

PARTON DISTRIBUTIONS: RECENT DEVELOPMENTS

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN



UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA



Istituto Nazionale di Fisica Nucleare

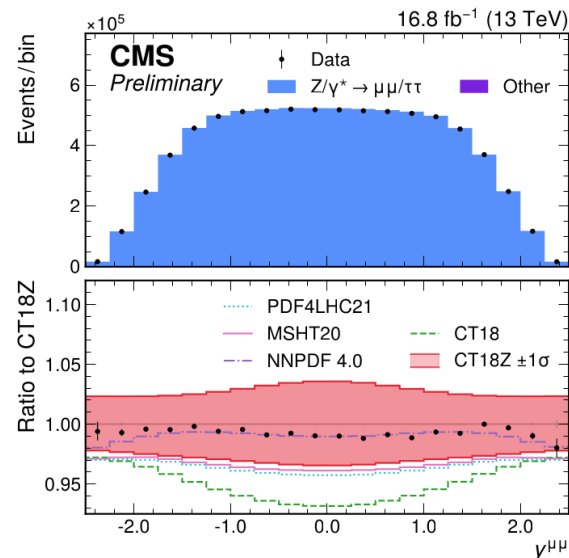
QCD@LHC

Oxford, 28 September 2026

THE IMPACT OF PDFs

W MASS DETERMINATION

Parton Distribution Functions



- **Good:** PDF sets are accompanied by uncertainty models with well defined correlations across phase space and between processes
- **Bad:** Different PDFs don't necessarily agree within their uncertainties
- Missing higher order uncertainties, resummation corrections in predictions usually not included
 - Partly mitigated by tolerance factors, etc

THIS TALK: **DISCLAIMER**

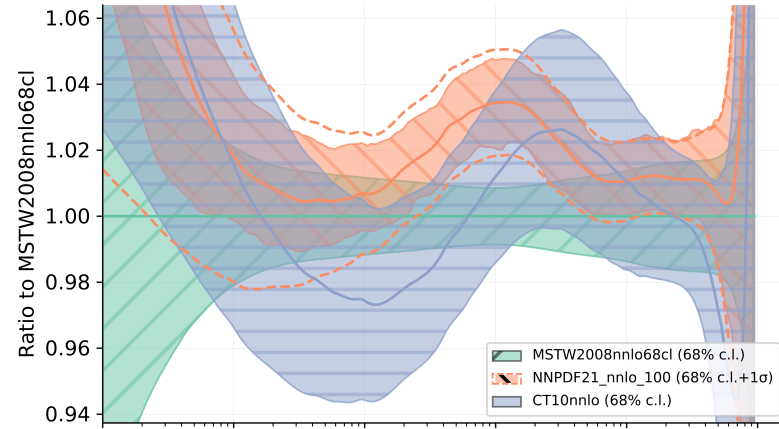
- a **personal view** on relevant issues
- **no attempt** at an exhaustive review $\Rightarrow$ see DIS conference, PDF4LHC meetings,...
- only unpolarized collinear PDFs, lattice results not covered

PDFs IN TIME

THE UP QUARK

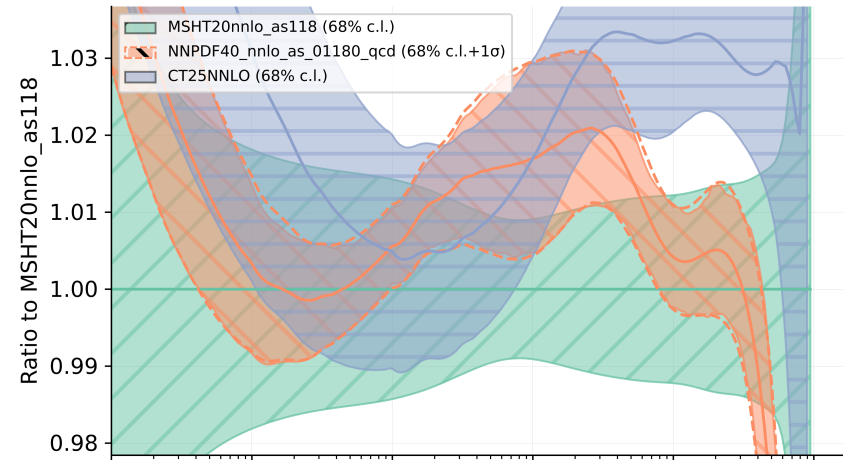
~2012

u at 100.0 GeV

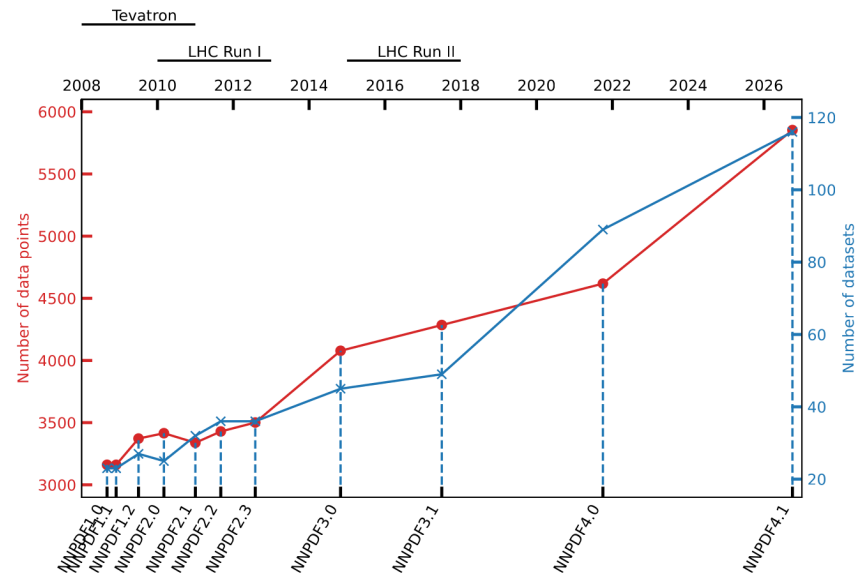


~2022

u at 100.0 GeV



THE DATA



- $\sim \times 2$ REDUCTION IN UNCERTAINTIES
- $\sim \times 1.5$ INCREASE IN # OF DATAPOINTS

CURRENT AND FUTURE “GLOBAL” PDF SETS

- **CT**
 - CT18 (NNLO)
 - * release on LHAPDF August 2019; preprint December 2019
 - * paper published January 2021
 - CT25 (NNLO): partially available on CT website since few months
- **MSHT**
 - MSHT20 (NNLO)
 - * release on LHAPDF & preprint December 2020
 - * paper published April 2021
 - $N^3\text{LO}$
 - * release on LHAPDF & preprint July 2022
 - * paper published March 2023
- **NNPDF**
 - NNPDF4.1 (NNLO)
 - * release on LHAPDF September 2021 & preprint September 2021
 - * paper published May 2022
 - $N^3\text{LO}$
 - * release on LHAPDF & preprint July 2024
 - * paper published July 2024
 - NNPDF4.1 ($N^3\text{LO}$): in preparation, due 2027

SOME MORE SPECIALIZED PDF SETS

- **ABMP**: simultaneous fit of QCD & EW parameters
- **ATLASPDF**: cross-correlation between ATLAS datasets
- **JAM**: large x , low scale JLAB data, higher twists

THE PROBLEM: PDF UNCERTAINTIES

- **DATA:** CORRELATIONS & CROSS-CORRELATIONS BETWEEN EXPERIMENTS
- **MODEL:** PDF IS ARE THE OUTPUT OF AN ILL-POSED REGRESSION PROBLEM
- **THEORY:** MISSING HIGHER ORDERS, TREATMENT OF HEAVY QUARKS (INCL. PARAMETRIZED CHARM), HIGHER TWISTS AND POWER CORRECTIONS, RESUMMATION, PHOTON PDF & EW CORRECTIONS

UNCERTAINTY BREAKDOWN

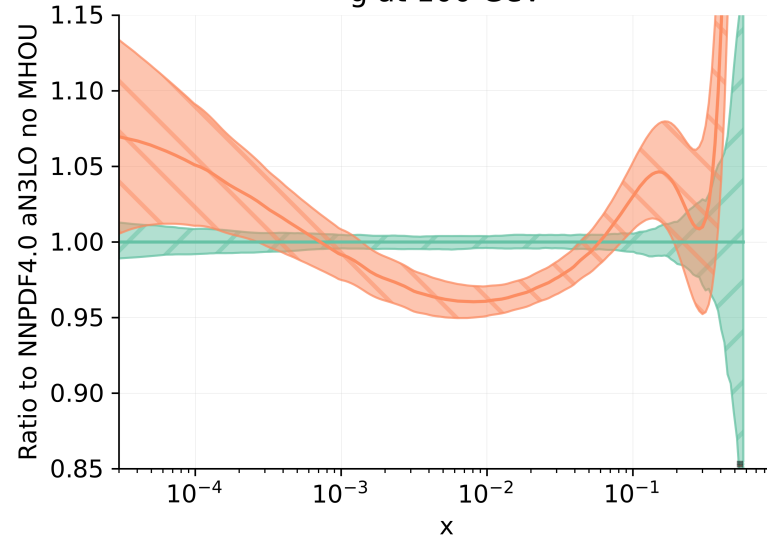
METHODOLOGY VS. DATA/THEORY

N^3 LO GLUON **MSHT20** VS **NNPDF4.0**

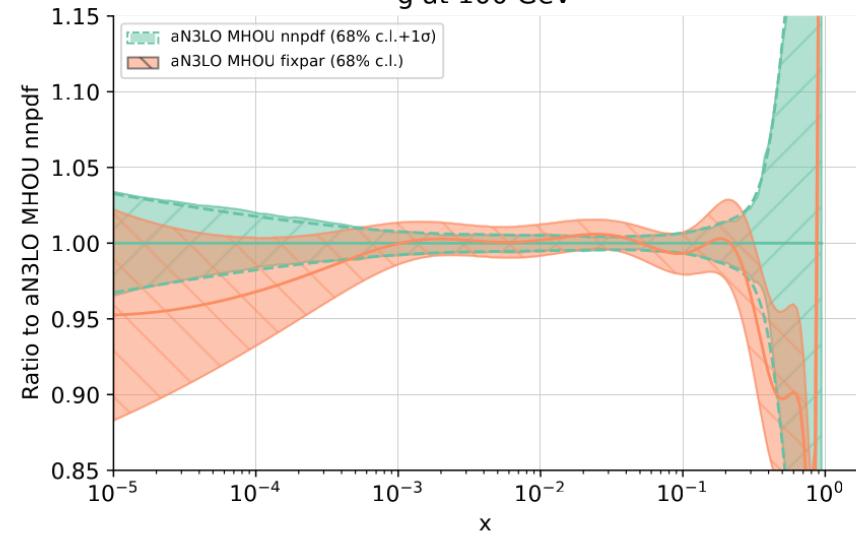
DIFFERENT METHOD., SAME DATA/THEORY

NATIVE PDFS

g at 100 GeV



g at 100 GeV



(Cruz-Martinez, Giani, Harland-Lang, 2026)

- **NATIVE**: PUBLISHED SETS
- **FIX**:
 - **SAME (NNPDF) DATA/THEORY**: N^3 LO TH, FITTED CHARM, MHOU, FONLL. . .
 - **DIFFERENT METHODOLOGY**: TOLERANCE, HESSIAN VS MC, POLY VS NN. . .

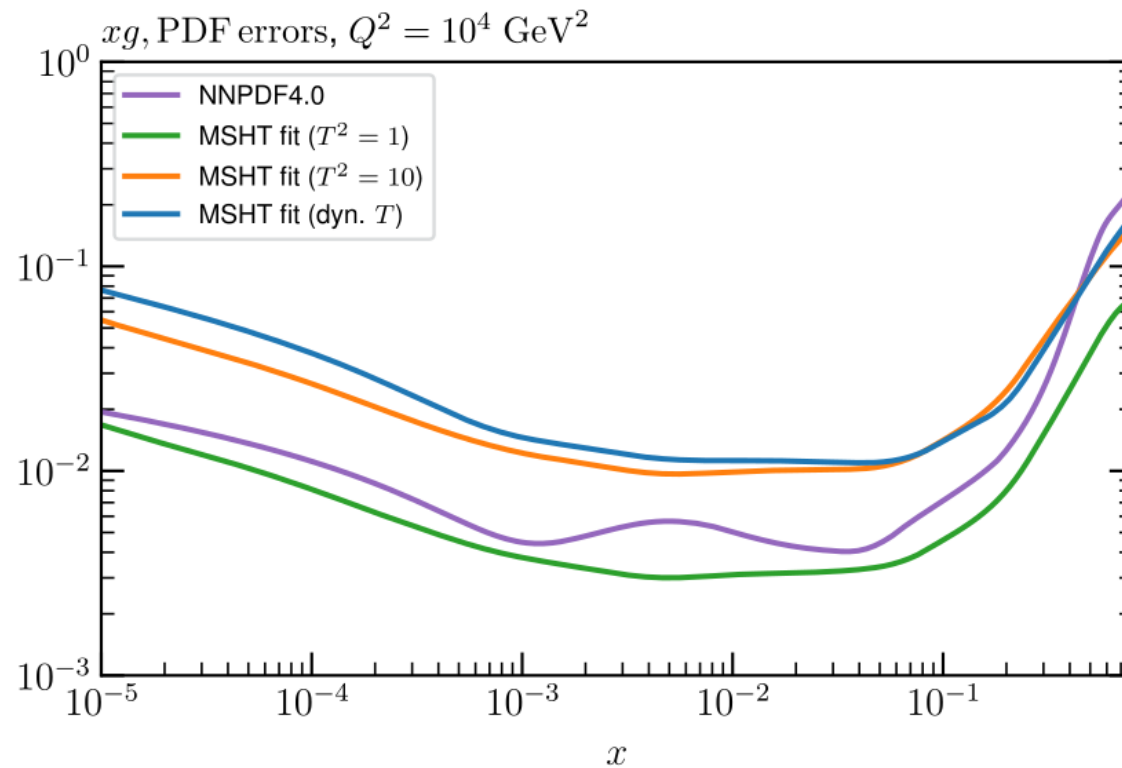
$\Rightarrow$ **GOOD AGREEMENT WITHIN UNCERTAINTY!**

METHODOLOGY & UNCERTAINTIES

THE TOLERANCE PROBLEM

- HESSIAN PDF UNCERTAINTY (CT, MS^{HT}) $\Rightarrow$ TOLERANCE
- $\Delta\chi^2 = T^2$, ACCOMMODATE FLUCTUATIONS BETWEEN DATASETS
- DYNAMICAL: DIFFERENT T ALONG DIFFERENT HESSIAN EIGENVECTORS

MSHT TOLERANCE IMPACT VS. NNPDF



(Harland-Lang, Cridge, Thorne, 2025)

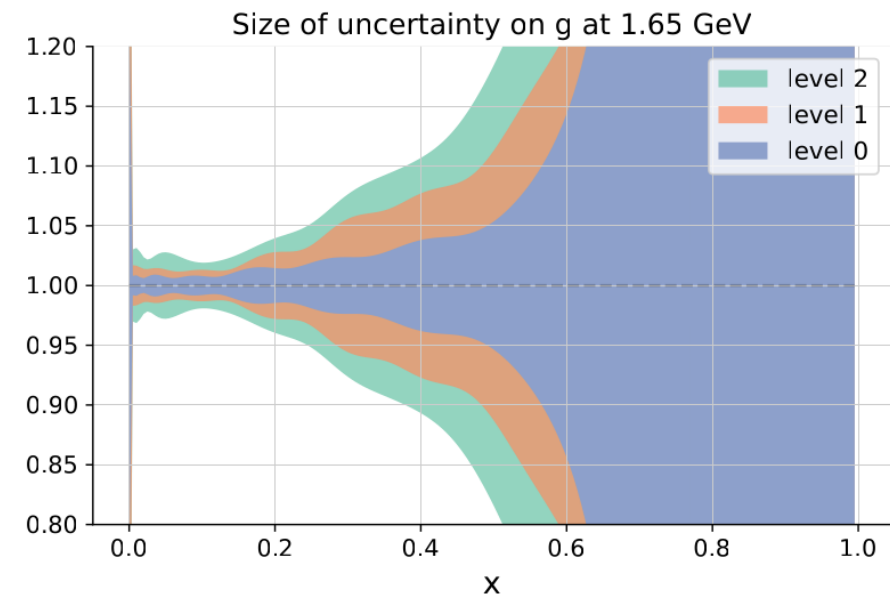
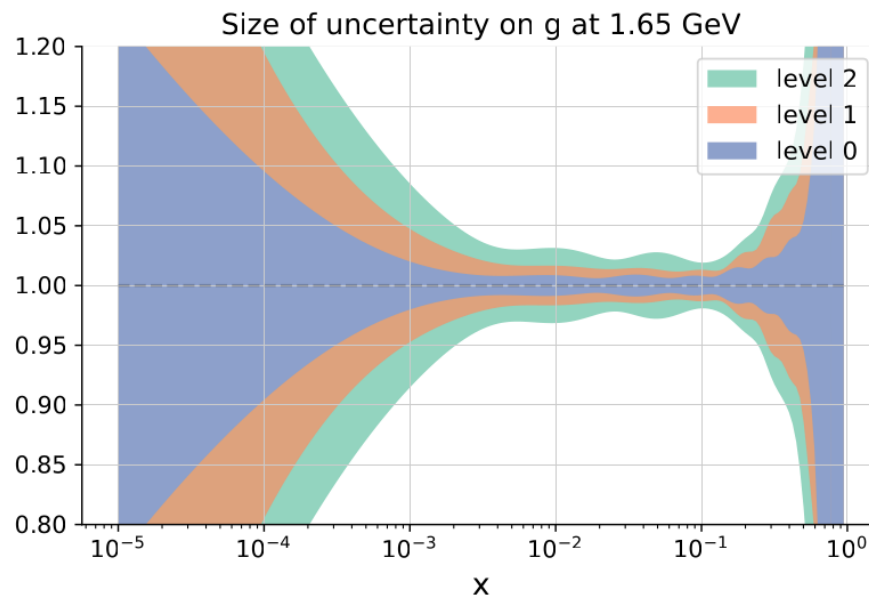
fixed NNPDF dataset, native NNPDF vs. MSHT methodology & theory

THE FUNCTIONAL UNCERTAINTY

ILL-POSED REGRESSION:

- INTERPOLATION/EXTRAPOLATION
- FUNCTIONAL: CHOICE OF FUNCTIONAL FORM
- DATA

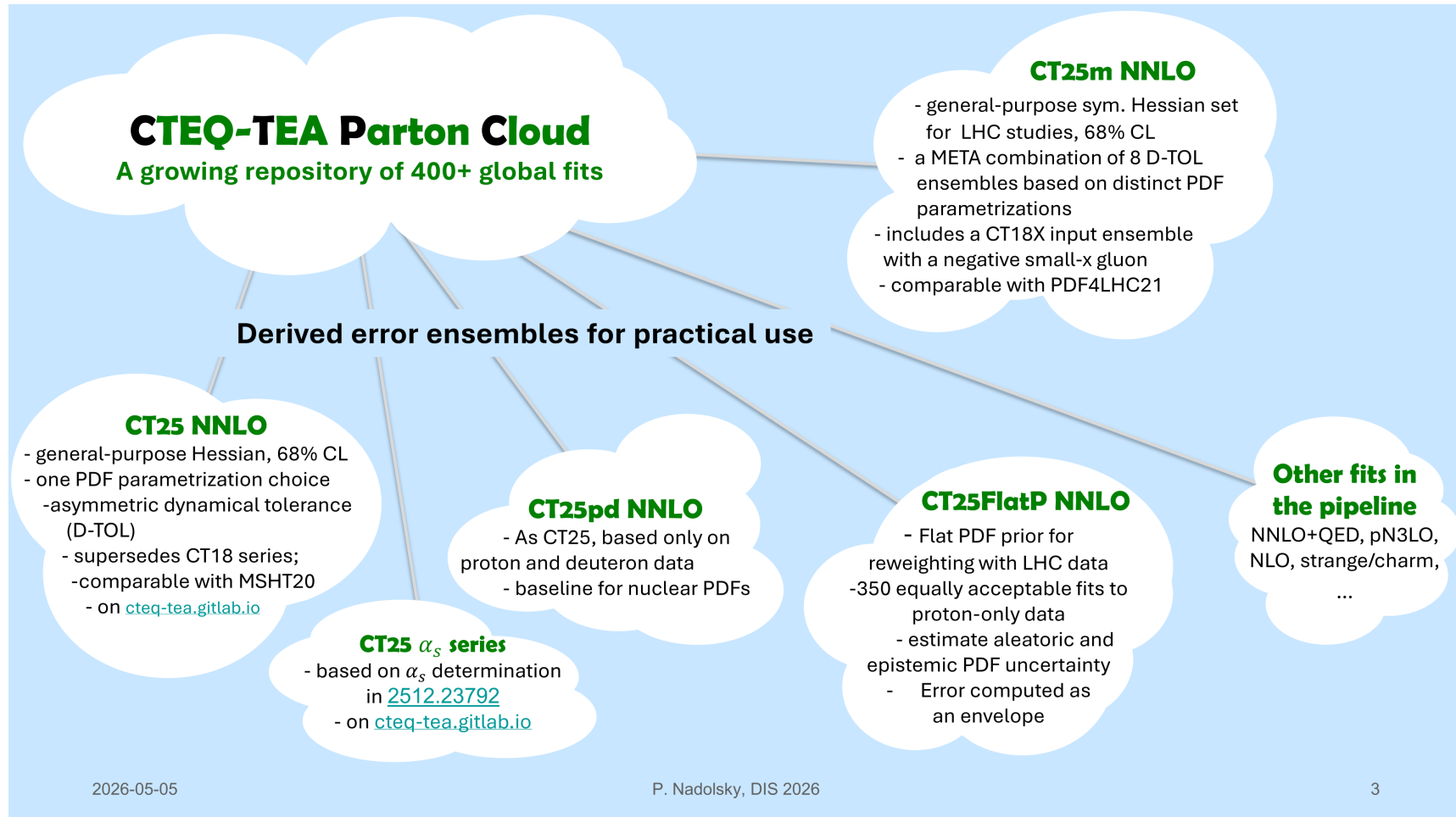
NNPDF4.0 CLOSURE TEST BREAKDOWN



(Del Debbio, Giani, Wilson, 2022)

THE FUNCTIONAL UNCERTAINTY & TOLERANCE

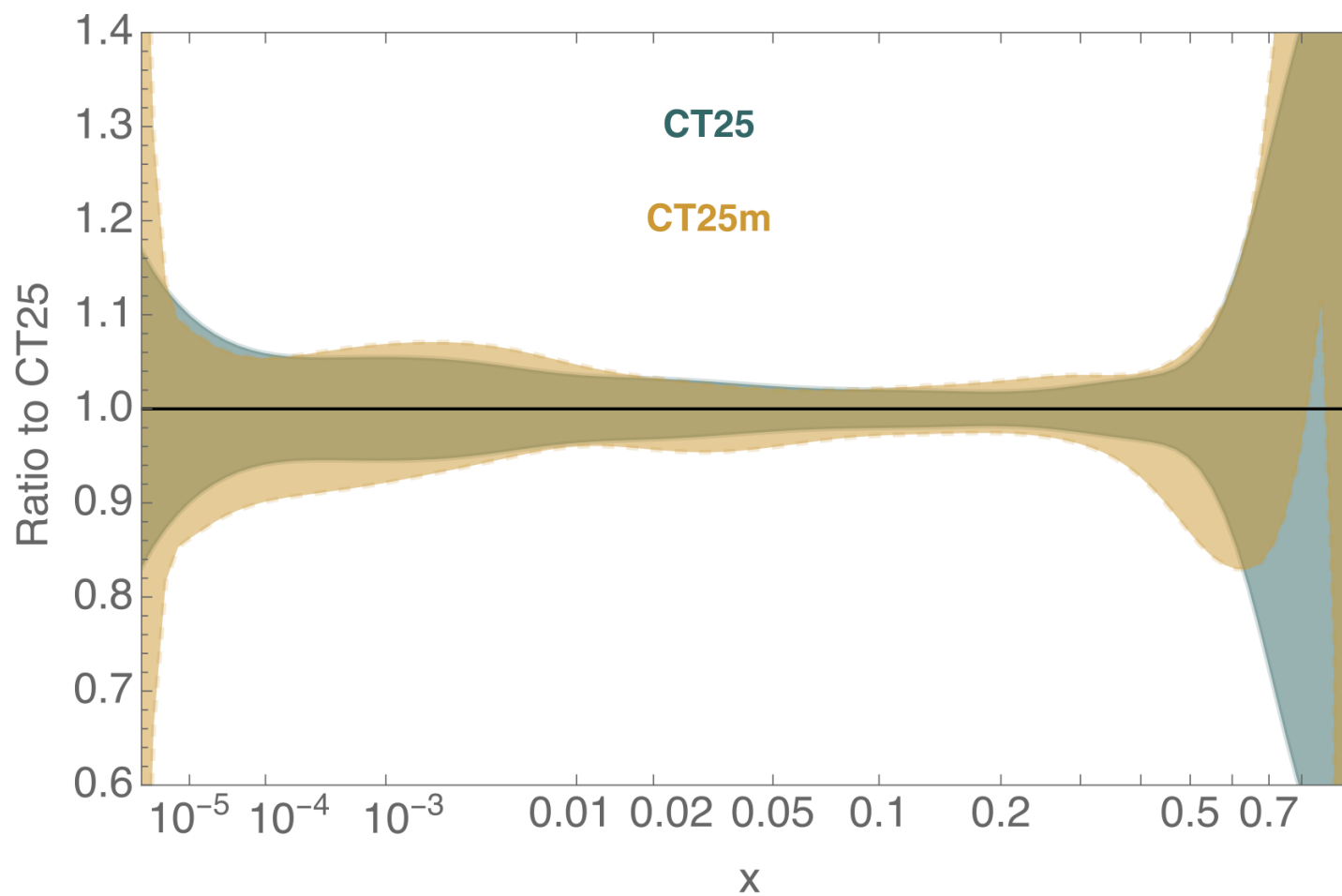
CT25 STUDY



(Nadolsky, DIS 2026)

ENSEMBLING VS. TOLERANCE

$x_d(x, Q)$ at $Q=2.0$ GeV (sym.err)

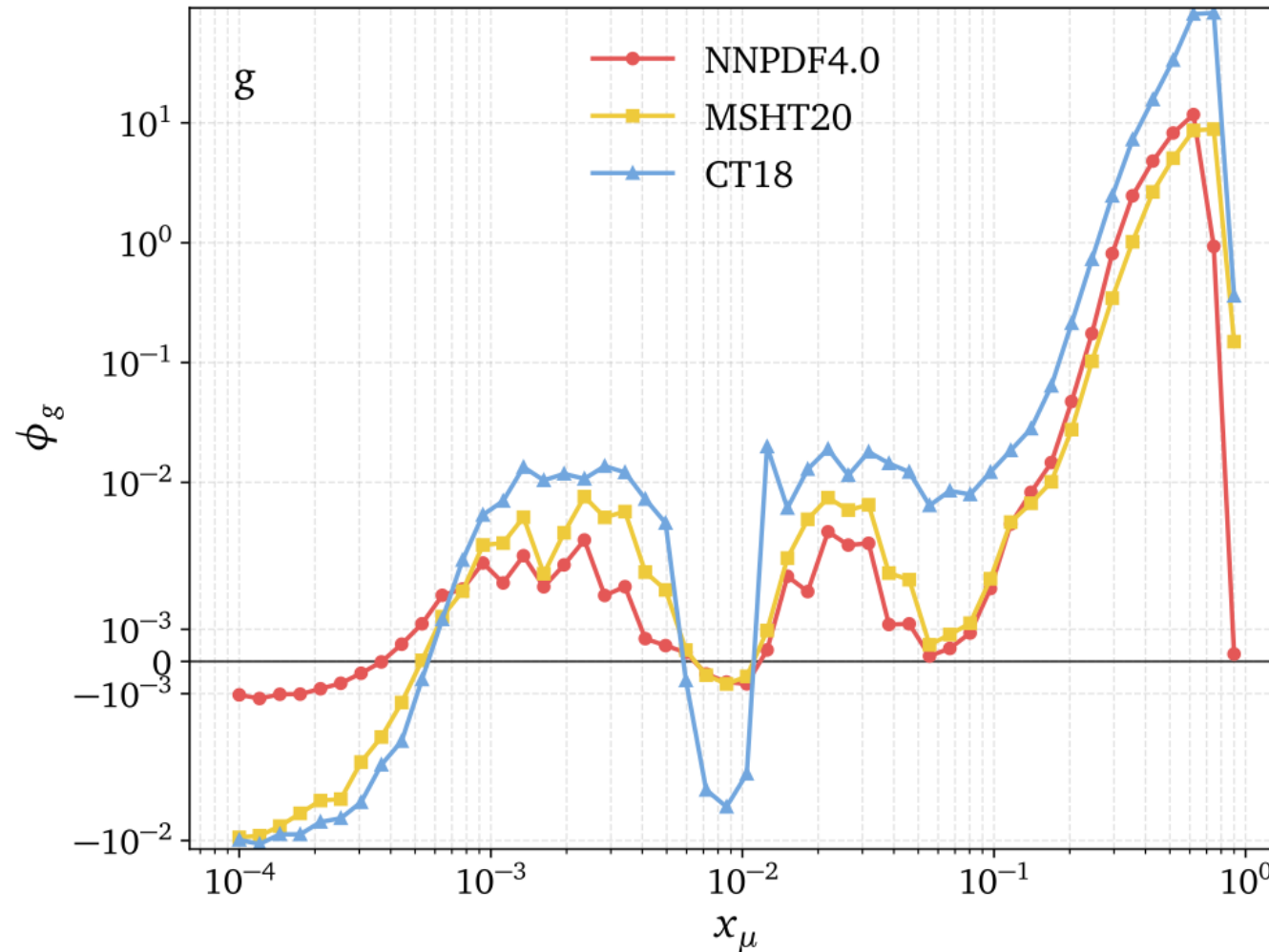


(Courtoy, DIS 2026)

TOLERANCE LARGELY ACCOUNTS FOR FUNCTIONAL UNCERTAINTY

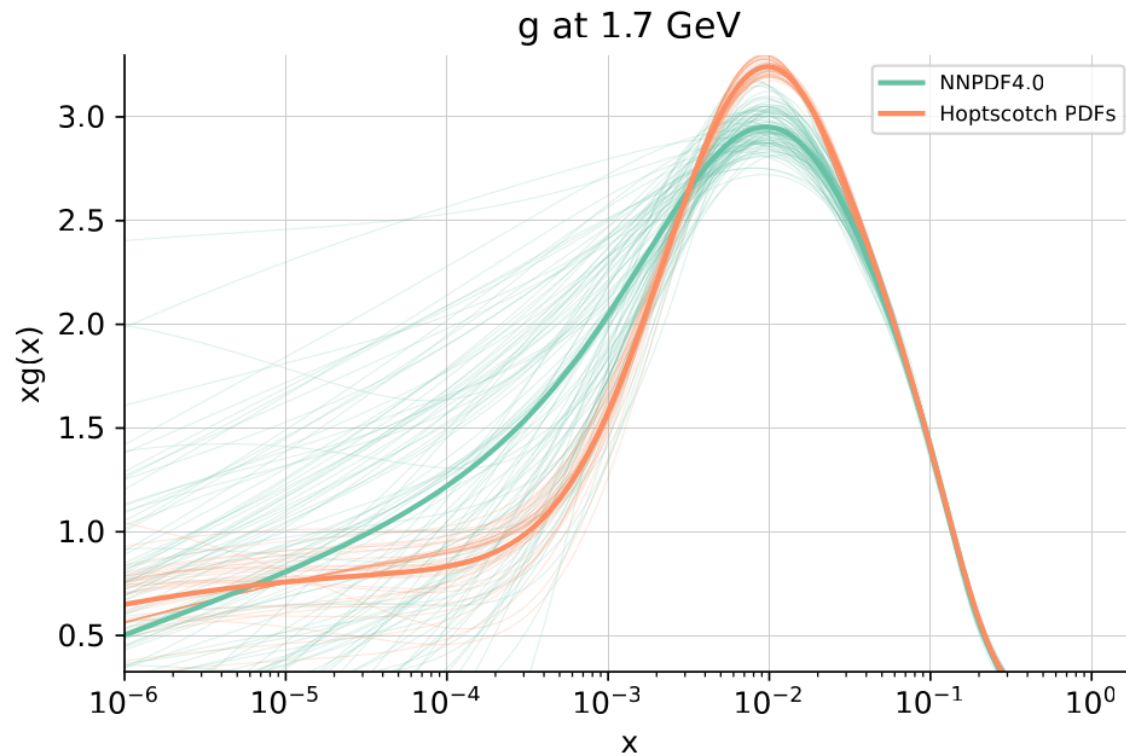
THE IMPACT OF ASSUMPTIONS

THE SHAPLEY INDEX: A MEASURE OF SENSITIVITY GLUON $\Leftrightarrow$ NNPDF4 DATA



(Bonnet-Guerrini, Carrazza, SF, Groenendijk, Piuri, Winterhalder , DIS 2026)

- LOSS OF SENSITIVITY FOR $x \sim 10^{-2}$
- SEEN IN “HOPSCOTCH” PDFs (Courtoy et al., 2022)

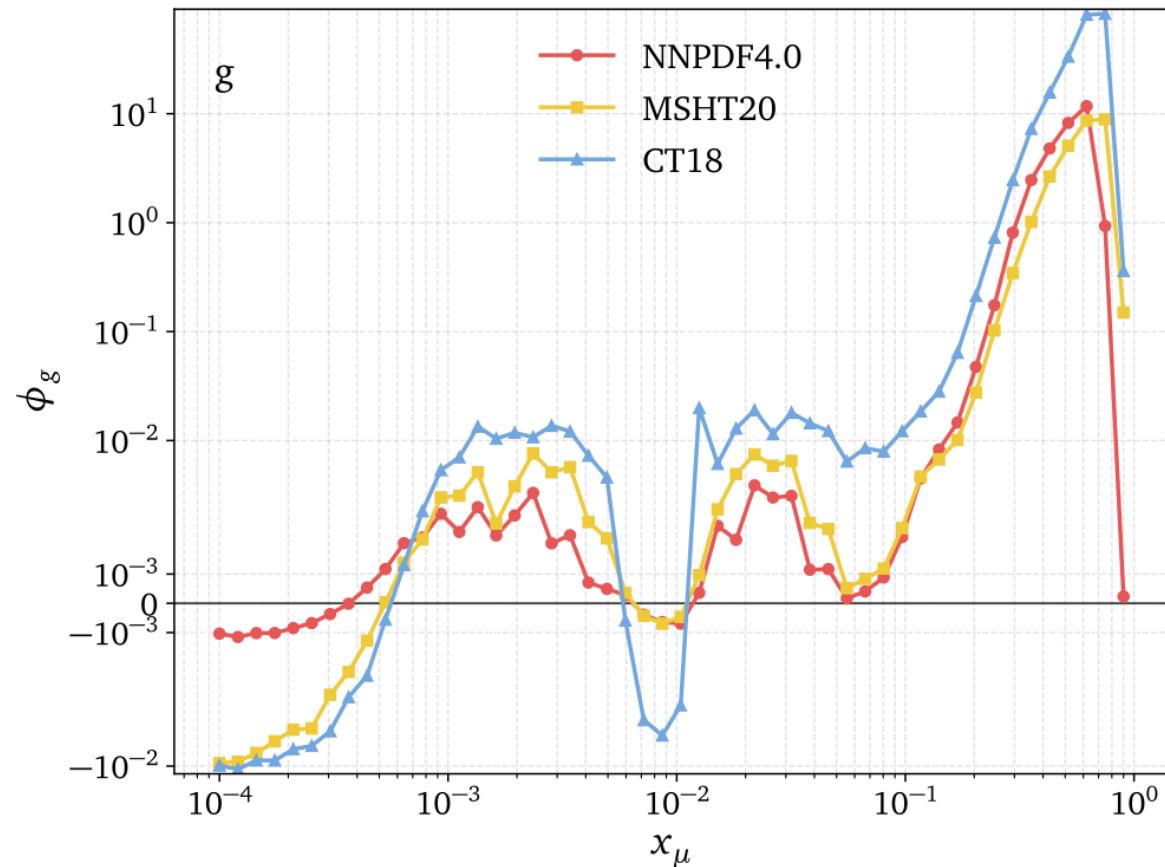


(NNPDF, 2022)

- NNPDF $\Rightarrow$ BUMP **DISFAVORED** BY ML MODEL (OVERLEARNING)
- CT/MSHT $\Rightarrow$ BUMP **FORBIDDEN** BY FUNCTIONAL FORM

THE IMPACT OF THEORY CONSTRAINTS

THE SHAPLEY INDEX: A MEASURE OF SENSITIVITY $\text{GLUON} \Leftrightarrow \text{NNPDF4 DATA}$

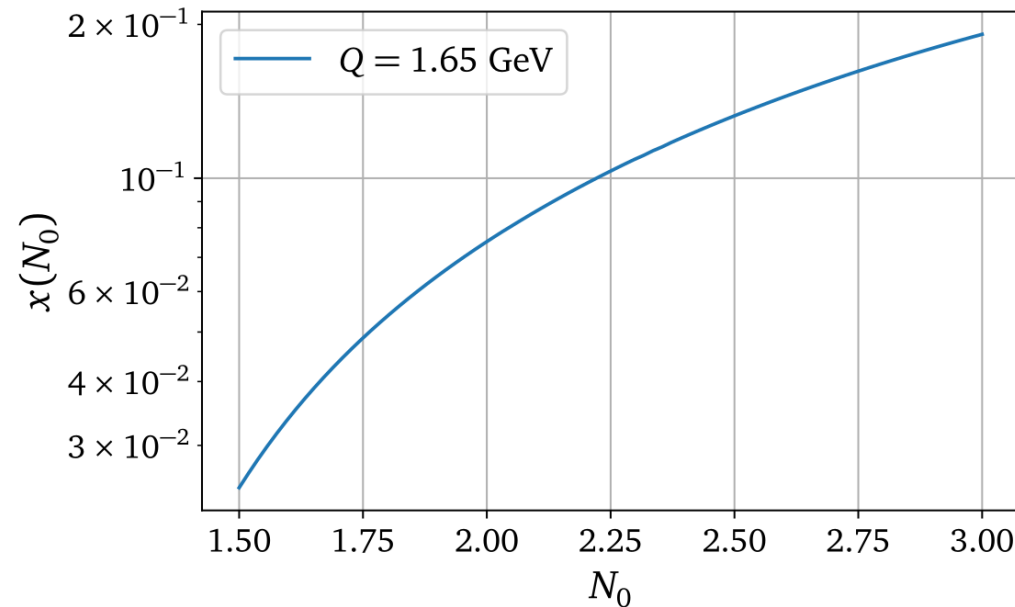


(Bonnet-guerrini, Carrazza, SF, Groenendijk, Piuri, Winterhalder , DIS 2026)

- LOSS OF SENSITIVITY FOR $x \sim 0.07$
- EXPLAINED BY MOMENTUM SR

x VALUE CORRESPONDING TO N BY SADDLE

Saddle point of gluon PDF



- ANOMALOUS DIMENSION MATRIX $\gamma_{ij}(2) \Rightarrow$ ZERO EIGENVALUE
- FLAT DIRECTION (GLUON)
- $N = 2$ CORRESPONDS TO $x = 0.07$ BY SADDLE MELLIN INVERSION

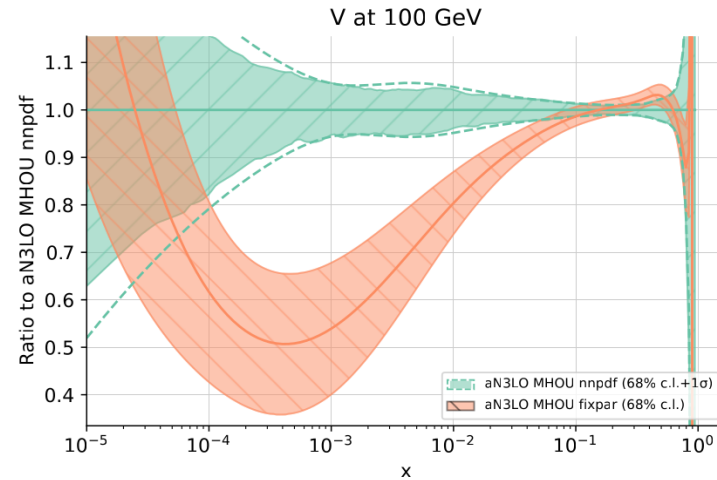
LESSON LEARNT

- A BSM BUMP COULD BE HIDDEN
- DIFFERENT DATA NEEDED

HIDDEN ASSUMPTIONS

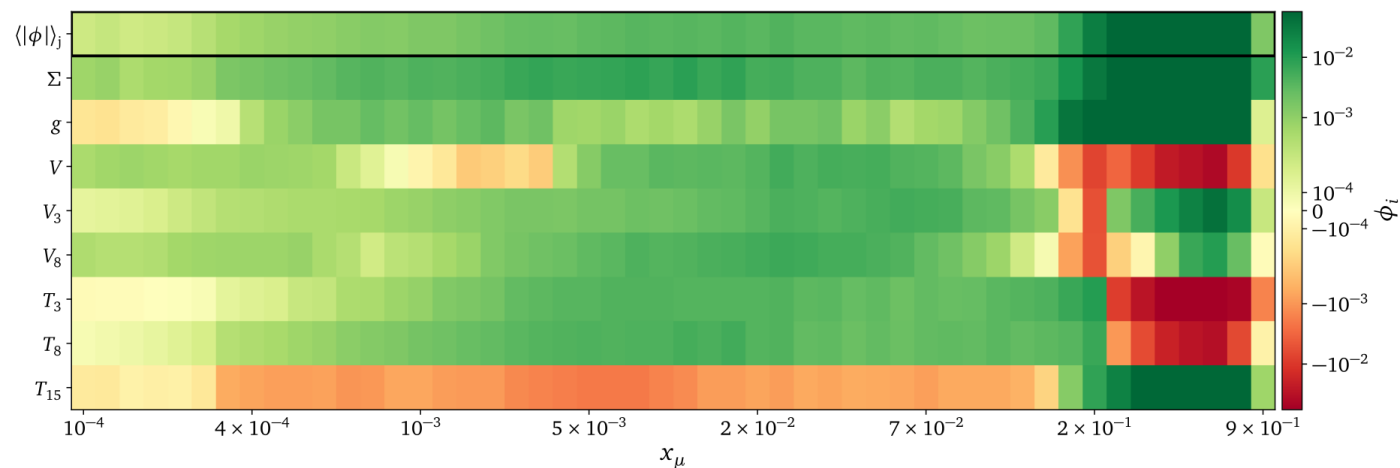
DIFFERENT METHODOLOGY, SAME DATA/THEORY AGAIN

THE TOTAL VALENCE (N³LO)



(Cruz-Martinez, Giani, Harland-Lang, 2026)

THE SHAPLEY ANALYSIS



- VALENCE LOSS OF SENSITIVITY
- DIFFERENT METHODOLOGIES LOCK ONTO DIFFERENT MININIMA

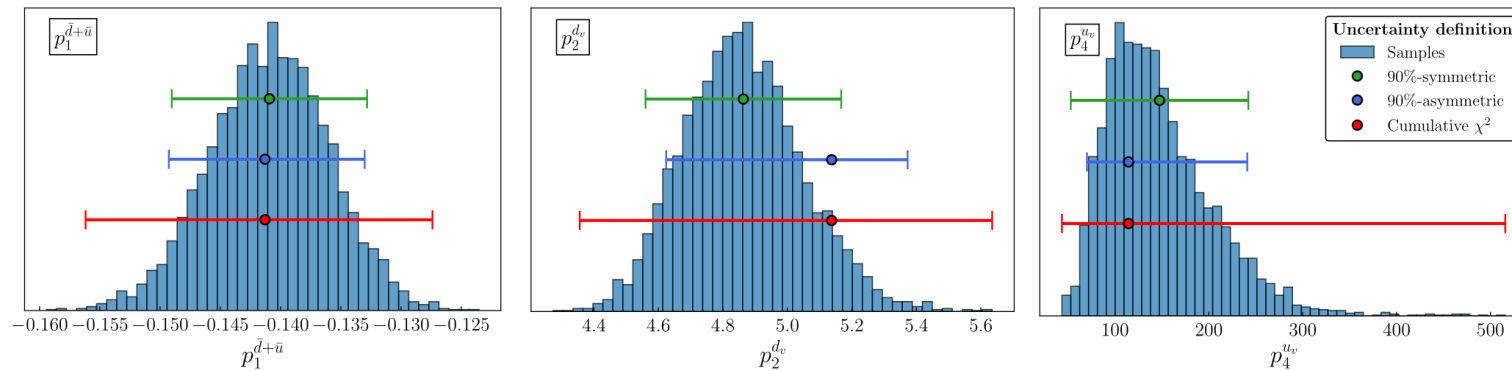
WHAT'S THE PROBABILITY DISTRIBUTION?

- PERFORM **BAYESIAN INFERENCE** FROM DATA USING SIMPLE PARM & PRIOR
- DETERMINE **POSTERIOR PROBABILITY** DISTRIBUTION (MCMC) (Kusina et al., 2026)

$$xf_i(x, Q_0) = p_0^i x^{p_1^i} (1 - x)^{p_2^i} (1 + p_3^i \sqrt{x} + p_4^i x)$$

PARAMETRIZATION

PROBABILITY DISTRIBUTIONS



(Risse, Derkashanian, Ježo, Kovařík, Kusina, 2026)

- SOME PARAMETERS **NONGAUSSIAN**
- **EFFECTIVE TOLERANCE** $T^2 \sim 10$

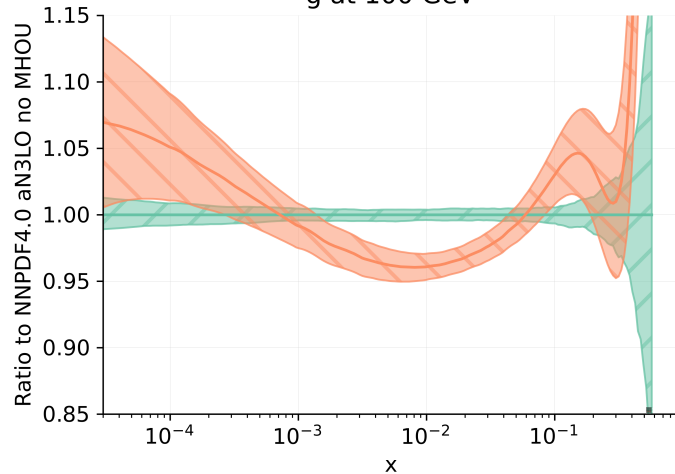
THEORY & UNCERTAINTIES

THE IMPACT OF N^3 LO

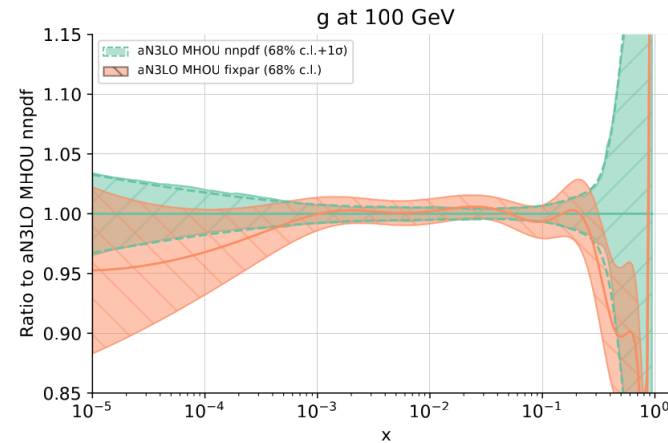
MSHT20 VS NNPDF4.0 GLUON, METHODOLOGY VS. DATA/THEORY AGAIN

N^3 LO

NATIVE PDFS
g at 100 GeV

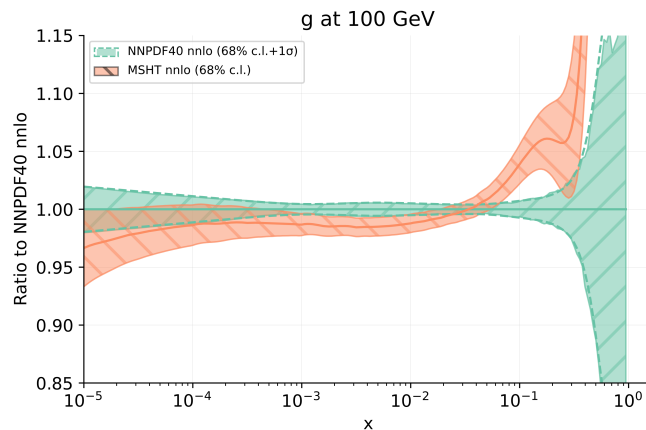


DIFFERENT METHOD., SAME DATA/THEORY

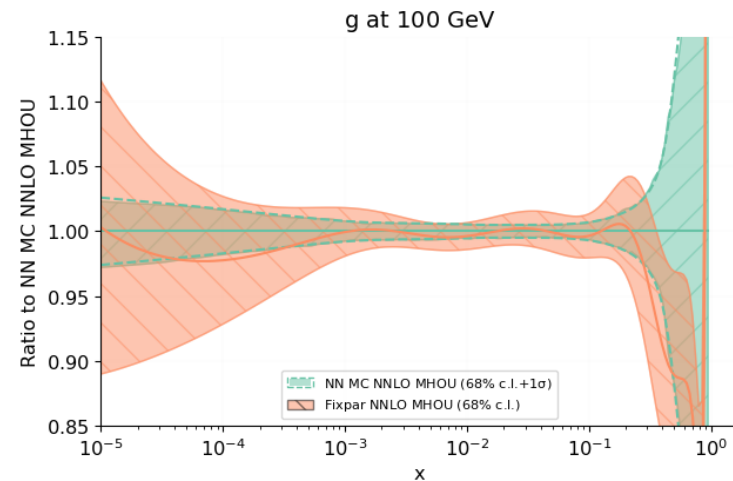


NNLO

(NO MHOU)



(WITH MHOU)

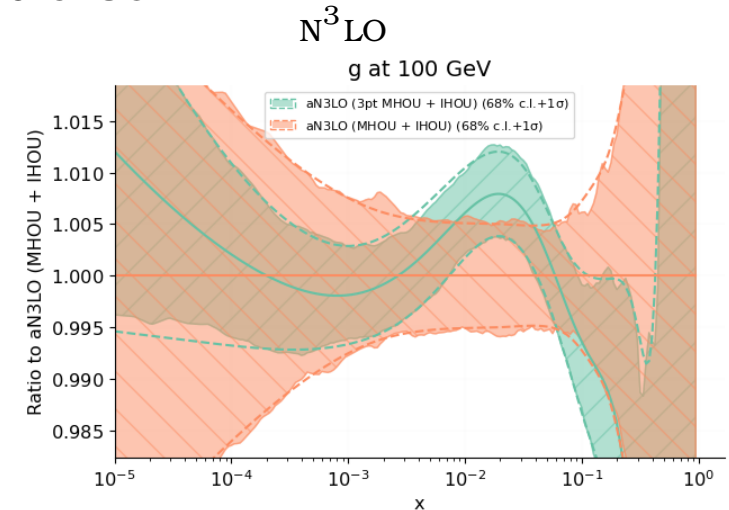
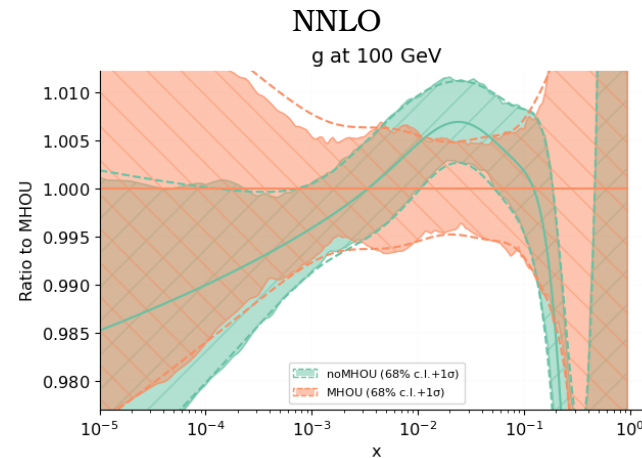
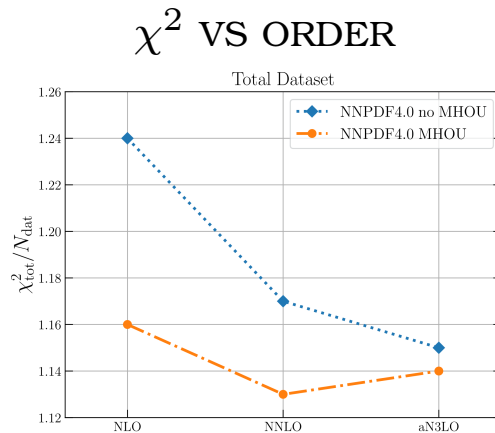


- N^3 LO DISCREPANCY LARGER THAN NNLO $\Rightarrow$ 2022 APPROXIMATION
- CURRENT N^3 LO THEORY EXACT: DIS, EVOLUTION, MATCHING;
HADRONIC PROCESSES MISSING $\Rightarrow$ MHOU

THE IMPACT OF MHOUS

- MSHT
 - INCLUDED ONLY AT N³LO THROUGH NUISANCE PARAMETERS
 - UNCERTAINTIES ESTIMATED BASED ON FORM OF MISSING TERMS
- NNPDF
 - INCLUDED AT ALL PERTURBATIVE ORDERS (2025) THROUGH COVMAT
 - MHOU COVMAT ESTIMATED THROUGH SCALE VARIATION
 - N³LO IHOU COVMAT FROM APPROX. SPREAD
- OTHER PDF SETS: NOT INCLUDED

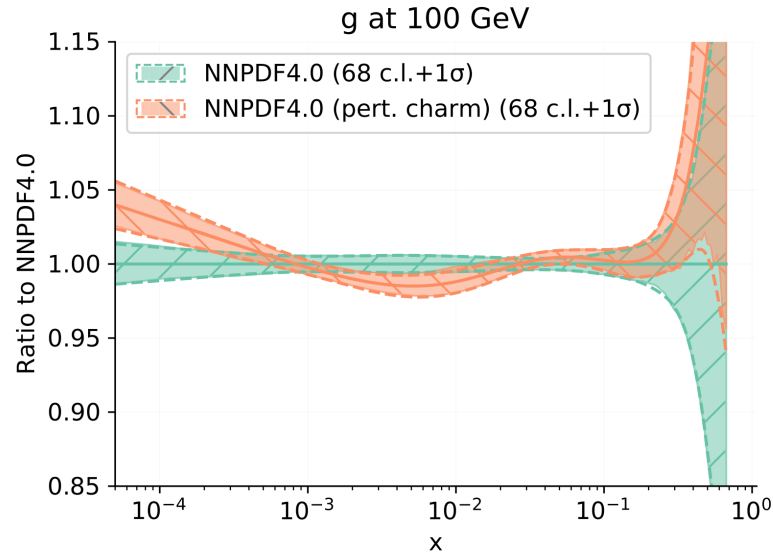
NO MHOU RATIO TO MHOU NNPDF4.0 GLUON



- IMPACT ON CENTRAL VALUE
- MHOU: N³LO (SLIGHTLY) SMALLER THAN NNLO
- N³LO: IHOU SMALLER THAN MHOU
- N³LO REQUIRED FOR HADRONIC PROCESS CONSISTENCY

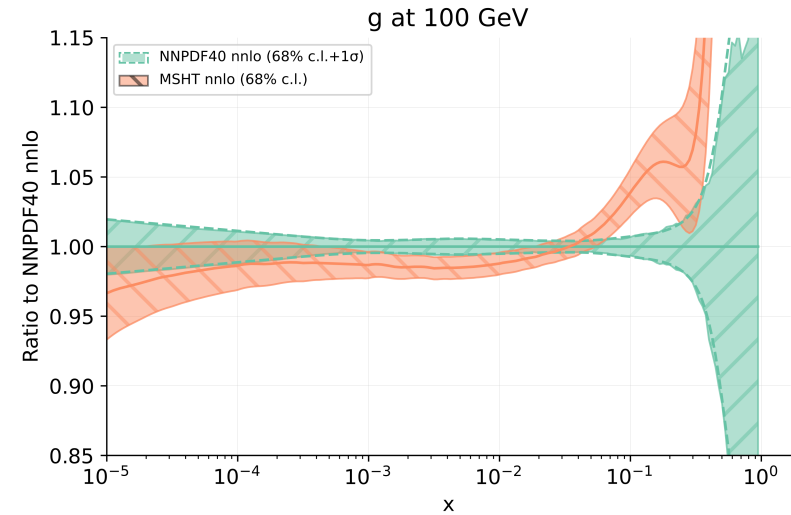
THE IMPACT OF FITTING CHARM

NNPDF4.0 FITTED VS PERT. CHARM



RECALL:

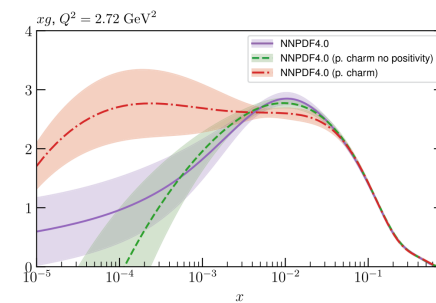
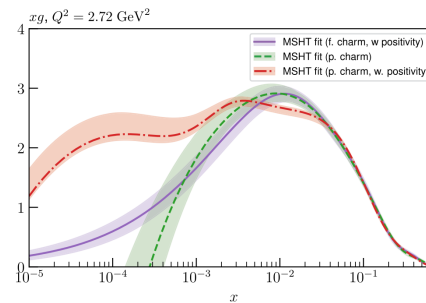
NNPDF4.0 VS MSHT20 NNLO NO MHOU



- **FITTED** CHARM $\Leftrightarrow$ **MISSING** HO **MATCHING**
- **EXPLAINS** RESIDUAL MSHT/NNPDF **DIFFERENCE** IN $x \gtrsim 0.003$ REGION

GLUON: **PERTURBATIVE** VS **FITTED** CHARM
MSHT NNPDF

- **MSHT VS NNPDF** FITTED CHARM **SAME IMPACT** ON GLUON
- **SAME** FITTED CHARM PDF, MSHT **SOMEWHAT LARGER UNC.**

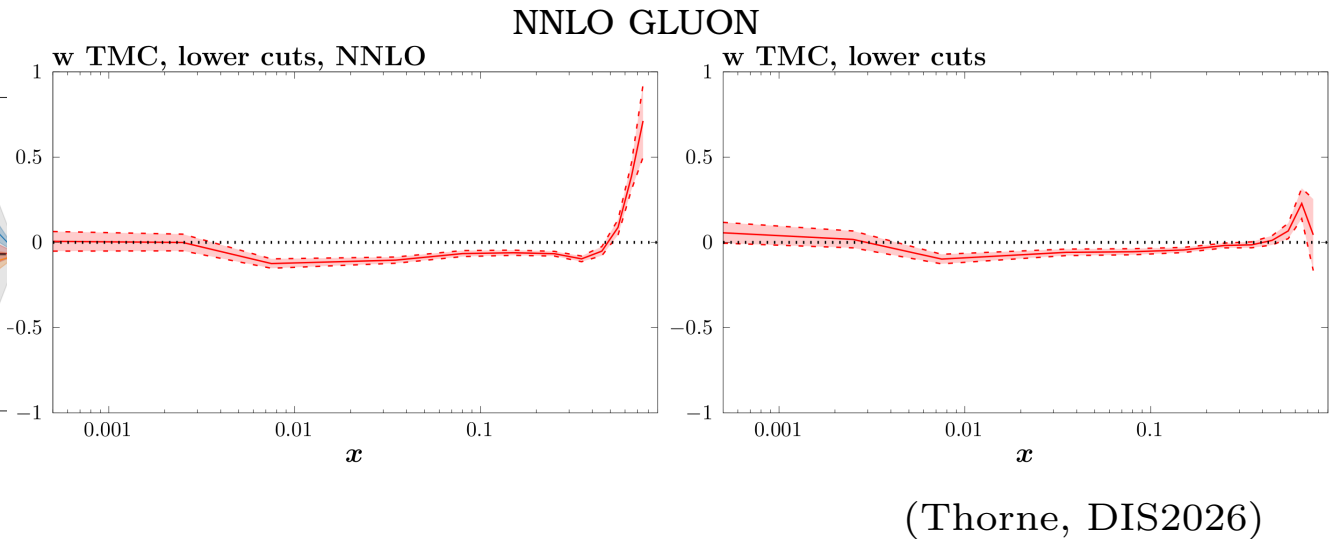
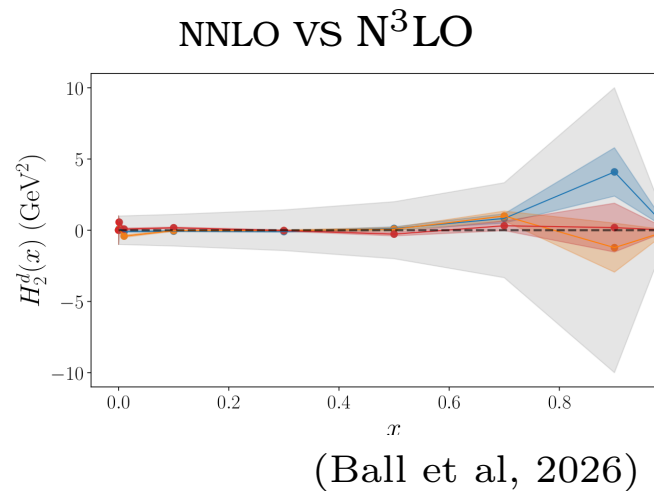


(Harland-Lang, Cridge, Thorne, 2025)

POWER CORRECTIONS

HIGHER TWISTS IN DIS

- EMPHASIZED IN PDF FITS INCLUDING LOW Q^2 DIS DATA (JAM, ABMP)
- TREATED WITH CUTS IN CURRENT GLOBAL FITS
- INCLUDED IN RECENT NNPDF, MSHT STUDIES: $\Rightarrow$ LARGELY REABSORBED AT $N^3\text{LO}$

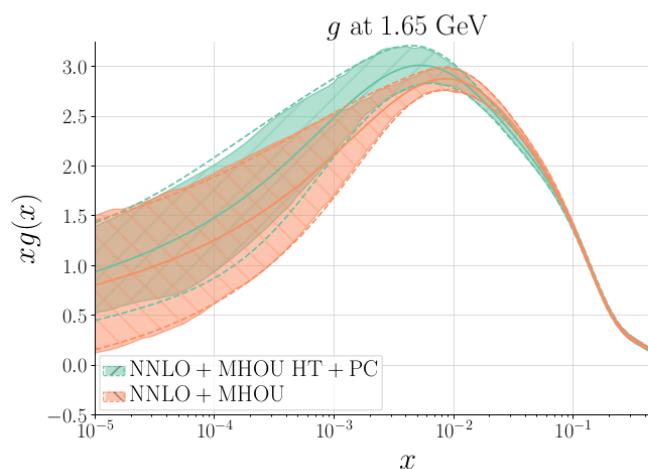


POWER CORRECTIONS

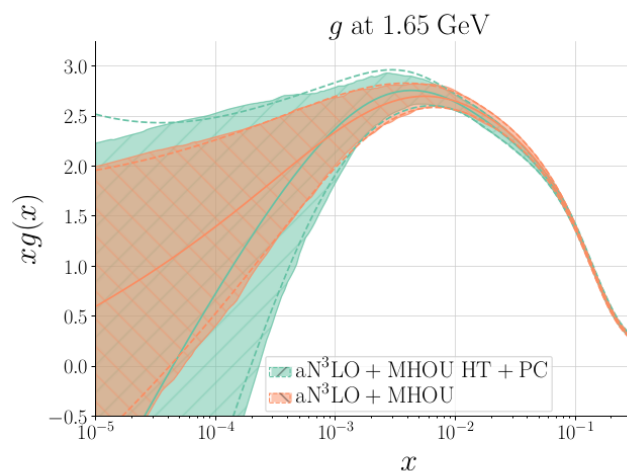
LINEAR PCs FOR JETS

- INCLUDED IN NNPDF STUDY
- MODERATE BUT **VISIBLE IMPACT** ON GLUON AT $10^{-3}x \lesssim x \lesssim 10^{-2}$

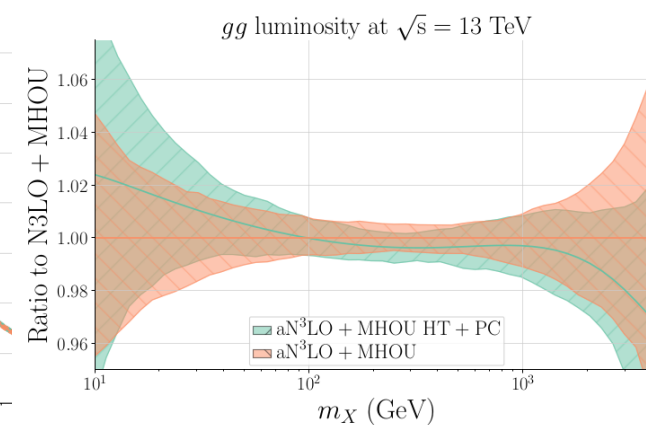
NNLO GLUON



N³LO GLUON



N³LO GG LUMI

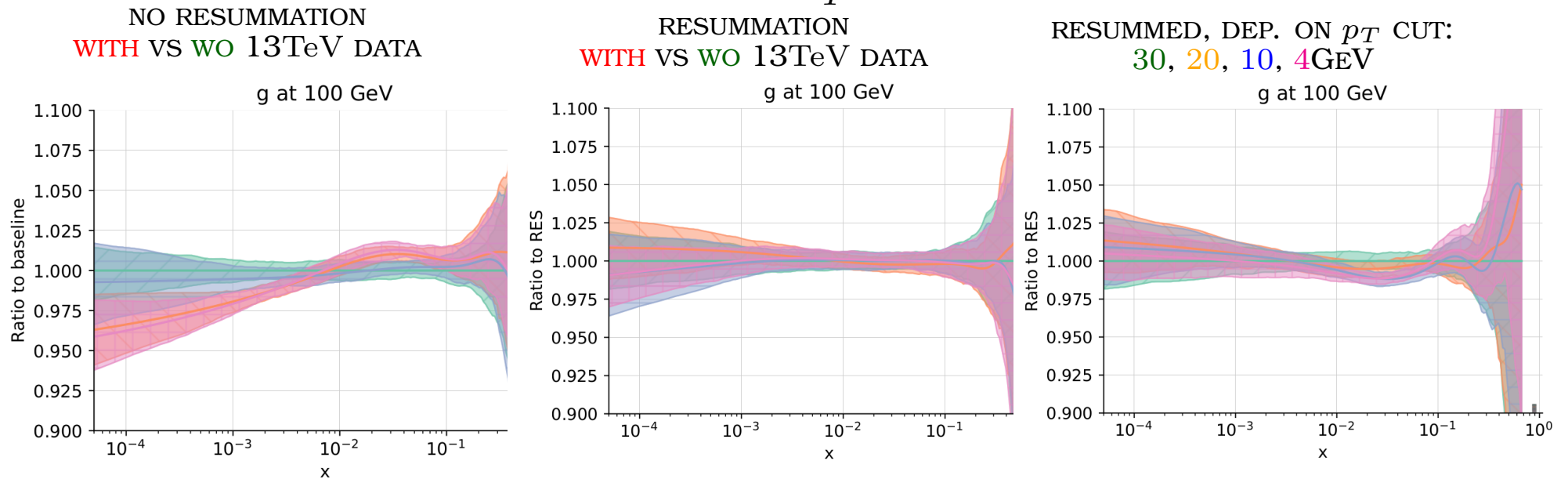


(Ball, Chiefa, Stegeman, 2026)

TRANSVERSE MOMENTUM RESUMMATION

- LITTLE IMPACT WITH 8TeV Z p_T DATA & $p_t > 30$ GeV
- NEEDED IN ORDER TO ACCOUNT FOR 13TeV ATLAS DATA
- GLUON STABLE UPON LOWERING p_t CUT DOWN TO $p_t > 4$ GeV

IMPACT OF Z P_T DATA

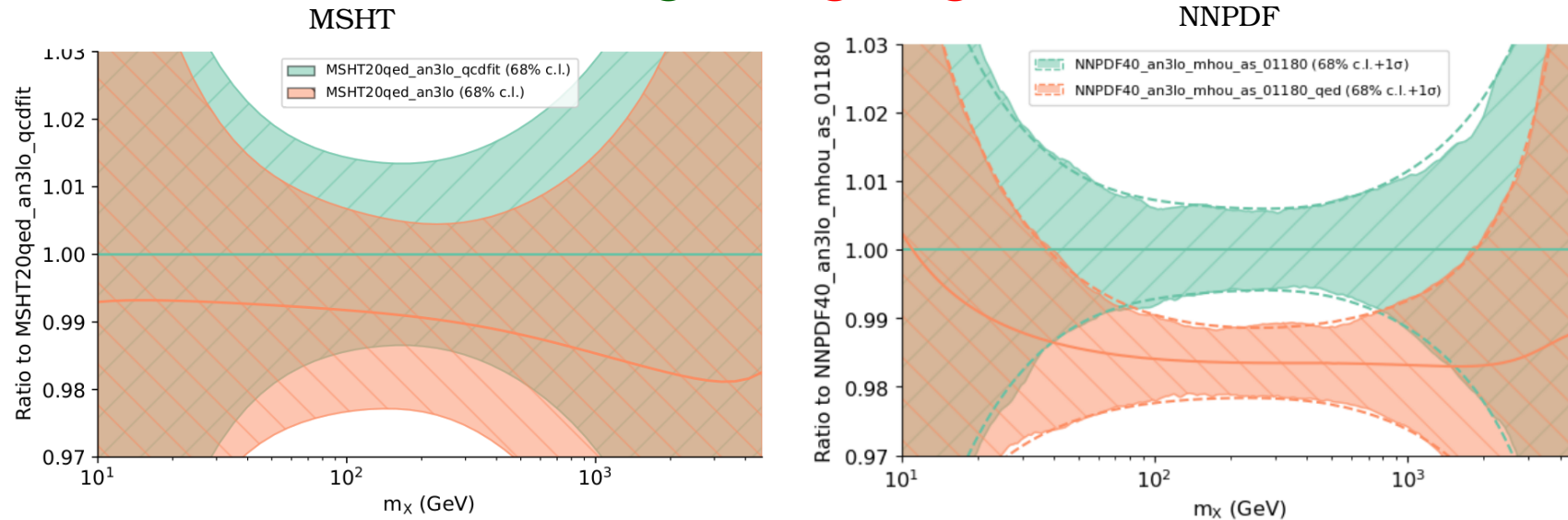


(Cruz-Martinez, Nocera, Rottoli, Torrielli, 2026)

QED & THE PHOTON PDF

- **COMPUTABLE** USING THE LUX FORMALISM (Manohar, Nason, Salam, Zanderighi 2017)
- **REDUCES GLUON MOMENTUM** FRACTION AT THE PERCENT LEVEL
- **PERCENT LEVEL IMPACT** ON GLUON LUMI
- RECENT LH STUDY $\Rightarrow$ MSHT-NNPDF **DIFFERENCE REDUCED** USING SMALLER (NNPDF3.1) DATASET

PURE QCD vs QCD+QED



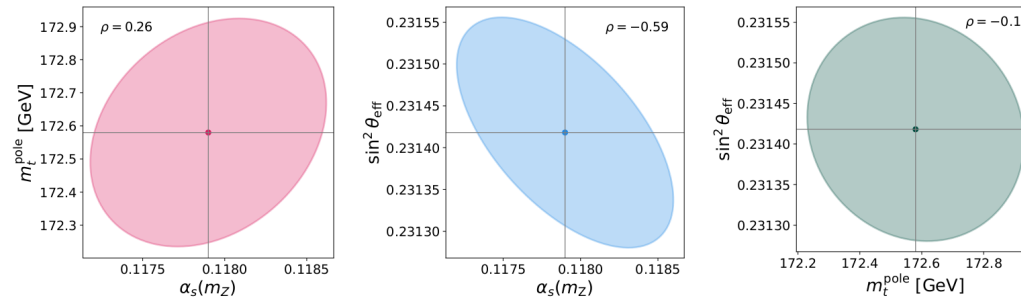
(Cridge, Cruz-Martinez, Huston, 2026)

PROGRESS

DETERMINING STANDARD MODEL PARAMETERS

- CANNOT DETERMINE SM PARAMETERS CORRELATED TO PDF WITHOUT REFITTING
- GLUON STRONGLY CORRELATED TO α_s , m_{top}

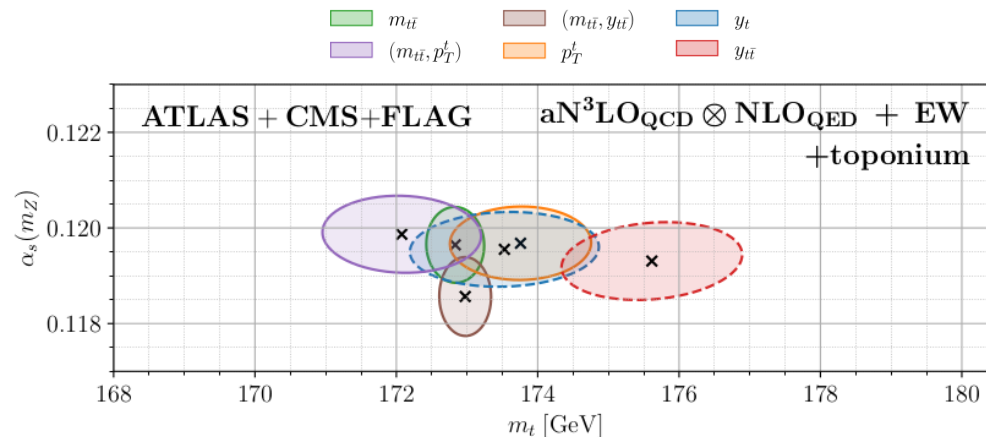
SIMULTANEOUS DETERMINATION OF PDFs, α_s , m_{top} , $\sin^2 \theta_w^{\text{eff}}$



(Morentin, Lipka, Limatola, Agrafiotis, Zenaiev, Shen, Moch, 2026)

- $\sin^2 \theta_w^{\text{eff}}$ CORRELATED TO α_s , m_{top}

SIMULTANEOUS DETERMINATION OF PDFs, α_s , m_{top}

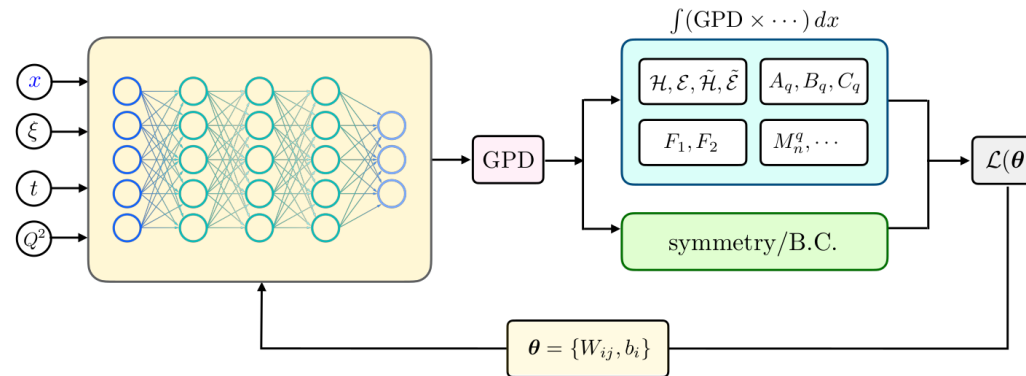


(Ball, ter Hoeve, Stegeman, 2026)

- GOOD AGREEMENT IF POOR OBSERVABLES DISCARDED

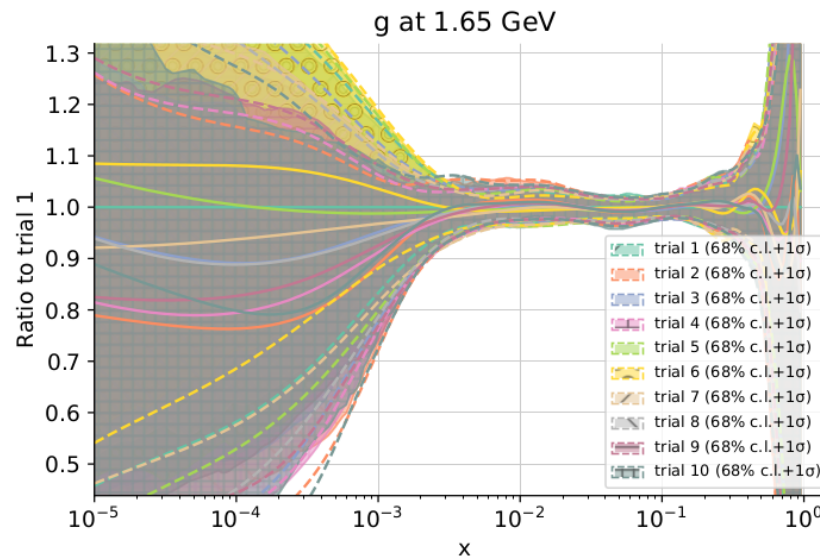
NEW IDEAS

- PHYSICS-INFORMED MACHINE LEARNING



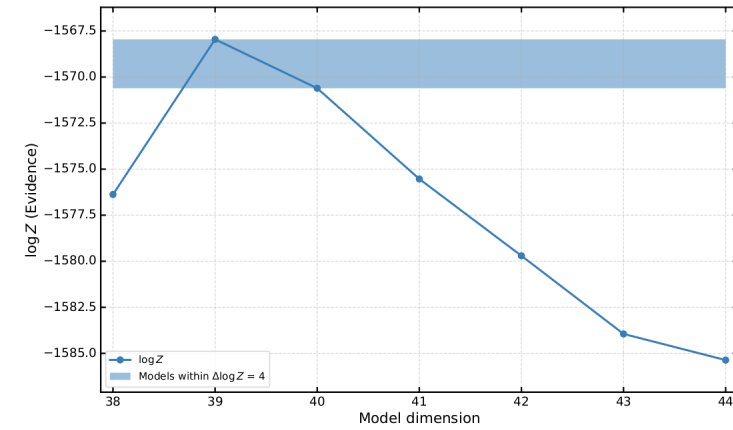
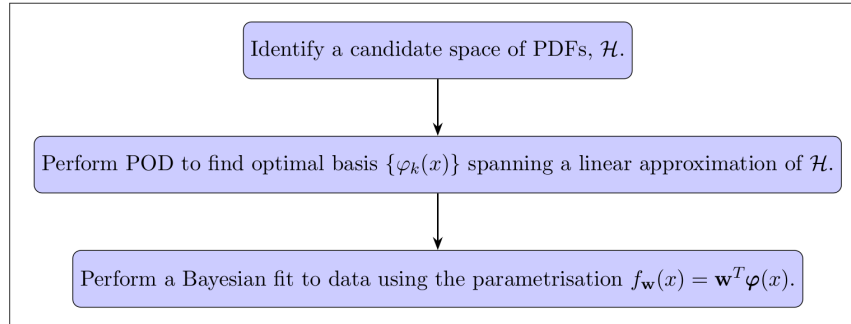
(EXCLAIM coll., 2026)

- ENSEMBLE HYPEROPTIMIZATION



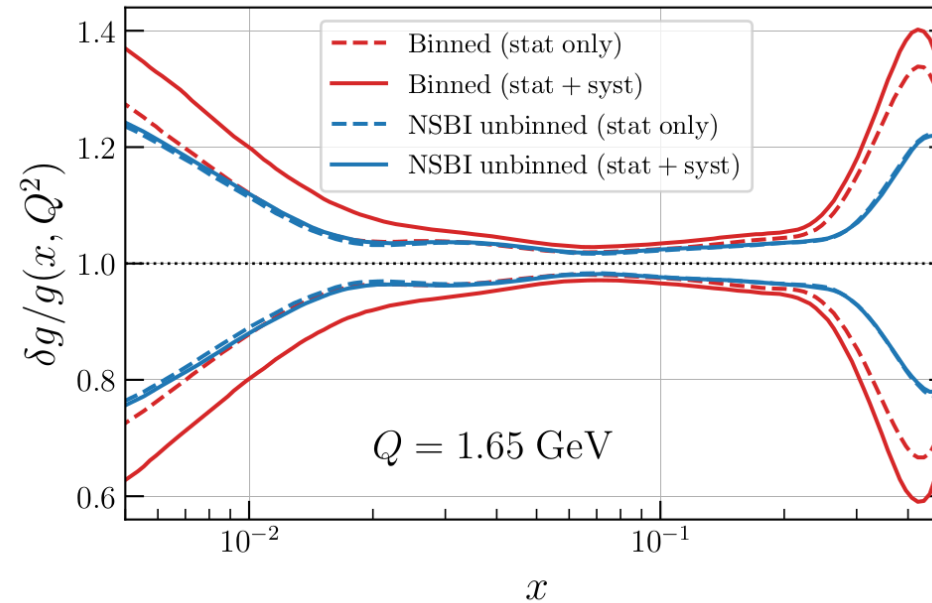
(Cruz-Martinez, Giani, Rabemananjara 2026)

- LINEAR BAYESIAN INFERENCE



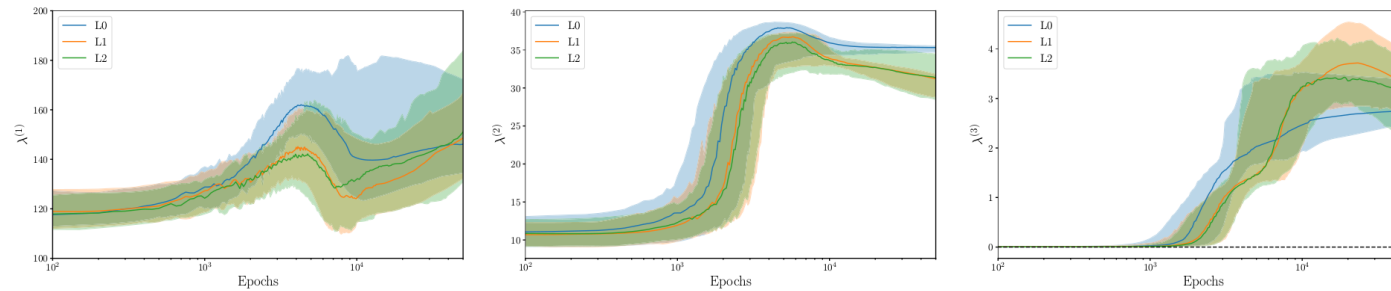
(Costantini, Mantani Moore, Ubiali, 2026)

- NEURAL SIMULATION-BASED INFERENCE



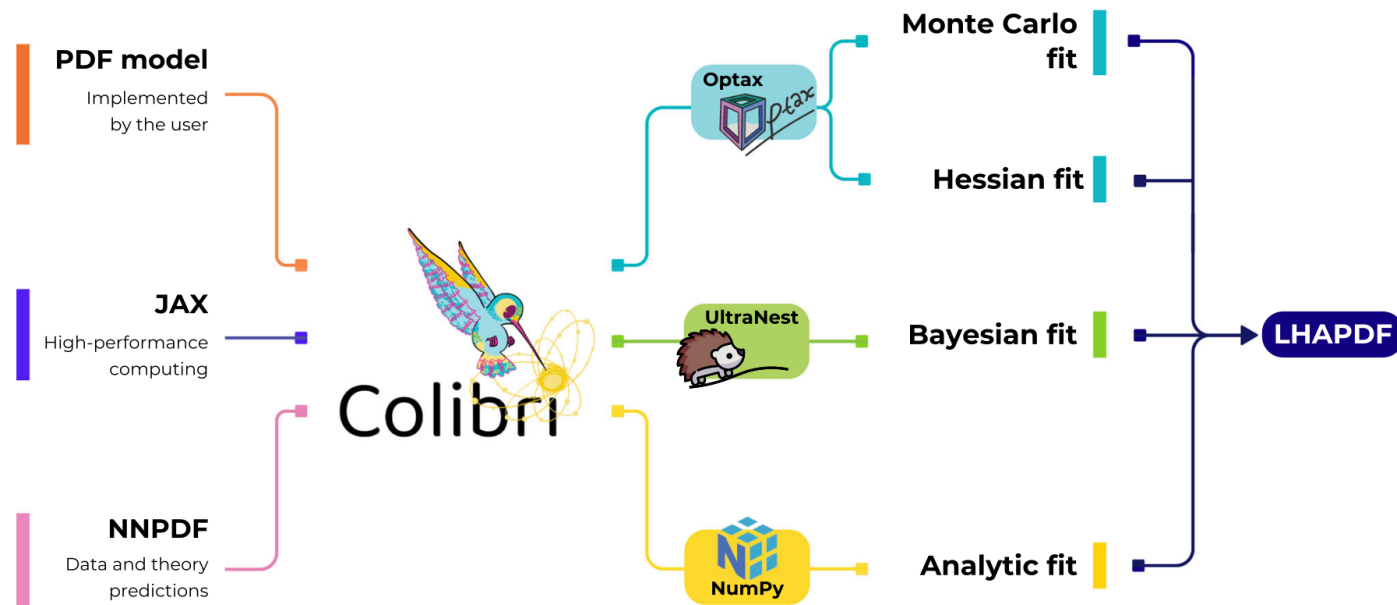
(Barrué et al, 2026)

- NEURAL TANGENT KERNEL



(Chiefa, Del Debbio, 2026)

- UNIVERSAL INFERENCE MACHINE



(Costantini, Mantani, Schütze Sánchez, Ubiali, 2026)

SUMMARY

- SUBOPTIMAL THEORY LEADS TO LARGE DISCREPANCIES
- PERCENT-LEVEL PRECISION

REQUIRES EXPLAINABLE METHODOLOGY