURACY

THE NNPDF COLLABORATION

RICHARD D. BALL, STEFANO CARRAZZA, JUAN CRUZ-MARTINEZ, LUIGI DEL DEBBIO,
STEFANO FORTE, TOMMASO GIANI, SHAYAN IRANIPOUR, ZAHARI KASSABOV,
JOSE I. LATORRE, EMANUELE R. NOCERA, ROSALYN L. PEARSON, JUAN ROJO,
ROY STEGEMAN, CHRISTOPHER SCHWAN, MARIA UBIALI, CAMERON VOISEY, MICHAEL WILSON

AMSTERDAM-CAMBRIDGE-EDINBURGH-INFN-MILAN-NIKHEF-SINGAPORE

CMS SM MEETING JULY 6, 2021

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

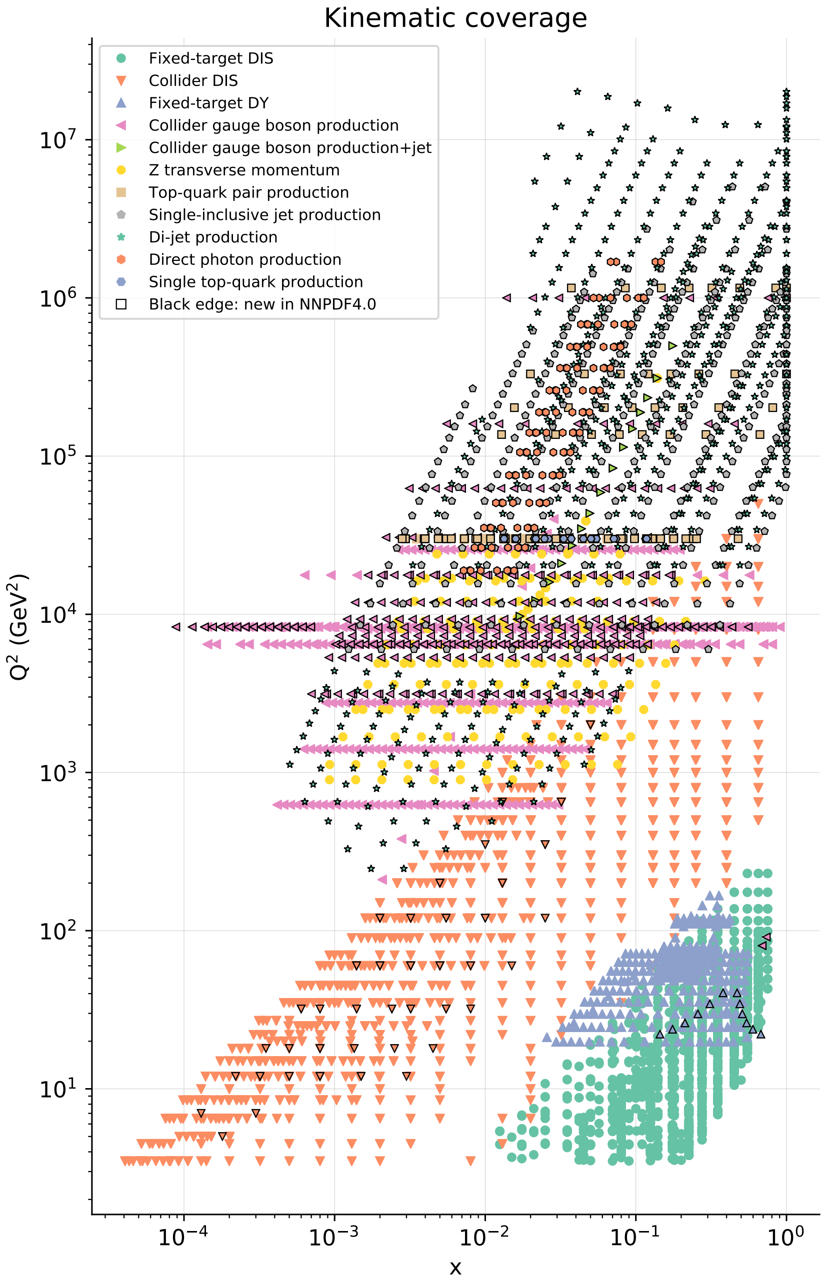
SUMMARY

NNPDF4.0

- DATA
- THEORY
- METHODOLOGY
- DATASET SELECTION
- UNCERTAINTIES
- VALIDATION
- STABILITY
- PHENOMENOLOGY
- DELIVERY

DATA

DATA



- ABOUT 50 NEW DATASETS & 400 EXTRA DATAPOINTS
- FULL DIS AND FT DY DATASET
  - AS IN NNPDF3.1: FINAL HERA, NMC, BCDMS, CHORUS, NUTEV
  - NOW ALSO NOMAD NEUTRINO
  - SEAQUEST DY
- FULL 7 TEV AND 8 TEV DATASET & EXTENSIVE USE OF 13 TEV DATA:
  - W, Z PRODUCTION: RAPIDITY DISTRIBUTIONS, ASYMMETRIES, Z p_T DISTRIBUTIONS
  - TOP PAIR PRODUCTION: ALL AVAILABLE DISTRIBUTIONS
  - SINGLE-INCLUSIVE JETS
- SEVERAL NEW PROCESSES:
  - PROMPT PHOTON
  - SINGLE TOP
  - DIJETS
  - HERA JETS

LHC DATA

LHCb

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
LHCb Z 940 pb	✓	✓	✗	✗	✓
LHCb $Z \to ee$ 2 fb	✓	✓	✓	✓	✓
LHCb $W, Z \to \mu$ 7 TeV	✓	✓	✓	✓	✓
LHCb $W, Z \to \mu$ 8 TeV	✓	✓	✓	✓	✓
LHCb $Z \to \mu\mu, ee$ 13 TeV	✓	✗	✗	✗	✗

ATLAS

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
ATLAS W, Z 7 TeV (2010)	✓	✓	✓	✓	✓
ATLAS W, Z 7 TeV (2011)	✓	✓	✗	✓	✓
ATLAS low-mass DY 7 TeV	✓	✓	✗	✗	✗
ATLAS high-mass DY 7 TeV	✓	✓	✗	✗	✓
ATLAS W 8 TeV	✓	✗	✗	✗	✓
ATLAS DY 2D 8 TeV	✓	✗	✗	✗	✓
ATLAS high-mass DY 2D 8 TeV	✓	✗	✗	✗	✓
ATLAS $\sigma_{W,Z}$ 13 TeV	✓	✗	✓	✗	✗
ATLAS W^++jet 8 TeV	✓	✗	✗	✗	✓
ATLAS Z p_T 8 TeV	✓	✓	✗	✓	✓
ATLAS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✓	✓	✓	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 13 TeV	✓	✓	✓	✗	✗
ATLAS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✓	✓
ATLAS $t\bar{t}$ dilepton 8 TeV	✓	✗	✗	✗	✓
ATLAS single-inclusive jets 7 TeV, R=0.6	✗	✓	✗	✓	✓
ATLAS single-inclusive jets 8 TeV, R=0.6	✓	✗	✗	✗	✗
ATLAS dijets 7 TeV, R=0.6	✓	✗	✗	✗	✗
ATLAS direct photon production 13 TeV	✓	✗	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	✓	✗	✓	✗	✗
ATLAS single top diff. 7, 8 TeV	✓	✗	✗	✗	✗
ATLAS single top diff. 8 TeV	✓	✗	✗	✗	✗

- CUTOFF DATE AROUND 06/2020
- DIJETS NOW INCLUDED ALONG WITH JETS
  CANNOT INCLUDE SIMULTANEOUSLY FROM SAME UNDERLYING DATASET

CMS

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
CMS W electron asymmetry 7 TeV	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	✓	✓	✗	✗	✓
CMS W rapidity 8 TeV	✓	✓	✓	✓	✓
CMS Z p_T 8 TeV	✓	✓	✗	✓	✗
CMS $W + c$ 7 TeV	✓	✓	✗	✗	✓
CMS $W + c$ 13 TeV	✓	✗	✗	✗	✗
CMS single-inclusive jets 2.76 TeV	✗	✓	✗	✗	✓
CMS single-inclusive jets 7 TeV	✗	✓	✗	✓	✓
CMS dijets 7 TeV	✓	✗	✗	✗	✗
CMS single-inclusive jets 8 TeV	✓	✗	✗	✓	✓
CMS 3D dijets 8 TeV	✗	✗	✗	✗	✗
CMS $\sigma_{tt}^{\rm tot}$ 5 TeV	✓	✗	✓	✗	✗
CMS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✓	✓	✓	✗	✓
CMS $\sigma_{tt}^{\rm tot}$ 13 TeV	✓	✓	✓	✗	✗
CMS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✗	✓
CMS $t\bar{t}$ 2D dilepton 8 TeV	✓	✗	✗	✓	✓
CMS $t\bar{t}$ lepton+jet 13 TeV	✓	✗	✗	✗	✗
CMS $t\bar{t}$ dilepton 13 TeV	✓	✗	✗	✗	✗
CMS single top $\sigma_t + \sigma_{\bar{t}}$ 7 TeV	✓	✗	✓	✗	✗
CMS single top R_t 8, 13 TeV	✓	✗	✓	✗	✗

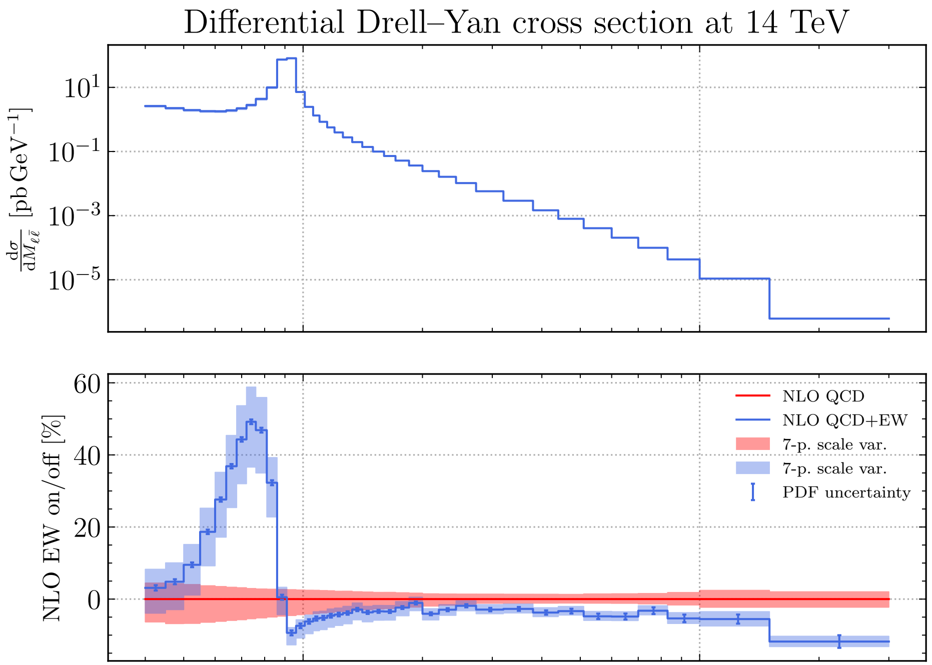
THEORY

ELECTROWEAK CORRECTIONS

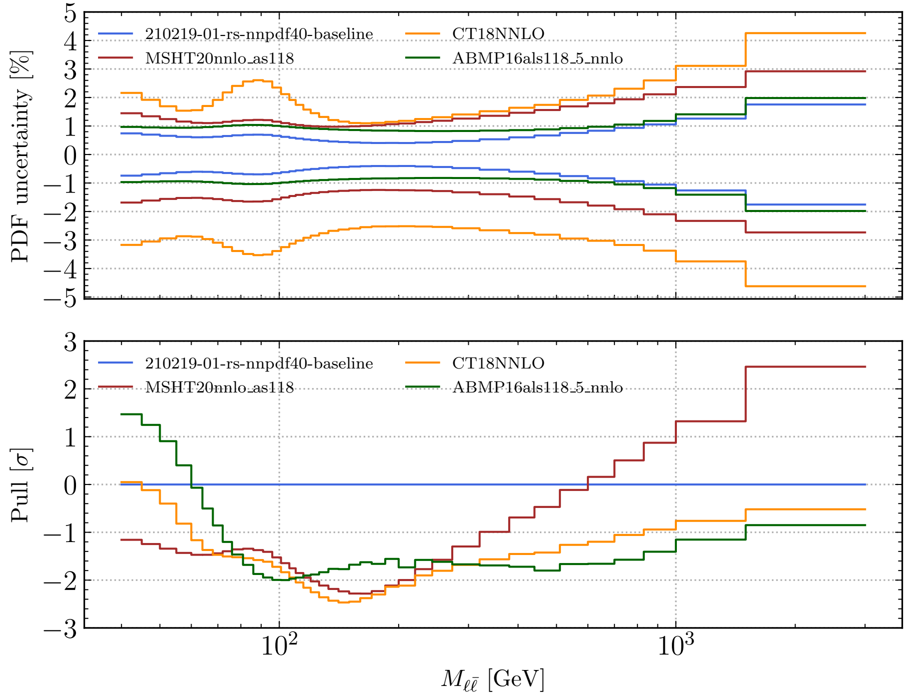
- PineAPPL FAST INTERFACE TO Madgraph5_aMC@NLO AVAILABLE
  (Schwan, Carrazza, Nocera, Zaro 2020)
  ⇒ FULL NLO EW+QCD POSSIBLE
- DATA W/O FSR & PHOTON-INITIATED SUBTRACTION OFTEN NOT AVAILABLE
- CURRENTLY USED FOR DATASET SELECTION:
  ⇒ DISCARDED IF EW CORRNS EXCEED THRESHOLD

EXAMPLE: DRELL-YAN AT 14 TEV

QCD+EW VS. QCD



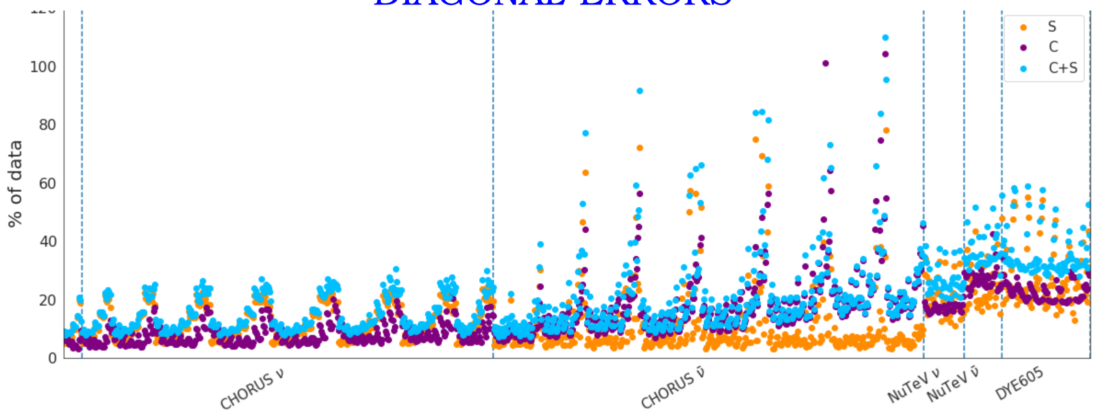
PDF UNCERTAINTIES



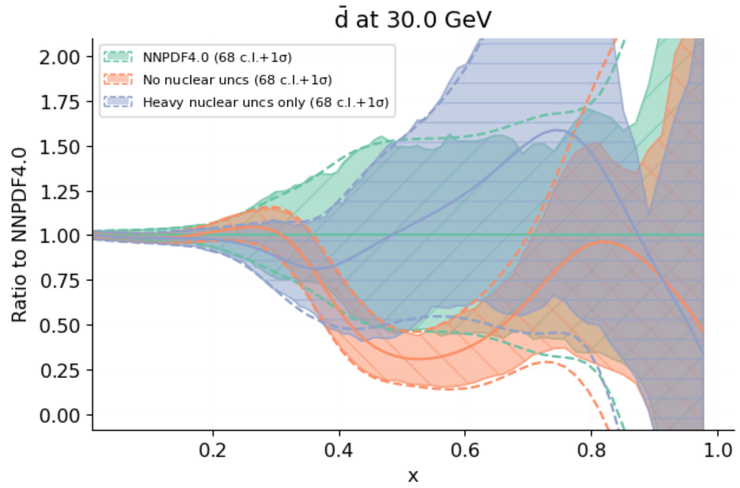
NUCLEAR CORRECTIONS

- INCLUDED AS CONTRIBUTION TO COVARIANCE MATRIX (FULLY CORRELATED) (Ball, Nocera, Pearson, 2019)
- COMPUTED AS SHIFT BETWEEN NUCLEAR & STANDARD PDF
- DEUTERIUM PDF DETERMINED FROM SELF-CONSISTENT NNPDF FIT (Ball, Nocera, Pearson, 2019)
- NUCLEAR PDFs FROM nNNPDF2.0 (Abdul Khalek, Ethier, Rojo, van Weelden, 2020)

DIAGONAL ERRORS



IMPACT ON PDFs

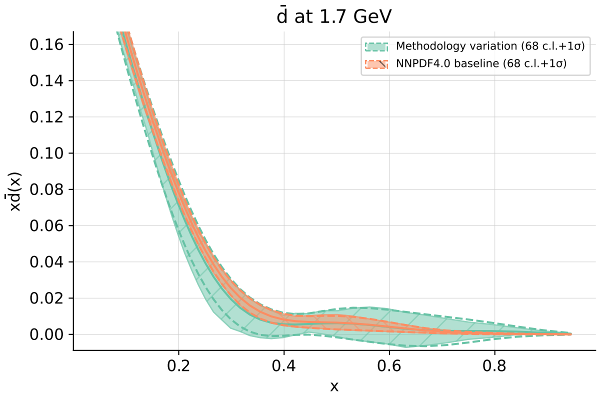


PDF POSITIVITY & INTEGRABILITY

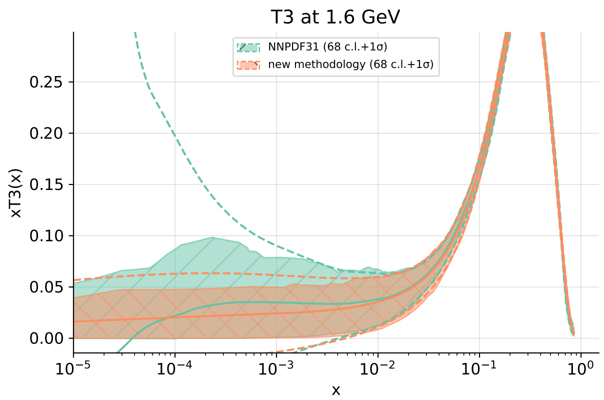
- $\overline{\text{MS}}$ PDFs ARE NON-NEGATIVE!(Candido, Hekhorn, Forte, 2020)
- PDF POSITIVITY IMPOSED (PREVIOUSLY: OBSERVABLE POSITIVITY)
  $\Rightarrow$ SMALLER LARGE x UNCERTAINTIES
- SEA NONSINGLET COMBINATIONS INTEGRABLE:
  GOTTFRIED $u + \bar{u} - (d + \bar{d})$
  STRANGENESS $u + \bar{u} + (d + \bar{d}) - 2(s + \bar{s})$
  $\Rightarrow$ SMALLER SMALL x UNCERTAINTIES

IMPACT ON PDFs

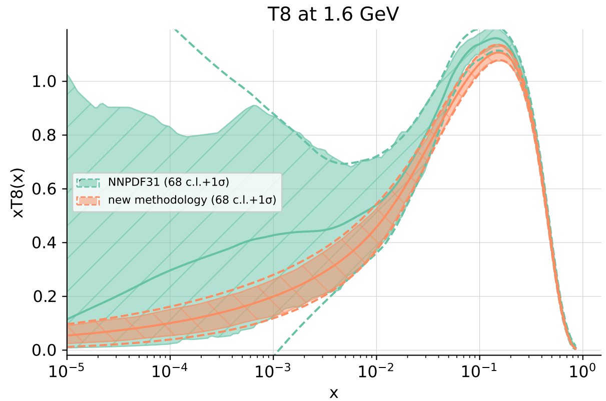
POS.: ANTIDOWN



INTEGR.: GOTTFRIED SR



INTEGR.: STRANGE SEA SR



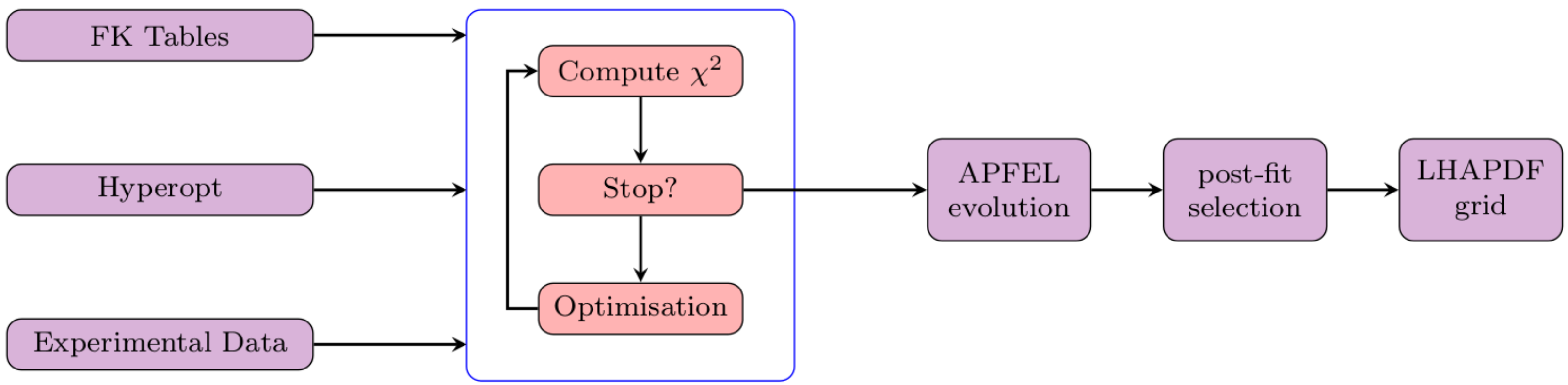
METHODOLOGY

THE NNPDF CODE STRUCTURE

- MODULAR PYTHON-BASED CODE
- HIGH DEGREE PARALLELIZATION & HARDWARE ACCELERATION

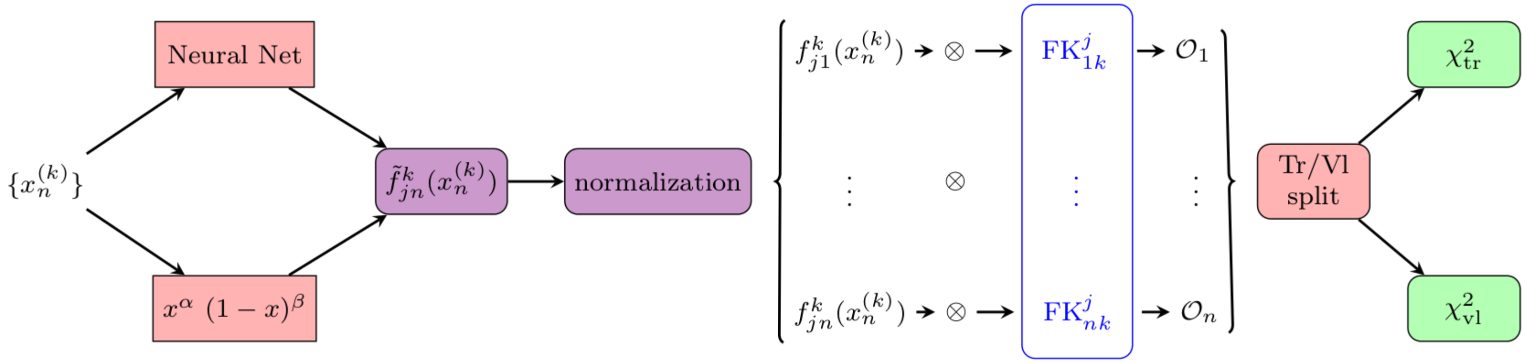
AVERAGE FITTING TIME PER REPLICA AND USE OF RESOURCES
SAME DATASET FOR OLD AND NEW METHODOLOGIES IN CPU AND GPU
CPU: INTEL(R) CORE(TM) I7-4770 AT 3.40GHZ; GPU: NVIDIA TITAN V

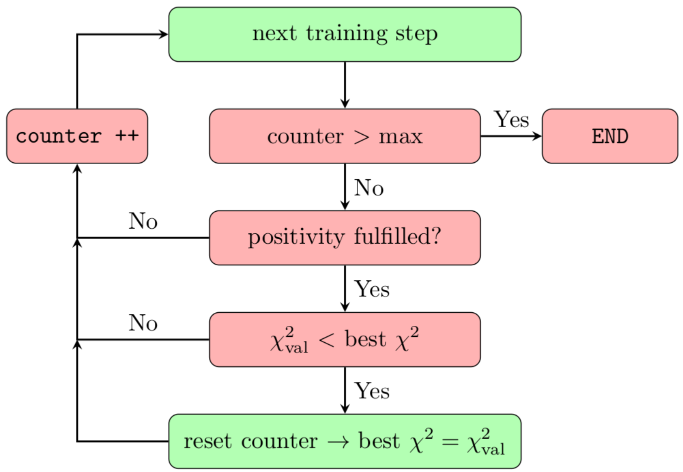
	NNPDF31 CODEBASE	NNPDF40 CODEBASE IN CPU	NNPDF40 CODEBASE IN GPU
TIME	15.2 H.	38 ± 5 MIN.	6.6 MIN.
RAM USE	1.5 GB	6.1 GB	NA



MINIMIZATION AND CROSS-VALIDATION

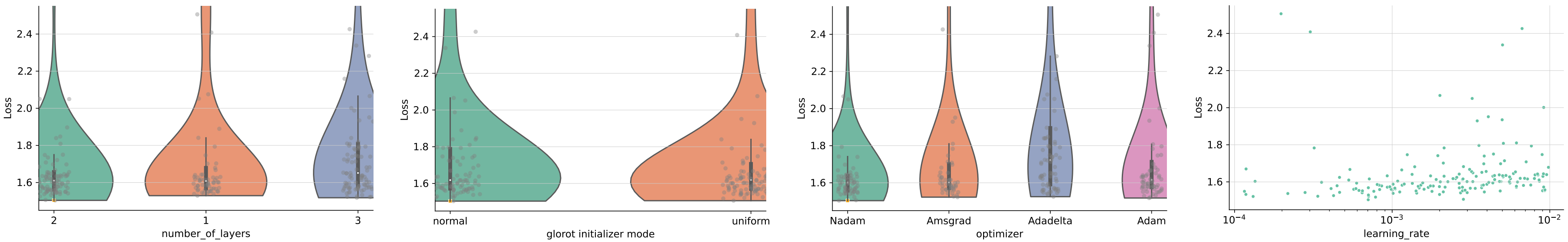
- DATA REPLICAS $\Rightarrow$ PDF REPLICAS
- EACH PDF REPLICA: PREPROCESSED NEURAL NET
- NEURAL NET $\Rightarrow$ OBSERVABLES
- RANDOM TRAINING-VALIDATION SPLIT, χ^2 TO TRAINING DATA REPLICAS MINIMIZED
- TRAINING STOPS IF VALIDATION χ^2 GROWS FOR A WHILE (PATIENCE)
- LOWEST VALIDATION χ^2 $\Rightarrow$ OPTIMAL FIT



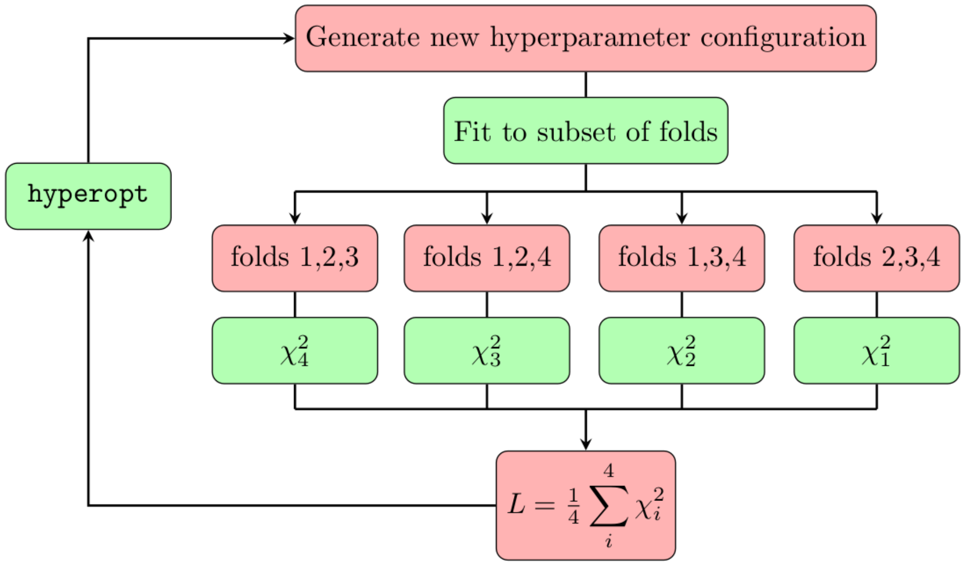


HYPEROPTIMIZATION

- PARAMETRIZATION AND MINIMIZATION PARAMETERS VARIED
- SCAN OF PARAMETER SPACE
- BAYESIAN UPDATING LEADS TO BEST METHODOLOGY



K-FOLDING



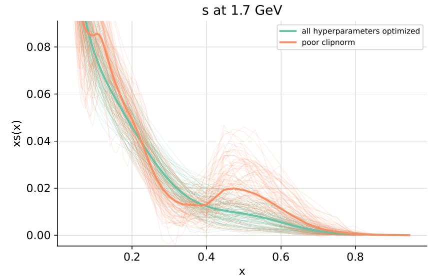
Fold 1		
CHORUS σ_{CC}^{ν}	HERA I+II inc NC e^+p 920 GeV	BCDMS p
LHCb Z 940 pb	ATLAS W, Z 7 TeV 2010	CMS Z p_T 8 TeV (p_T^{ll}, y_{ll})
DY E605 σ_{DY}^{p}	CMS Drell-Yan 2D 7 TeV 2011	CMS 3D dijets 8 TeV
ATLAS single-$\bar{t}$ y (normalised)	ATLAS single top R_t 7 TeV	CMS $t\bar{t}$ rapidity $y_{t\bar{t}}$
CMS single top R_t 8 TeV		

Fold 2		
HERA I+II inc CC e^-p	HERA I+II inc NC e^+p 460 GeV	HERA comb. $\sigma_{b\bar{b}}^{\mathrm{red}}$
NMC p	NuTeV $\sigma_c^{\bar{\nu}}$	LHCb $Z \to ee$ 2 fb
CMS W asymmetry 840 pb	ATLAS Z p_T 8 TeV (p_T^{ll}, M_{ll})	D0 $W \to \mu\nu$ asymmetry
DY E886 σ_{DY}^{p}	ATLAS direct photon 13 TeV	ATLAS dijets 7 TeV, R=0.6
ATLAS single antitop y (normalised)	CMS $\sigma_{tt}^{\mathrm{tot}}$	CMS single top $\sigma_t + \sigma_{\bar{t}}$ 7 TeV

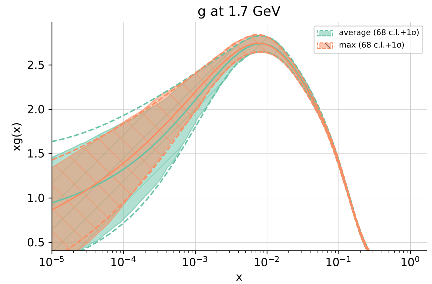
Fold 3		
HERA I+II inc CC e^+p	HERA I+II inc NC e^+p 575 GeV	NMC d/p
NuTeV σ_c^{ν}	LHCb $W, Z \to \mu$ 7 TeV	LHCb $Z \to ee$
ATLAS W, Z 7 TeV 2011 Central selection	ATLAS W^++jet 8 TeV	ATLAS HM DY 7 TeV
CMS W asymmetry 4.7 fb	DYE 866 $\sigma_{\mathrm{DY}}^{d}/\sigma_{\mathrm{DY}}^{p}$	CDF Z rapidity (new)
ATLAS $\sigma_{tt}^{\mathrm{tot}}$	ATLAS single top y_t (normalised)	CMS $\sigma_{tt}^{\mathrm{tot}}$ 5 TeV
CMS $t\bar{t}$ double diff. $(m_{t\bar{t}}, y_t)$		

Fold 4		
CHORUS $\sigma_{CC}^{\bar{\nu}}$	HERA I+II inc NC e^+p 820 GeV	LHCb $W, Z \to \mu$ 8 TeV
LHCb $Z \to \mu\mu$	ATLAS W, Z 7 TeV 2011 Fwd	ATLAS W^-+jet 8 TeV
ATLAS low-mass DY 2011	ATLAS Z p_T 8 TeV (p_T^{ll}, y_{ll})	CMS W rapidity 8 TeV
D0 Z rapidity	CMS dijets 7 TeV	ATLAS single top y_t (normalised)
ATLAS single top R_t 13 TeV	CMS single top R_t 13 TeV	

NO K-FOLDING



K-FOLDING VARIATION



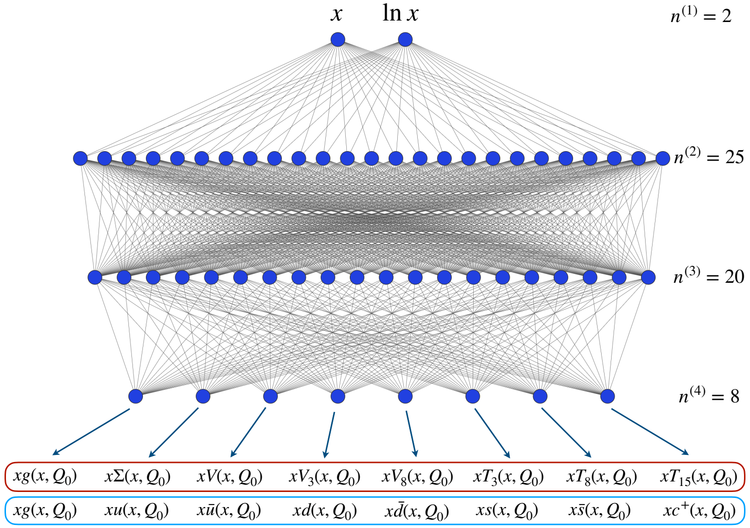
- HYPEROPTIMIZATION $\Rightarrow$ OVERFITTING (χ^2 TOO GOOD)
- CHECK GENERALIZATION POWER: K-FOLDING
  - DIVIDE DATA IN FOLDS
  - EXCLUDE ONE FOLD IN TURN FROM FIT
  - OPTIMIZE ON THE χ^2 OF THE EXCLUDED FOLDS
  - BEST AVERAGE OR BEST WORST

THE ML METHODOLOGY

HYPEROPTIMIZED PARAMETERS

Parameter	NNPDF4.0	L as in Eq. (3.21)	Flavour basis Eq. (3.2)
Architecture	25-20-8	70-50-8	7-26-27-8
Activation function	hyperbolic tangent	hyperbolic tangent	sigmoid
Initializer	`glorot_normal`	`glorot_uniform`	`glorot_normal`
Optimizer	`Nadam`	`Adadelta`	`Nadam`
Clipnorm	6.0×10^{-6}	5.2×10^{-2}	2.3×10^{-5}
Learning rate	2.6×10^{-3}	2.5×10^{-1}	2.6×10^{-3}
Maximum # epochs	17×10^{3}	45×10^{3}	45×10^{3}
Stopping patience	10% of max epochs	12% of max epochs	16% of max epochs
Initial positivity $\Lambda^{(\mathrm{pos})}$	185	106	2
Initial integrability $\Lambda^{(\mathrm{int})}$	10	10	10

NN ARCHITECTURE



- HYPEROPT ADAPTS TO EXTERNAL CHOICES (E.G. PARAMETRIZATION BASIS)
- SIMILAR RESULTS CAN BE OBTAINED WITH RATHER DIFFERENT SETTINGS

DATASET SELECTION

DATA “TENSION”

PROBLEMATIC DATA

$\frac{\chi^2-1}{\sigma_{\chi^2}} \gg 1 \Leftrightarrow$ POOR FIT QUALITY

- MISSING HIGHER-ORDER CORRECTIONS
- NO RESUMMATION WHERE NEEDED
- ILL-CONDITIONED COVARIANCE MATRIX
- EXPERIMENTAL ISSUES

THE WEIGHTED FIT METHOD

- FLAG PROBLEMATIC DATASETS:
  - LARGE χ^2
  - LARGE FROBENIUS NUMBER OF COVMAT (EIGENVALUES TOO SMALL)
- REPEAT GLOBAL FIT WITH LARGE WEIGHT GIVEN TO EACH PROBLEMATIC DATASET IN TURN

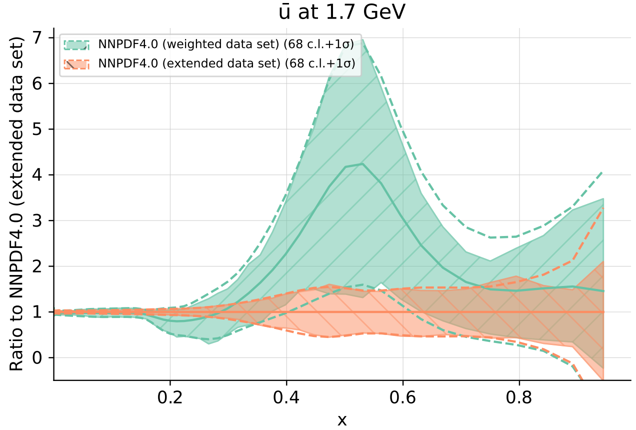
- χ^2 OF DATASET
  - UNCHANGED $\Rightarrow$ INTERNAL INCONSISTENCY
  - DECREASES $\Rightarrow$ TENSION
- GLOBAL χ^2
  - UNCHANGED $\Rightarrow$ CONSISTENT, KEEP
  - INCREASES $\Rightarrow$ INCONSISTENT, DISCARD

DATA "TENSION"

PROBLEMATIC DATA

$$\frac{\chi^2 - 1}{\sigma_{\chi^2}} \gg 1 \Leftrightarrow \text{POOR FIT QUALITY}$$

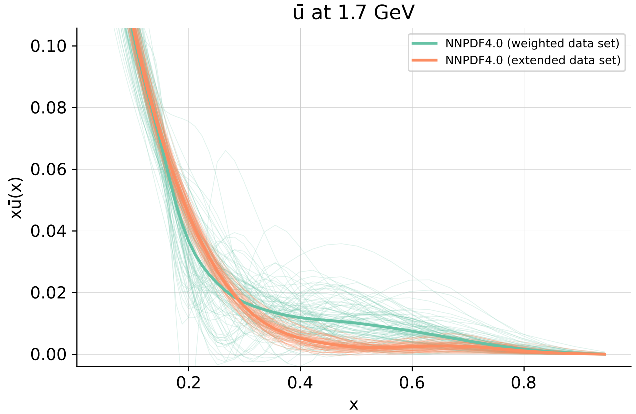
INCONSISTENT!



- MISSING HIGHER-ORDER CORRECTIONS
- NO RESUMMATION WHERE NEEDED
- ILL-CONDITIONED COVARIANCE MATRIX
- EXPERIMENTAL ISSUES

THE WEIGHTED FIT METHOD

INCONSISTENT!



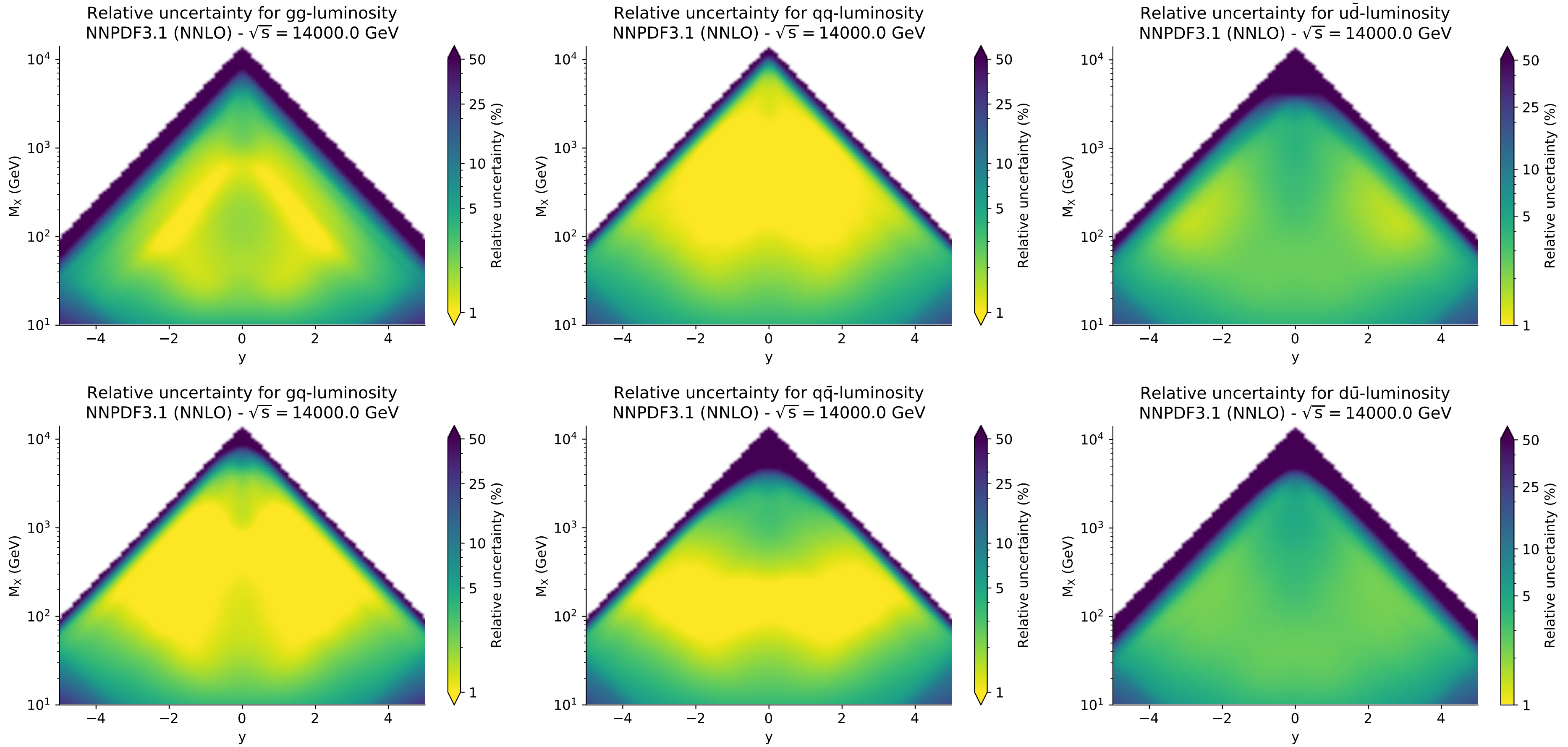
- FLAG PROBLEMATIC DATASETS:
  - LARGE χ^2
  - LARGE FROBENIUS NUMBER OF COVMAT (EIGENVALUES TOO SMALL)
- REPEAT GLOBAL FIT WITH LARGE WEIGHT GIVEN TO EACH PROBLEMATIC DATASET IN TURN

- χ^2 OF DATASET
  - UNCHANGED $\Rightarrow$ INTERNAL INCONSISTENCY
  - DECREASES $\Rightarrow$ TENSION
- GLOBAL χ^2
  - UNCHANGED $\Rightarrow$ CONSISTENT, KEEP
  - INCREASES $\Rightarrow$ INSONSISTENT, DISCARD

UNCERTAINTIES

UNCERTAINTIES: FROM NNPDF3.1...

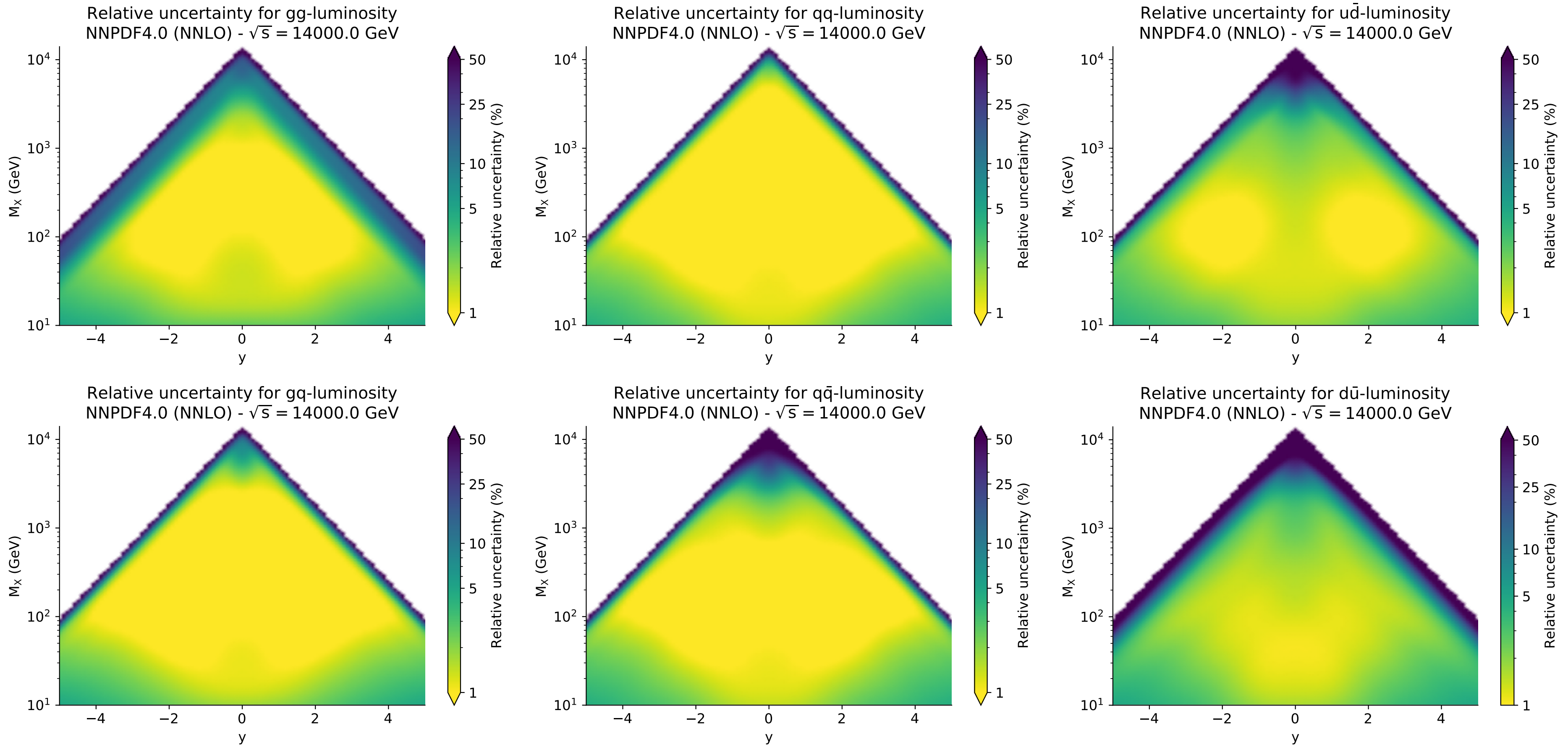
GLUON | SINGLET | FLAVORS



- TYPICAL UNCERTAINTIES IN DATA REGION: SINGLET $\sim 3\%$, NONSINGLET $\sim 5\%$
- DATA REGION: $10^2 \lesssim M_X \lesssim 10^3$ TeV, $-2 \lesssim y \lesssim 2$

UNCERTAINTIES: ...TO NNPDF4.0

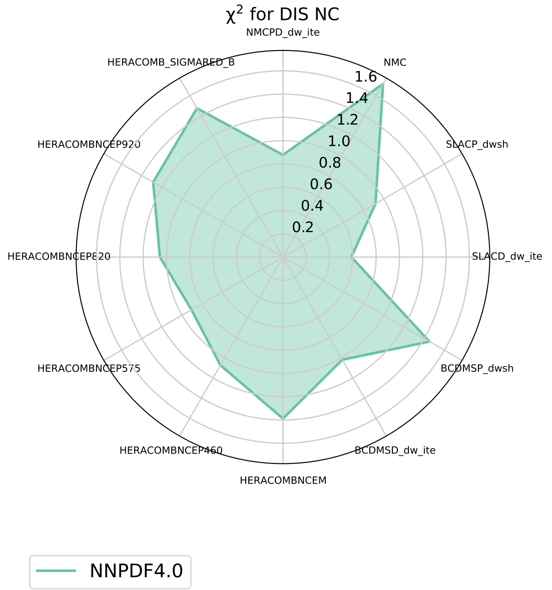
GLUON SINGLET FLAVORS



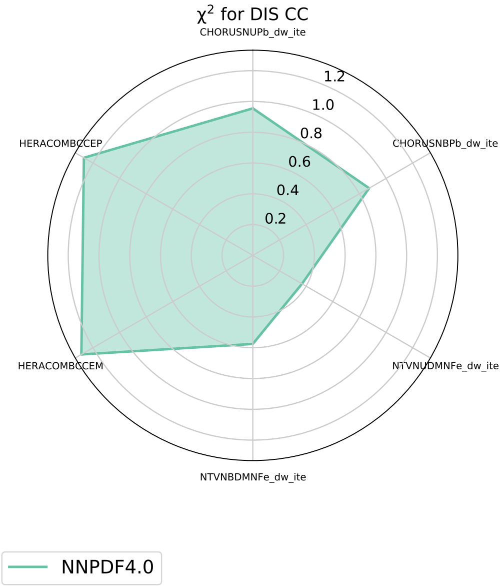
- TYPICAL UNCERTAINTIES IN DATA REGION: SINGLET $\sim 1\%$, NONSINGLET $\sim 2-3\%$
- DATA REGION: $10 \lesssim M_X \lesssim 3 \cdot 10^3$ TeV, $-4 \lesssim y \lesssim 4$

NNPDF4.0 FIT QUALITY

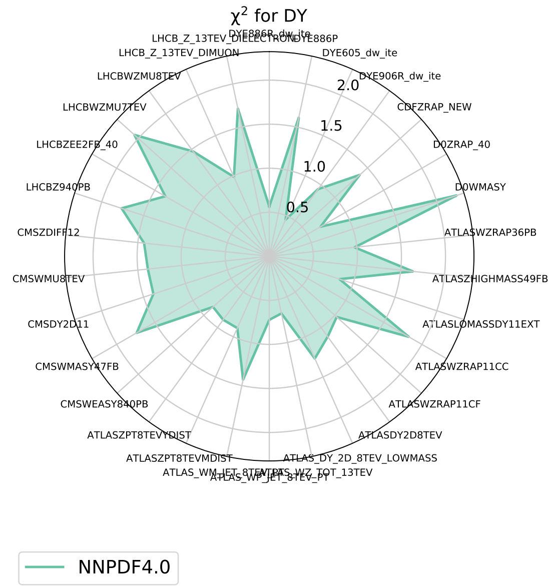
DIS NC



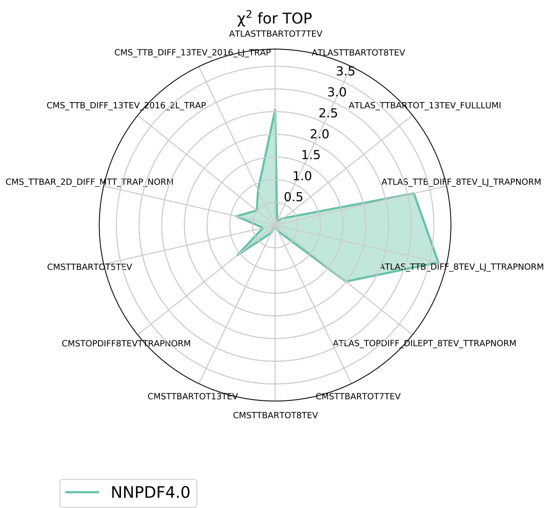
DIS CC



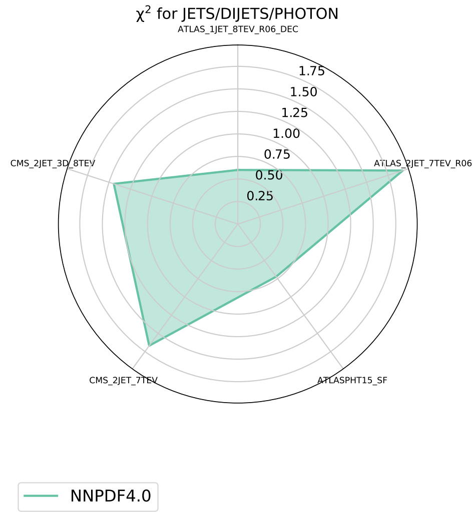
DY



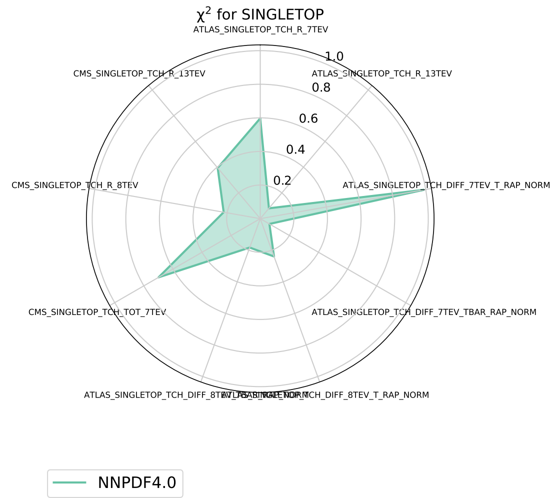
TOP PAIR



JETS



SINGLE TOP



- MOST χ^2 VALUES OF ORDER ONE PER DATAPOINT
- OUTLIERS $\Rightarrow$ FLAGGED DATASETS
- LARGE DATASETS (DIS) WELL FITTED

VALIDATION

CLOSURE TESTS

FAITHFUL UNCERTAINTIES IN DATA REGION?

- ASSUME "TRUE" UNDERLYING PDF $\Rightarrow$ E.G. SOME RANDOM PDF REPLICA
- GENERATE DATA DISTRIBUTED ACCORDING TO EXPERIMENTAL COVARIANCE MATRIX
- RUN WHOLE METHDOLOGY ON THESE DATA
- DO STATISTICS ON "RUNS OF THE UNIVERSE", POSSIBLE THANKS TO EFFICIENT METHDOLOGY: COMPARE TO TRUE PDFS, OR TO TRUE VALUES OF OBSERVABLES (NOT FITTED)
  - BIAS/VARIANCE: MEAN SQUARE DEVIATION WR TO TRUTH VS UNCERTAINTY
  - IS TRUTH WITHIN ONE SIGMA 68% OF TIMES?

RESULTS

BIAS/VARIANCE

experiment	$\mathcal{R}_{bv}$	bootstrap error
SeaQuest	0.99	0.10
ATLAS	1.08	0.04
CMS	1.04	0.06
LHCb	0.92	0.07
Total	1.03	0.05

ONE SIGMA

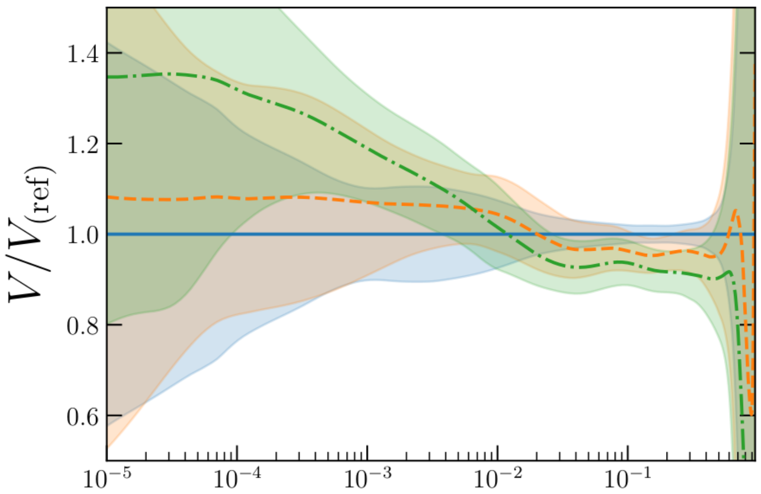
experiment	$\xi_{1\sigma}^{(\mathrm{data})}$	bootstrap error	$\mathrm{erf}(\mathcal{R}_{bv}/\sqrt{2})$	bootstrap error
SeaQuest	0.67	0.05	0.69	0.05
ATLAS	0.69	0.02	0.64	0.02
CMS	0.68	0.02	0.67	0.03
LHCb	0.68	0.04	0.72	0.03
Total	0.69	0.02	0.67	0.03

FUTURE TESTS

FAITHFUL UNCERTAINTIES IN EXTRAPOLATION?

- DETERMINE PDFS FROM A SUBSET OF CURRENT DATA: “PRE-HERA”, “PRE-LHC”,. . .
- COMPUTE χ^2 TO THE FULL CURRENT DATASET:
  - WITHOUT PDF UNCERTAINTIES $\Rightarrow$ IF $\gg 1$, MISSING INFORMATION
  - WITH PDF UNCERTAINTY $\Rightarrow$ IF ~ 1, MISSING INFO REPRODUCED BY UNCERTAINTY

VALENCE: PRE-HERA VS PRE-LHC VS NNPDF4.0



χ^2 WITHOUT/**WITH PDF UNC.**

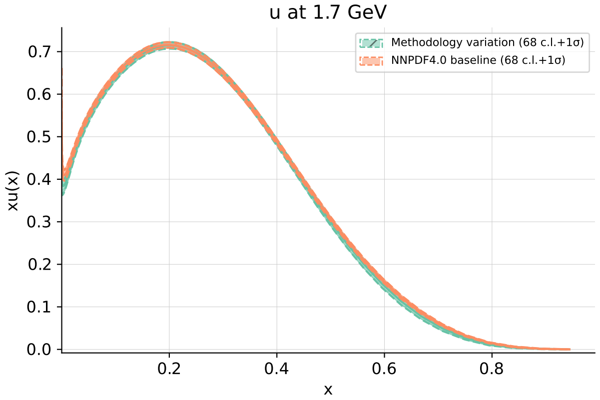
PROCESS	PRE-HERA	PRE-LHC	3.1-LIKE	4.0-GLOB
FT DIS (NC)	1.04	1.17	1.17	1.26
FT DIS (CC)	0.80	0.86	0.88	0.90
FT DY	0.93	1.27	1.43	1.59
HERA	24.01/**1.12**	1.22	1.21	1.21
COLL. DY (TEV.)	5.31/**1.08**	0.96	0.95	1.13
COLL. DY (LHC)	15.50/**1.37**	2.64/**1.54**	1.39	1.54
TOP QUARK	23.35/**1.08**	1.29/**0.86**	0.82	0.98
JETS	6.18/**1.21**	3.66/**1.29**	2.07/**1.37**	1.26
TOTAL	9.70	1.44	1.22	1.17

STABILITY

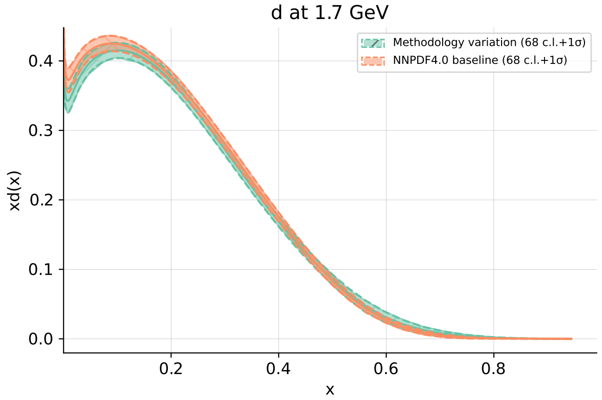
PARAMETRIZATION BASIS

- PDFs by default parametrized in "EVOLUTION BASIS":
  SINGLET $\Sigma = \sum_i q_i + \bar{q}_i$, VALENCE $V = \sum_i q_i - \bar{q}_i$, TRIPLET $T_3 = u + \bar{u} - (d + \bar{d})$, . . .
- WHAT IF ONE CHOOSES THE "FLAVOR BASIS": $u,\ \bar{u},\ d,\ \bar{d}$?
- COMPLETE STABILITY OF RESULTS!

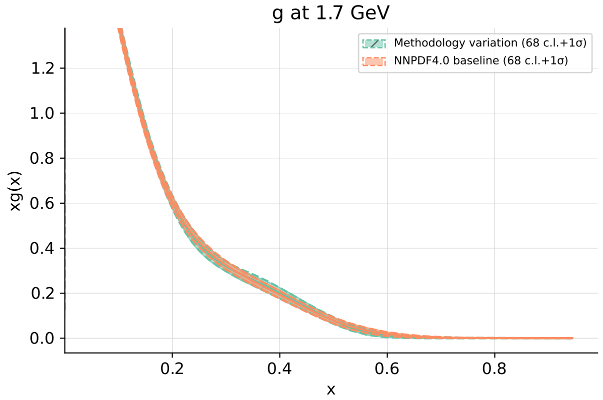
UP



DOWN



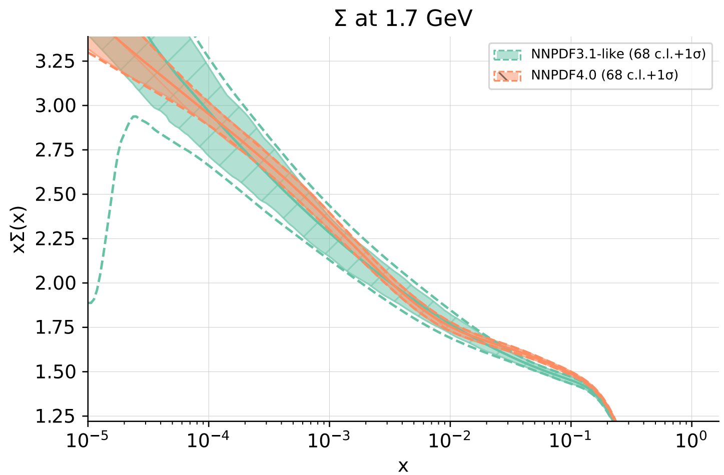
GLUON



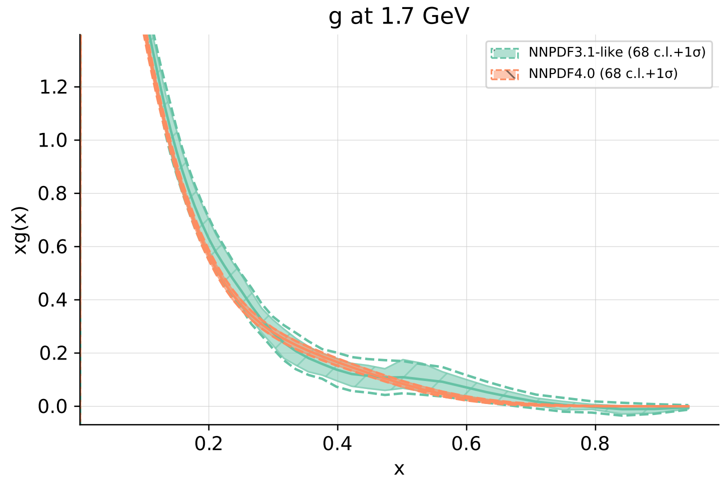
NNPDF4.0 vs. NNPDF3.1

- FULL BACKWARD COMPATIBILITY
- SUBSTANTIAL REDUCTION IN UNCERTAINTY

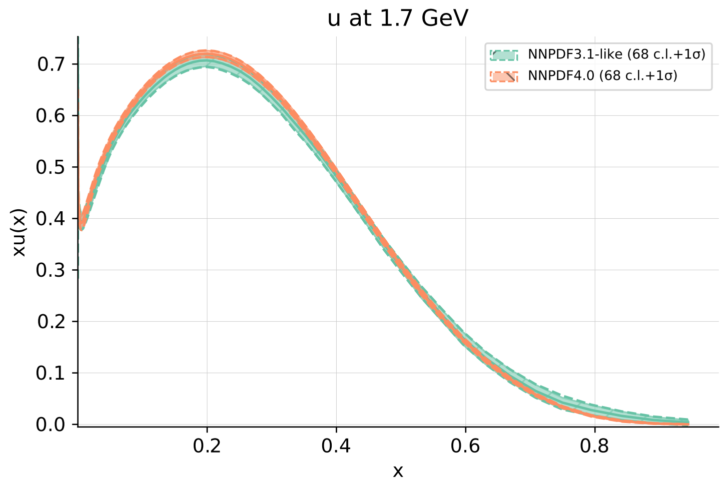
SINGLET



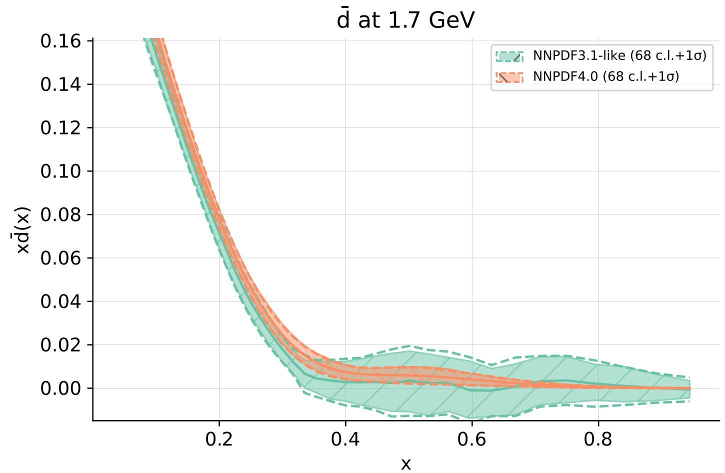
GLUON



UP



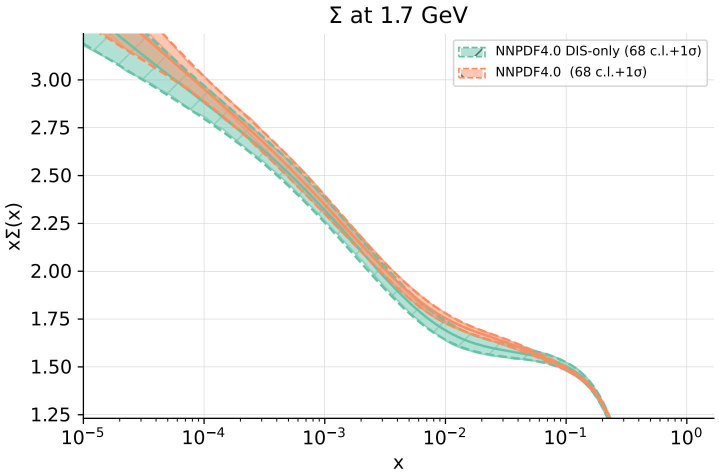
ANTIDOWN



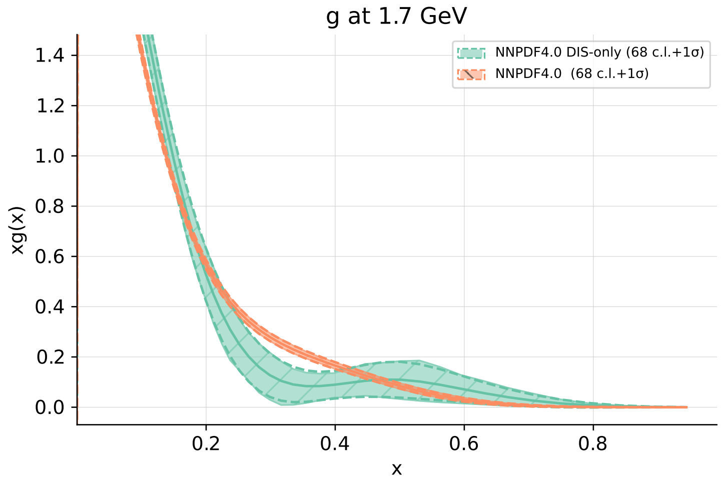
NNPDF4.0 VS DIS-ONLY

- DIS-ONLY FIT NO LONGER COMPETITIVE
- HADRONIC DATA NEEDED FOR PRECISION

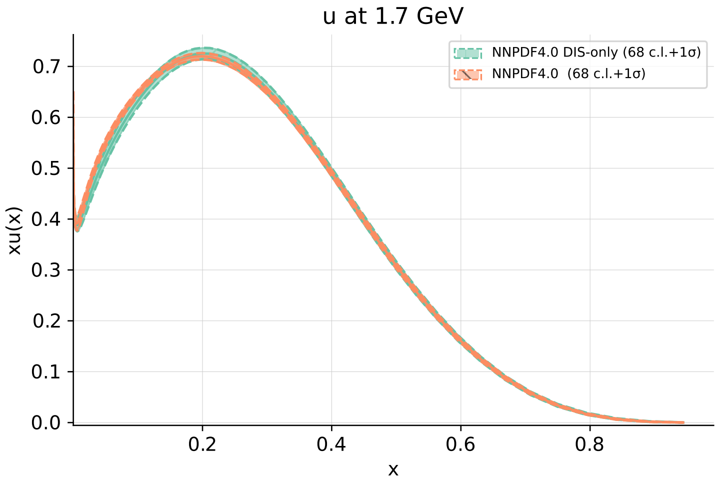
SINGLET



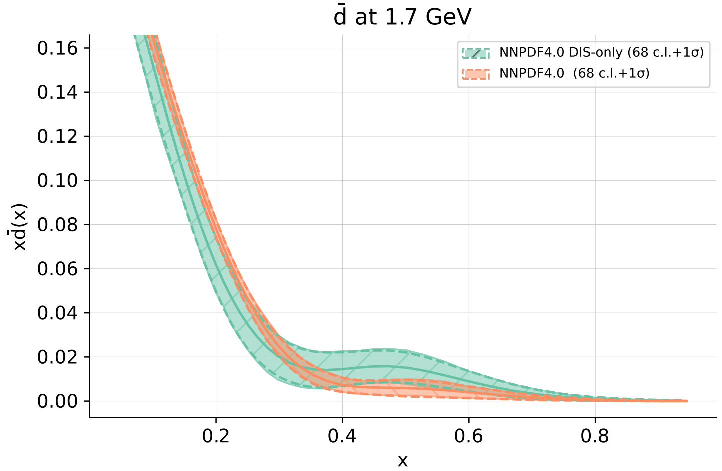
GLUON



UP



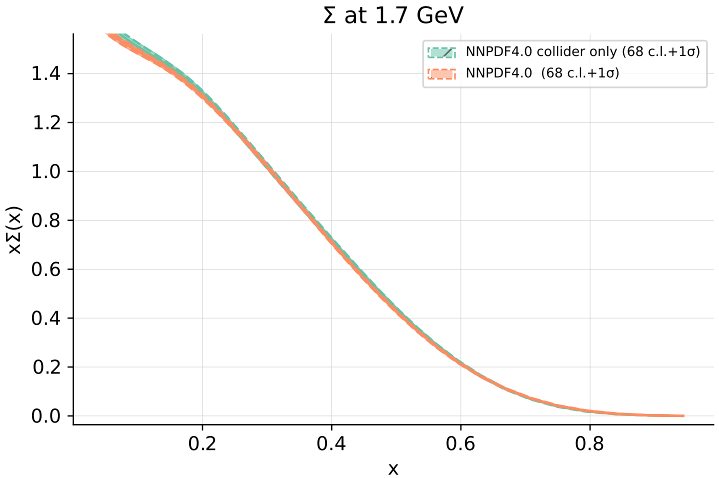
ANTIDOWN



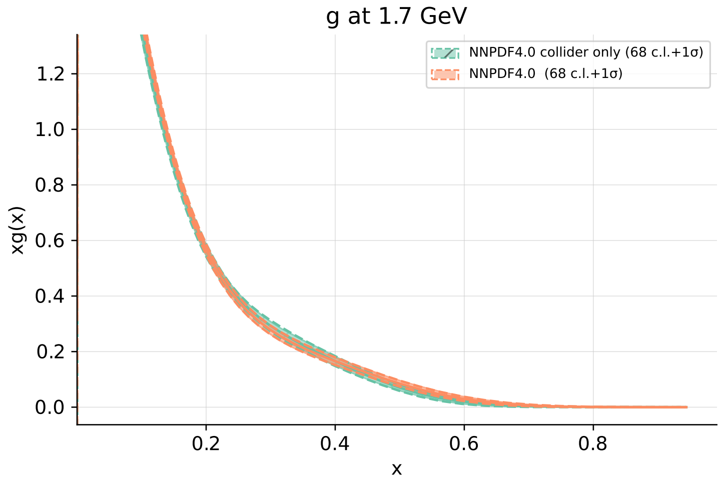
NNPDF4.0 VS COLLIDER ONLY

- COLLIDER ONLY COMPETITIVE!
- ONLY DEUTERIUM FIXED-TARGET DATA STILL RELEVANT

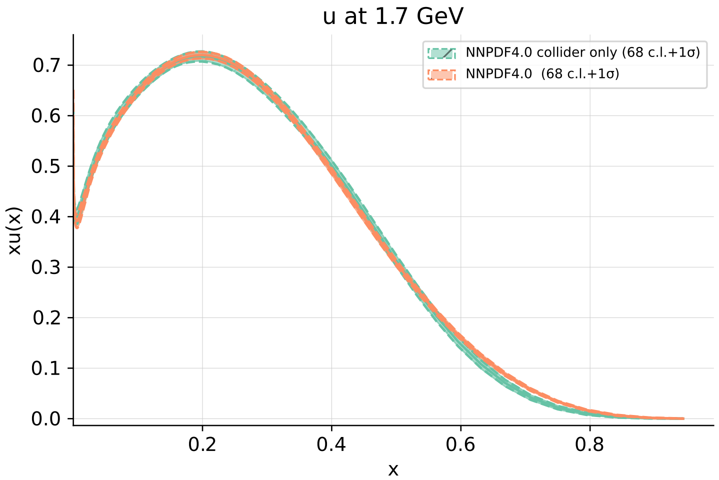
SINGLET



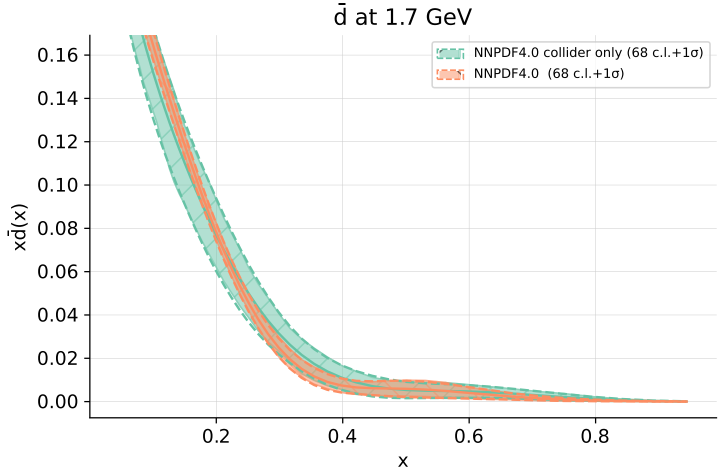
GLUON



UP



ANTIDOWN

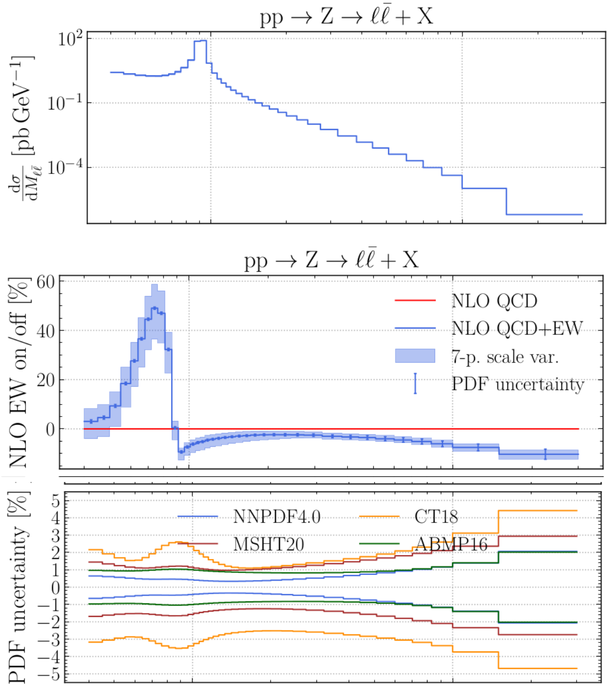


PHENOMENOLOGY

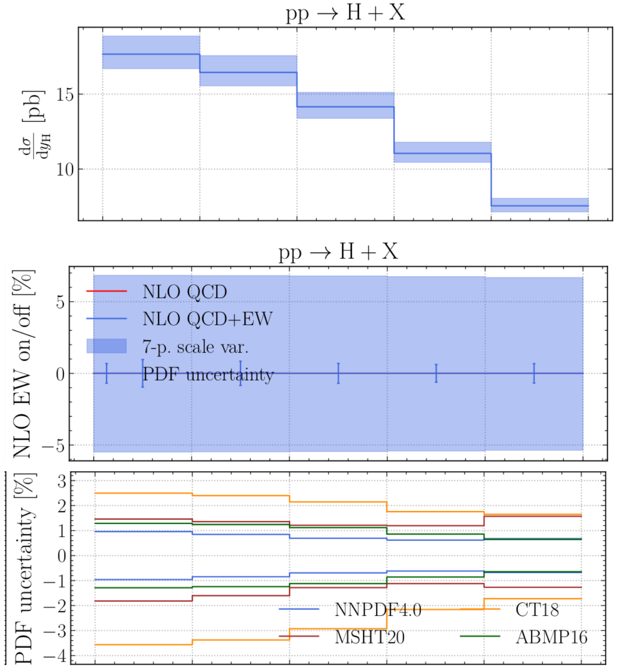
REPRESENTATIVE PROCESSES

- EW CORRECTIONS UNDER CONTROL
- SMALL UNCERTAINTIES

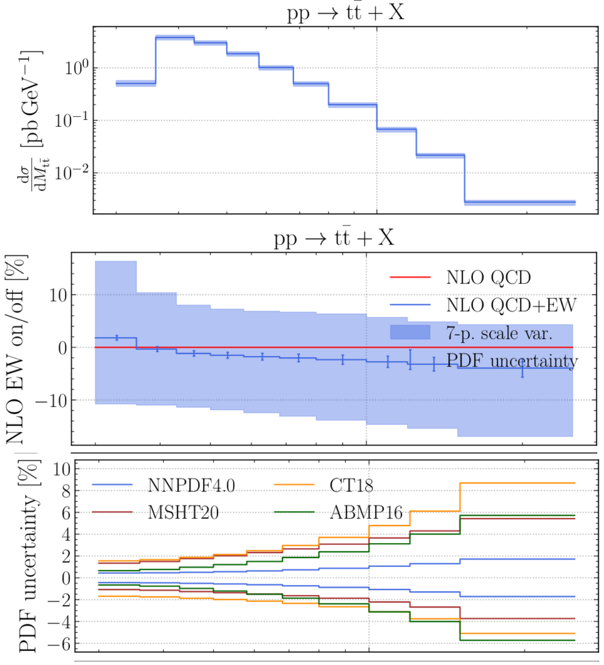
Z INVARIANT MASS



HIGGS RAPIDITY



TOP PAIR INV. MASS



DELIVERY

AN OPEN SOURCE CODE!

- THE FULL NNPDF CODE WILL BE MADE PUBLIC!
- INCLUDING HYPEROPTIMIZATION, EVOLUTION, THEORY, FITTING, VISUALIZATION
- FULLY DOCUMENTED CODE

An open-source machine learning framework for global analyses of parton distributions

The NNPDF Collaboration: Richard D. Ball · Stefano Carrazza · Juan Cruz-Martinez · Luigi Del Debbio · Stefano Forte · Tommaso Giani · Shayan Iranipour · Zahari Kassabov · Jose I. Latorre · Emanuele R. Nocera · Rosalyn L. Pearson · Juan Rojo · Roy Stegeman · Christopher Schwan · Maria Ubiali · Cameron Voisey · Michael Wilson

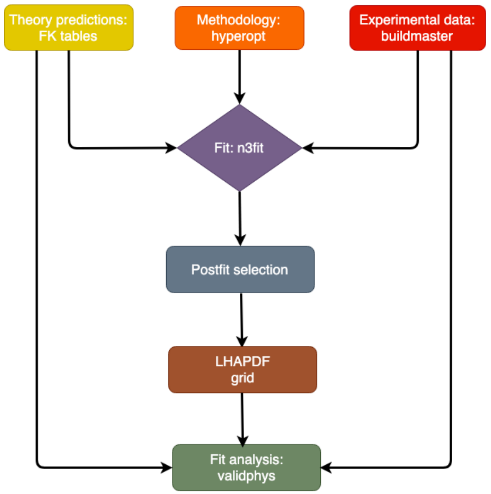


Fig. 2.1. Workflow for an NNPDF fit

OUTLOOK

THE ACCURACY CHALLENGE

- AT 1%, NOT ALL REDUCED DATASETS AGREE
- MUST INCLUDE MISSING HIGHER-ORDER CORRECTION UNCERTAINTIES
- INCLUDE EW CORRECTIONS
- GO BEYOND K-FACTORS
- WEIGHT SMALL DATASETS

THE ACCURACY CHALLENGE

- AT 1%, NOT ALL REDUCED DATASETS AGREE
- MUST INCLUDE MISSING HIGHER-ORDER CORRECTION UNCERTAINTIES
- INCLUDE EW CORRECTIONS
- GO BEYOND K-FACTORS
- WEIGHT SMALL DATASETS

STAY TUNED FOR NNPDF4.1!!