Exploring the impact of high-precision top-quark pair production data on the structure of the proton at the LHC

Alim Ablat,[1] Marco Guzzi[🟢],[2,*] Keping Xie,[3,†] Sayipjamal Dulat[🟢],[1] Tie-Jiun Hou[🟢],[4] Ibrahim Sitiwaldi[🟢],[1] and C.-P. Yuan[5]

[1]*School of Physics Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046 China*
[2]*Department of Physics, Kennesaw State University, Kennesaw, Georgia 30144, USA*
[3]*Pittsburgh Particle Physics Astrophysics and Cosmology Center, Department of Physics and Astronomy,*
University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
[4]*School of Nuclear Science and Technology, University of South China, Hengyang, Hunan 421001, China*
[5]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*

The impact of recent LHC top-quark pair production single differential cross section measurements at 13 TeV collision energy on the structure of the proton is explored. In particular, the impact of these high-precision data on the gluon and other parton distribution functions (PDFs) of the proton at intermediate and large partonic momentum fraction x is analyzed. This study extends the CT18 global analysis framework to include these new data. The interplay between top-quark pair and inclusive jet production as well as other processes at the LHC, is studied. In addition, a study of the impact of scale choice on the theory description of the new 13 TeV $t\bar{t}$ measurements is performed.

I. INTRODUCTION

The physics of the top quark, discovered by the D0 and CDF collaborations at the Tevatron proton-antiproton collider [1,2] in 1995, is central to a large number of precision programs for theory and experiments at the Large Hadron Collider (LHC) and its future upgrades.

In proton-proton collisions at the LHC, properties of the top quark are being thoroughly investigated as they are critical to many endeavors in high-energy physics, e.g., test the QCD theory and structure of the proton with unprecedented accuracy, unravel details of the electroweak (EW) sector of the Standard Model (SM) at high energies, and search for new physics interactions.

The mass of the top quark ($m_t = 172.69 \pm 0.30$ GeV from direct measurements [3]) is very close to that of the recently discovered [4,5] Higgs boson ($m_H = 125.25 \pm 0.17$ GeV) and quantum corrections to these masses are deeply related to one another. In particular, the mass of the top quark is an important ingredient in the determination of absolute stability conditions for the EW vacuum as top-quark mass radiative corrections drive the couplings of the

Higgs boson [6–14]. In addition, the top quark provides us with unique opportunities to search for signatures of new physics interactions at the TeV scale and beyond (see for instance Refs. [15,16] and references therein).

The top quark mainly decays into a real W boson and a b quark before hadronization occurs. Therefore, precision measurements of top-quark pair production cross sections allow us to set stringent tests on perturbative QCD and to explore the structure of the proton with higher precision. In fact, more than 90% of the production rate is due to the gluon-gluon fusion channel at the LHC with $\sqrt{S} = 13$ TeV of center-of-mass collision energy. This makes top-quark pair production a unique probe of proton's parton content, especially the gluon, at intermediate and large partonic momentum fraction x, which motivates this work.

The ATLAS and CMS collaborations performed high-precision measurements of both total inclusive and differential top-quark pair production cross section at $\sqrt{S} = 7$, 8 [17–22] and 13 TeV [23–28] of center-of-mass collision energy, at parton and particle level, in different channels (e.g., dilepton, semileptonic, all-hadronic, and lepton + jets) at different integrated luminosity (IL).

The higher precision LHC data samples obtained at 13 TeV are at the core of this analysis which aims at studying their impact on parton distribution functions (PDFs) of the proton in global QCD analyses using the CTEQ framework. High-precision theoretical calculations to predict top-quark pair production cross sections for a variety of kinematic distributions obtained with different bin resolutions are therefore crucial for this task.

*mguzzi@kennesaw.edu
†xiekeping@pitt.edu

Progress in theoretical calculations for $t\bar{t}$ cross sections.
Radiative corrections to heavy-quark production at the next-to-leading order (NLO) $\mathcal{O}(\alpha_s^3)$ in QCD have been known for a long time [29–34], and NLO calculations for $t\bar{t}$ total and differential cross sections have been implemented in various parton-level Monte Carlo computer programs, e.g., MCFM [35], MC@NLO [36], POWHEG [37], MadGraph/MadEvent [38,39].

The progress made in the past two decades on both the theoretical and computational sides by many groups has been remarkable (see for instance Refs. [40–50] and references therein). The fixed-order calculation for $t\bar{t}$ production at next-to-next-to-leading order (NNLO) at $\mathcal{O}(\alpha_s^4)$ in QCD [51–56] has been accomplished based on the STRIPPER subtraction [57–59] method. An independent calculation for the inclusive and differential $t\bar{t}$ production cross section has recently been accomplished based on the q_T-subtraction [60–62] method and has been incorporated in the publicly available computer program MATRIX [63]. The inclusion of top-quark decay effects has recently been studied in Refs. [64,65].

The impact of logarithmic enhancements in threshold resummation on the cross section has also been extensively studied in the past decades, and relevant work can be found in Refs. [66–83], while soft collinear effective field theory (SCET) [84–87] has been used in Refs. [88–96].

The impact of electroweak (EW) corrections on $t\bar{t}$ production from both the weak and QED sectors has been studied in Refs. [97–115].

Transverse momentum resummation for top-quark pair production has also been studied and recently documented in Refs. [116–122], while higher-order QCD corrections have been combined with parton shower (PS) simulations in [123–125], based on the general computer framework described in Refs. [126–128].

Top-quark pair production in global PDF analyses.
Top-quark pair production cross section measurements at the LHC and Tevatron are now a staple part of the dataset baseline in modern global analyses at NNLO and beyond in QCD to determine proton PDFs. Examples of these analyses are ABMP [129], CT18 [130], MSHT20 [131,132], and NNPDF [133,134]. In particular, measurements of $t\bar{t}$ differential cross sections at the ATLAS and CMS experiments have been included in the most recent analyses [130,131,133,134] at NNLO where, together with inclusive jet production measurements, they play an important role in constraining the gluon PDF in the intermediate to large-x region as they complement each other. Though the $t\bar{t}$ and inclusive high-p_T jet productions largely overlap in the Q-x plane, their matrix elements and phase-space suppression are different so that their constraints on the gluon are placed at different values of x. However, the presence of tensions between experiments which can potentially result in different pulls on the gluon at intermediate/large x, and the strong correlation between the top-quark mass m_t, the strong coupling α_s and the gluon itself, make PDFs extraction particularly challenging [129,135–137].

Another complication arises from the amount of information on the statistical and systematical uncertainties published by the experimental collaborations. These uncertainties can be expressed in terms of either the covariance matrix or nuisance parameter representation and conversion from covariance matrix to nuisance parameters is not unique. Complete information on the statistical, uncorrelated and correlated systematic uncertainties (and their sources) is critical to maximize constraints from the data in PDF determinations.

A large number of studies appeared in the literature that have used total and differential $t\bar{t}$ cross section measurements at the LHC to constrain the gluon and other PDFs. Recent and less recent analyses can be found in Refs. [135,136,138–145].

Main goals of this analysis. In this work, we shall study the impact of particular selections of 13 TeV $t\bar{t}$ single differential distributions at ATLAS (in the all-hadronic and lepton + jets channels) [23,24] and CMS (in the dilepton and lepton + jets channels) [26,28], on NNLO PDFs obtained by using the same framework (i.e., strategy, and tolerance criteria definitions for the uncertainties) adopted in the CT18 analysis [130]. In particular, PDFs are extracted by using optimal combinations of $t\bar{t}$ absolute differential cross sections measurements with different IL added on top of the CT18 baseline. The impact on the global fit from individual and combined kinematic distributions is analyzed by selecting different renormalization μ_R and factorization μ_F scale choices in the theory predictions. EW corrections are also considered, however their impact is found to be negligible.

In addition, we analyze the impact of 13 TeV $t\bar{t}$ double differential distributions at ATLAS and CMS using the ePump (error PDF updating method package) framework [146,147]. In Sec. II E we discuss the observed impact from double differential distributions on PDFs, and find that this is comparable to that from single differential ones. However, the treatment and interpretation of correlated systematic uncertainties in the analysis with double differential distributions is more challenging and complicates the data vs theory description. This work mainly concerns the study of the impact of 13 TeV $t\bar{t}$ single differential distributions on PDFs extracted in CTEQ global QCD analyses. A thorough and more extensive investigation of $t\bar{t}$ double differential distributions in PDF determinations will be presented in the future, in a separate work. A previous study investigating $t\bar{t}$ double differential cross sections at the LHC at 8 TeV and their impact on PDFs with ePump, is discussed in Ref. [140].

Single top production. Single (anti)top production cross sections in the t- and s-channel have also been measured at ATLAS and CMS at 7, 8, and 13 TeV collision energies. The impact of t-channel single (anti)top production has

been explored in Ref. [148] where it is found that an optimal combination of single-top data constrains the light quark and gluon PDFs with a reduction of their relative uncertainty by a fraction of a percent in the region $10^{-3} \leq x \leq 0.5$ with even more pronounced reduction on the ratio u/d around $x \approx 0.1$. Part of these measurements are also included in the NNPDF4.0 global analysis [134]. An investigation of the impact of single top-quark production cross section measurements at the LHC on CTEQ PDFs will also be presented in a separate work.

The single and double differential cross sections of top-quark pair production as well as single (anti)top production cross sections will be ingredients of high importance in the next generation of PDF determinations which are going to use higher IL data.

The rest of this paper is organized as follows: in Sec. II A we summarize the findings from recent global QCD analyses using 7 and 8 TeV top-quark pair production measurements, while in Sec. II B we describe the LHC measurements from ATLAS and CMS at 13 TeV considered in this study. In Sec. II C, we describe the theoretical framework where details of the calculations such as NNLO QCD corrections, EW corrections, and scale dependence are discussed. In Sec. II D, we discuss the impact on PDFs from the bin-by-bin statistical correlations in the ATLAS 13 TeV lepton + jets channel measurements, while in

Sec. II E we discuss single vs double distributions at the LHC. In Sec. III, we describe the impact on PDFs from individual 13 TeV $t\bar{t}$ datasets considered in separate fits, and in Sec. IV we present the main results of this analysis obtained from two optimal combinations of 13 TeV $t\bar{t}$ measurements. We will conclude in Sec. V, while the details of theoretical calculations are presented in Appendix A, and the treatment of the correlated systematics are summarized in Appendix D.

II. TOP-QUARK PAIR PRODUCTION AT THE LHC RUN I, AND II

We start with a brief overview of top-quark pair production measurements in recent global QCD analyses of PDFs which include LHC 7 and 8 TeV $t\bar{t}$ data from ATLAS and CMS, and summarize their findings. These measurements with respective analyses are also reported in Table I. Next, we discuss the LHC 13 TeV data that are at the core of this work and will play an important role in all post-CT18 PDF determinations.

A. Top-quark data in the CT18 era

The CT18NNLO analysis [130] includes two ATLAS absolute single-differential cross sections [19] $d\sigma/dp_T^t$ and $d\sigma/dm_{t\bar{t}}$ for the invariant mass with 20.3 fb^{-1} of IL, and

TABLE I. $t\bar{t}$ measurements at the LHC used in the ABMP16 [129], CT18 [130], MSHT20 [131], and NNPDF4.0 [134] global analyses.

Data	Lumi [fb^{-1}]	References	CT18	MSHT20	ABMP16	NNPDF4.0
ATLAS 8 TeV lep + jets (norm.)	20.3	[19]	$p_{T,t}, m_{t\bar{t}}$	$\cdots$	$\cdots$	$y_t, y_{t\bar{t}}$
ATLAS 8 TeV lep + jets (abs.)	20.3	[19]	$\cdots$	$m_{t\bar{t}}, y_{t\bar{t}}, p_{T,t}, y_t$	$\cdots$	$\cdots$
ATLAS 8 TeV dilep (norm.)	20.2	[149]	$\cdots$	$y_{t\bar{t}}$	$\cdots$	$y_{t\bar{t}}$
CMS 8 TeV lep + j and dilep (norm.)	19.7	[150]	$\cdots$	$p_{T,t} + y_t$	$\cdots$	$y_{t\bar{t}}$
CMS 8 TeV dilep 2D (norm.)	19.7	[22]	$(p_{T,t}, y_t)$	$(p_{T,t}, y_t)$	$\cdots$	$(y_t, m_{t\bar{t}})$
ATLAS 13 TeV lep + jets	36	[23]	This work	$\cdots$	$\cdots$	$\cdots$
ATLAS 13 TeV all hadronic	36.1	[24]	This work	$\cdots$	$\cdots$	$\cdots$
CMS 13 TeV dilep	35.9	[26]	This work	$\cdots$	$\cdots$	y_t
CMS 13 TeV lep + jets	35.8	[27]	$\cdots$	$\cdots$	$\cdots$	y_t
CMS 13 TeV lep + jets	137	[28]	This work	$\cdots$	$\cdots$	$\cdots$
Total cross sections						
CMS 5.02 TeV lep + jets	0.0274	[151]	$\cdots$	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$
CMS 5.02 TeV lep + jets	0.026	[152]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$
ATLAS 7 TeV	4.66/4.6/1.67/4.7	[153–157]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$
ATLAS 7/8 TeV	4.6/20.3	[158]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\sigma_{t\bar{t}}$
ATLAS 8 TeV lep + jets	20.3	[159]	$\cdots$	$\sigma_{t\bar{t}}$	$\sigma_{t\bar{t}}$	$\cdots$
CMS 7 TeV	2/2.2/3.9/3.54	[160–164]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$
CMS 7/8 TeV	19.6/19.7	[160,161]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$
CMS 8 TeV	19.6/18.4	[165,166]	$\cdots$	$\sigma_{t\bar{t}}$	$\sigma_{t\bar{t}}$	$\cdots$
CMS 8 TeV dilep	5.3	[21]	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$	$\cdots$
ATLAS 13 TeV dilep	3.2	[167]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$
ATLAS 13 TeV lep + jets	139	[168]	$\cdots$	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$
CMS 13 TeV lep + jets	19.7	[169]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\sigma_{t\bar{t}}$
CMS 13 TeV	0.042/2.2/2.53	[170–172]	$\cdots$	$\cdots$	$\sigma_{t\bar{t}}$	$\cdots$

the CMS normalized double-differential cross section [22] $d^2\sigma/dp_{T,t}dy_t$, with 19.7 fb^{-1}. The two ATLAS measurements are combined into one single dataset which includes the combination of the $e +$ jets and $\mu +$ jets channels for the $p_{T,t}$ and $m_{t\bar{t}}$ distributions with statistical correlations. These distributions are chosen according to their best compatibility within the global fit. In fact, the impact from the single-differential y_t and $y_{t\bar{t}}$ rapidity distributions (absolute or normalized) at ATLAS [19] is also explored and tension is found with some other datasets. For example, the y_t and $y_{t\bar{t}}$ rapidity distributions show agreement with HERA DIS data, but have opposite trend as compared to the CMS $d^2\sigma/dp_{T,t}dy_t$ and ATLAS $p_{T,t}$ and $m_{t\bar{t}}$ combined distributions. Their inclusion, either in the single-differential or double-differential form does not reduce PDF errors. The resulting $t\bar{t}$ impact on the CT18 PDFs is found to be modest with a preference for a softer gluon at large x in the $0.1 \lesssim x \lesssim 0.3$ range, and with changes that have no statistically significant amount.

The MSHT20 analysis [131], in addition to the same CMS double-differential distribution considered in the CT18 study, includes ATLAS single-differential cross section measurements for the $m_{t\bar{t}}, y_{t\bar{t}}, p_{T,t}$, and y_t distributions [19] in the lepton + jets channel combined with statistical correlations, ATLAS measurements of $y_{t\bar{t}}$ distribution [149] in the dilepton channel, as well as the CMS normalized $y_{t\bar{t}}$ distribution in the lepton + jets channel [150]. In addition, four total cross section measurements from ATLAS [159] and CMS [21,165,166] are included. For the ATLAS single-differential measurements in the lepton + jets channel, the correlated systematic parton-shower (PS) error across the four distributions has been decorrelated according to the procedure described in Ref. [141]. The impact on the MSHT20 gluon PDF from top-quark pair production data results in a suppressed high-x gluon ($x \gtrsim 0.1$) with a complicated interplay/tension with the Z-p_T and LHC jet data. PDF uncertainties at large x are slightly reduced. Most of the impact is from the ATLAS combination of single differential distributions [19,149] in the lepton + jets channel, and the treatment of systematic uncertainties plays a significant role in the description of data in terms of χ^2.

The NNPDF4.0 analysis [134] includes the ATLAS normalized y_t and $y_{t\bar{t}}$ distributions [19] in the lepton + jets channel, as well as the ATLAS normalized $y_{t\bar{t}}$ distribution in the dilepton channel [149]. From CMS, it includes the normalized $y_{t\bar{t}}$ distribution in the lepton + jets channel [150] as well as the normalized $1/\sigma d^2\sigma/dy_t dm_{t\bar{t}}$ distribution in the dilepton channel [22]. Two measurements of $t\bar{t}$ total cross section from ATLAS [158] and CMS [173] are also considered. In addition to these measurements at $\sqrt{S} = 8$ TeV, the NNPDF4.0 analysis includes two CMS rapidity y_t distributions at $\sqrt{S} = 13$ TeV in the lepton + jets [27] and dilepton channel [26], respectively, and other total cross section measurements

at different center-of-mass energy: $\sigma_{t\bar{t}}$ at CMS [151] with $\sqrt{S} = 5.02$ TeV, at ATLAS [158] and CMS [173] with $\sqrt{S} = 7$ TeV, and at ATLAS [168] and CMS [169] with $\sqrt{S} = 13$ TeV. t-channel single top total and differential production cross section measurements are also considered in the NNPDF4.0 global fit. The resulting gluon is more suppressed at $x \gtrsim 0.1$ as compared to CT18 and MSHT20. total and differential production cross section measurements are also considered in the NNPDF4.0 global fit. The resulting gluon is more suppressed at $x \gtrsim 0.1$ as compared to CT18 and MSHT20. However, top-quark pair production data have overall a moderate impact on the NNPDF4.0 global fit. In a fit with no top-quark data, the gluon is slightly enhanced at $x \gtrsim 0.1$, but well within the NNPDF4.0 uncertainty.

The ABMP16 analysis [129] considers selected measurements of top-quark pair production total cross section only at the LHC and Tevatron with different collision energies: one data point at CMS with $\sqrt{S} = 5.02$ [152], six data points at ATLAS [153–158] and five at CMS [160–164] at $\sqrt{S} = 7$ TeV, two data points at ATLAS [158,159] and four at CMS [160,161,165,166] at $\sqrt{S} = 8$ TeV, one data point at ATLAS [167], and four at CMS [169–172] at $\sqrt{S} = 13$ TeV. The ABMP16 PDF parameters are fitted simultaneously with α_s and m_t. These total cross section measurements lead to an increase in the gluon central value at large x in the $0.05 \leq x \leq 0.35$ range, with 10–20% increase at $x \gtrsim 0.1$ in both the $n_f = 4$ and $n_f = 5$ calculations, depending on the factorization scale choice. These variations are well within the ABMP PDF uncertainty. However, the impact on the gluon PDF uncertainty is found to be small.

B. The 13 TeV top-quark data in the post-CT18 era

In this section we describe the top-quark pair production differential cross section measurements at ATLAS and CMS that are considered in this study.

In 2018, the CMS collaboration published differential cross section measurements at 13 TeV in the dilepton [26] and lepton + jets [27] channels. The lepton + jets channel measurements of Ref. [27] have recently been superseded by new measurements with a higher IL of 137 fb^{-1} [28] which we use in this analysis. In parallel, the ATLAS collaboration published two 13 TeV measurements based on the lepton + jets [23] and the all-hadronic [24] channels respectively. These measurements are described below.

ATLAS lepton + jets *channel (ATL13lj).* We explored the impact of absolute top-quark pair production single differential distributions based on lepton + jet events measured at ATLAS 13 TeV with 36 fb^{-1} of IL [23], which we label "ATL13lj." We use full phase-space results at parton-level, and consider reconstructed measurements at parton-level in the resolved topology which are expected to provide direct constraints on PDFs. We do not consider the transverse

momentum $p_{T,t\bar{t}}$ distribution of the $t\bar{t}$ pair because final-state interactions (FSI) between hard partons and beam remnants from the initial state may lead to substantial corrections to pair invariant mass (PIM) kinematic distributions when recoiling radiation is suppressed (see for instance the discussion in Ref. [174]). These contributions may manifest as higher-order perturbative corrections to the factorized cross section, and as nonperturbative corrections that are suppressed by powers of perturbative scales.

The correlated systematic uncertainties associated to these measurements [23] are presented in terms of both the covariance matrix and nuisance parameter representations. We adopt the nuisance parameter representation because this is the default treatment for correlated systematic uncertainties in all CTEQ PDF analyses.

We study the impact of bin-by-bin statistical correlations between these distributions released for various combinations. The results are shown in Sec. II E where we find that the statistical correlations for these measurements have some impact on the χ^2 description, but their overall impact on the PDFs and their errors is negligible.

In addition, we study the impact of combinations of double-differential distributions, such as $d\sigma/dm_{t\bar{t}}dy_{t\bar{t}}$, and confront the results with those obtained by including the corresponding statistically combined single differential $d\sigma/dm_{t\bar{t}}$, $d\sigma/dy_{t\bar{t}}$ distributions.

To facilitate comparisons with measurements at CMS, the ATL13lj measurements are released by using two different bin resolutions: (1) the original ATLAS bin resolution, and (2) the CMS bin resolution which shares the size of its bins and the number of points with the CMS 13 TeV $t\bar{t}$ measurements in the dilepton channel [26]. These two bin resolutions differ in number of bins and bin size. Moreover, bin-by-bin statistical correlations are made available only for bin choice (1). As we shall see in the next sections, this has a non-negligible impact on the data-vs-theory description for single distributions using these two bin resolutions.

ATLAS all-hadronic channel (ATL13had). For the ATLAS 13 TeV all-hadronic channel with an integrated luminosity of 36.1 fb^{-1} [24], we consider absolute single differential cross sections for the reconstructed top quark in terms of $p_{T,t_1(t_2)}$, $|y_{t\bar{t}}|$, $m_{t\bar{t}}$, and $H_T^{t\bar{t}} = p_{T,t} + p_{T,\bar{t}}$, where $p_{T,t_1(t_2)}$ is the transverse momentum of the leading (trailing) top quark. We label this dataset as "ATL13had." Combinations of double-differential cross sections for this channel are not available. For these measurements, correlated systematic uncertainties are made available in terms of nuisance parameters which are included in our global analysis.

CMS dilepton channel (CMS13ll). We study the impact of differential cross sections of top-quark pair production in the dilepton channel measured at CMS 13 TeV with 35.9 fb^{-1} of IL [26]. These are labeled as "CMS13ll." They are published in terms of absolute and normalized

single-differential distributions for the reconstructed top quarks. We consider only absolute cross section measurements in the full phase space at the top-quark level. This allows us to simplify the calculation of the theoretical predictions. Measurements relative to the decayed particle in the fiducial phase space will be analyzed in a future work as they require an additional effort to obtain complete predictions at the NNLO accuracy. In this study, we consider single-differential cross sections in terms of the $p_{T,t}(p_{T,\bar{t}})$, $y_t(y_{\bar{t}})$, $y_{t\bar{t}}$, and $m_{t\bar{t}}$ distributions.

Correlated systematic uncertainties are presented in terms of the covariance matrix. In accordance with the default treatment of systematic correlations in the CTEQ framework, we convert the covariance matrix into the nuisance parameter representation by using a version of the iterative $\Sigma + K$ decomposition method, adopted in the CT18 analysis [130]. A slightly extended discussion is in Appendix D.

Bin-by-bin statistical correlations between measurements are not available to date, to the best of our knowledge. Therefore, we cannot exploit combinations of single differential distributions for these measurements in the global fit.

CMS lepton + jets channel (CMS13lj). In 2018, the CMS collaboration published 13 TeV measurements with 35.8 fb^{-1} of IL in the lepton + jets channel [27]. These measurements are now superseded by new measurements with higher precision with 137 fb^{-1} of IL [28], which we use in this analysis and are labeled as "CMS13lj." We examine the impact from the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ single differential distributions in the full phase space. Bin-by-bin statistical correlations are not provided also for these measurements. Statistical and correlated systematic uncertainties are given in terms of the covariance matrix which we convert to nuisance parameter representation as discussed above. We have also examined the p_T^t and y_t distributions. These are for hadronically reconstructed tops. They are relative to a top quark decaying into a b quark and a W boson with a subsequent hadronic decay of the W boson. The theory predictions at NNLO in QCD for these distributions through MATRIX are therefore challenging to be obtained as it produces distributions for a stable top. However, in our preliminary investigation we found that these distributions were either poorly described by the MATRIX theory (e.g., p_T^t), or their impact was negligible. Therefore, we do not consider these distributions.

C. Theoretical framework

Details of the theoretical framework used in this analysis are given below. Additional details and comparisons are given in Appendix A.

Global PDF analyses necessitate fast, precise, and accurate theory predictions that are compared to experimental data in the χ^2-minimization procedure. To reduce the CPU turn-around time, fast theory predictions are obtained as
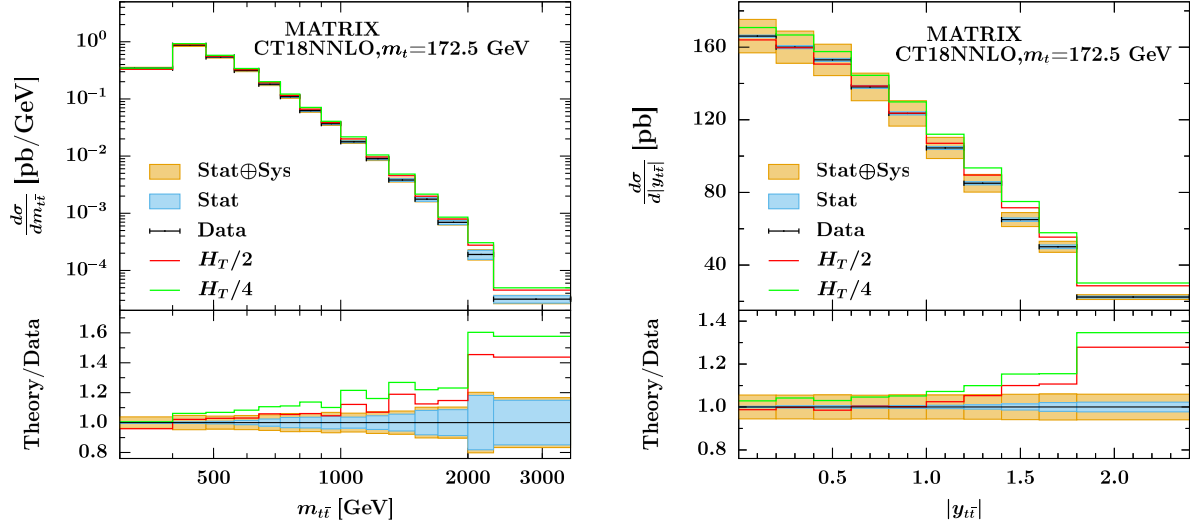
FIG. 1. Central scale choices for the $m_{t\bar{t}}$ (left) and $|y_{t\bar{t}}|$ (right) theory predictions compared to the CMS13lj measurements. The two central scales are represented by solid lines of different colors: $\mu_R = \mu_F = H_T/2$ is red, while $H_T/4$ is green.

interpolating tables through the `fastNLO` [175–178] and `APPLGrid` [179] frameworks.

NNLO QCD and NLO EW corrections. The theory predictions at NNLO in QCD used in this work are based on two independent calculations. One is the numerical calculation described in Refs. [55,56], based on the `STRIPPER` subtraction method [57–59] and implemented in `fastNLO` tables [180,181]; the other one is described in Refs. [60–62], based on the q_T-subtraction method [182] and implemented in the computer program `MATRIX` [63,183], that is publicly available.

The theory predictions utilized for the CMS13ll [26] differential distributions which share the same bin resolution of some of the ATL13lj [23] distributions, are generated at NNLO with `fastNLO` [181]. The theory predictions for the ATL13lj distributions resolved in terms of the ATLAS bins, are instead obtained with `MATRIX`. To optimize the calculation and minimize the CPU's turn-around time in the global fit, the `MATRIX` NNLO theory is constructed through bin-by-bin NNLO/NLO K-factors defined as $K = (\hat{\sigma}^{(\mathrm{NNLO})} \otimes \mathcal{L}^{(\mathrm{NNLO})})/(\hat{\sigma}^{(\mathrm{NLO})} \otimes \mathcal{L}^{(\mathrm{NNLO})})$, where $\mathcal{L}^{(\mathrm{NNLO})}$ represents the corresponding NNLO PDF luminosity. The NLO theory calculation is obtained from fast lookup `APPLGrid` tables [179] generated with `MCFM` [35,184]. A detailed comparison between the NNLO theory obtained with `fastNLO` tables [180] and that obtained with `MATRIX`, as well as a consistency check, are in Appendix A 1.

The theory predictions relative to the ATL13had [24] and CMS13lj [27,28] measurements in the global fit are obtained with `MATRIX` in similar manner.

The EW corrections and their implementation are discussed in Appendix A 2. Overall, the impact of the EW corrections on PDF determination is found to be negligible given the current size of the experimental errors.

Impact of different central scales in the theory prediction. It is interesting to explore the impact of different choices for the central scale in the theory prediction for the $t\bar{t}$ differential cross sections at 13 TeV and make a comparison with the experimental uncertainties. To this purpose, we consider the CMS13lj measurements [28] with 137 fb^{-1} of IL as they are the most precise data in this analysis. In Fig. 1, we confront to the experimental data with theory predictions for the $m_{t\bar{t}}$ (left) and $|y_{t\bar{t}}|$ (right) distributions calculated with the CT18NNLO PDFs and with different scale choices. The theory predictions are computed with central scales $\mu_F = \mu_R = H_T/2$ and $H_T/4$ respectively, represented by solid lines of different color in Fig. 1. For both distributions, the agreement with data deteriorates at large $m_{t\bar{t}}$ ($m_{t\bar{t}} \gtrsim 1.5$ TeV) and large $|y_{t\bar{t}}|$ ($|y_{t\bar{t}}| \gtrsim 1.5$), although it is better in correspondence of $H_T/2$. These two CMS13lj measurements are very sensitive to the gluon and play a major role in the global fit which is discussed later.

In this work, independent fits with central-scale choices $H_T/4$, $H_T/2$, and H_T are examined. Differences in the predictions obtained with these scale choices can be used to quantify part of the theoretical uncertainty in the $t\bar{t}$ theory predictions in the global analysis.

Hessian profiling with the `ePump` *framework.* A preliminary assessment of the impact of new measurements on existing PDFs sets before performing the global fit can be done by using different tools relying on a variety of statistical procedures based on the Monte Carlo or the Hessian method. Examples of these are: `ePump` (error PDF updating method package) [146,147], Hessian profiling [185], Bayesian reweighting techniques [186–188], and the `PDFSense` method [189].

We perform a preliminary investigation to assess the impact of the 13 TeV $t\bar{t}$ on PDFs using the `ePump` package.
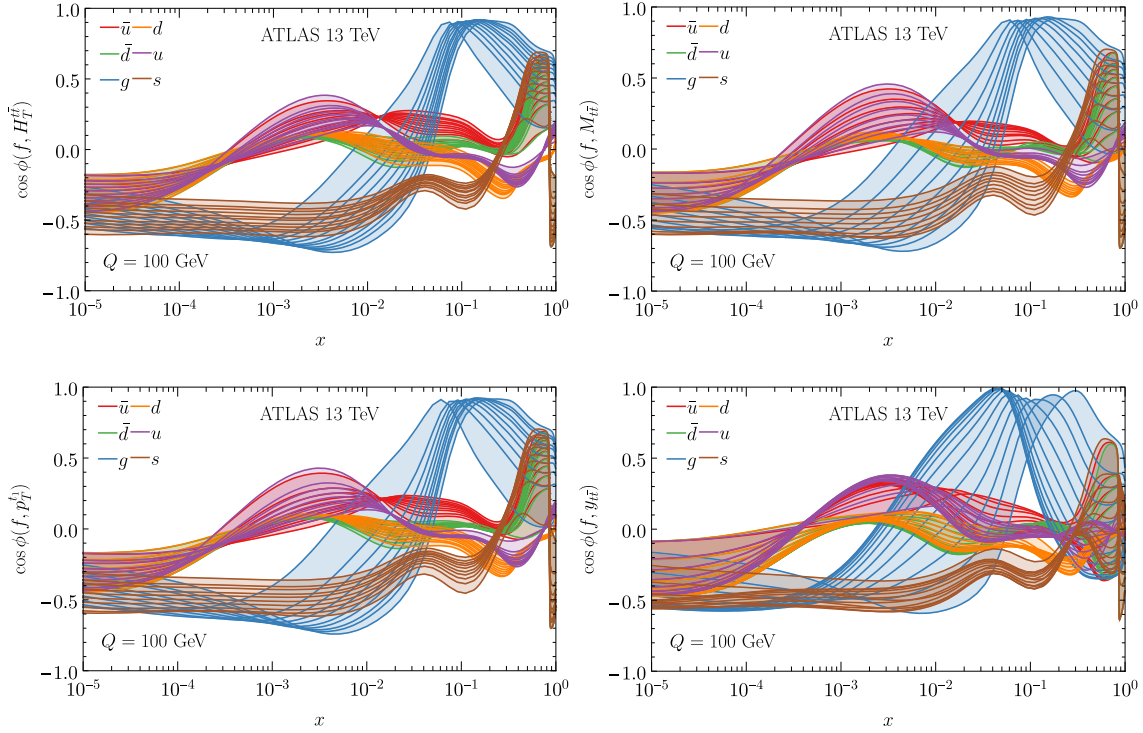
FIG. 2. PDF correlation-cosine plots [193] as a function of the momentum fraction x for the ATL13had [24] measurements added on top of the CT18NNLO baseline at $Q = 100$ GeV.

ePump allows one to obtain the updated best-fit PDF and relative PDF errors using the Hessian-profiling method. Previous studies using ePump are presented in Refs. [139,140,147,190–192]. ePump calculates the minimum of the χ^2 function within the Hessian method using some approximations (e.g., quadratic dependence on the parameters for the original data) which may restrict its applicability in certain cases, especially in presence of tensions among data (see related discussion in Sec. II E of Ref. [146]). ePump is used in this work as a fast and powerful tool to assess the impact of new data on PDFs, but it cannot replace a full PDF global analysis of experimental data and deviations from the results obtained in a global fit are expected, in particular, when subsets of data in the baseline are incompatible (see Appendix C for more details).

In Figs. 2 and 3, we present plots for the correlation-cosine between experimental data and PDFs for various differential distributions as a function of the parton momentum fraction x for the ATL13had [24] and CMS13ll [26] measurements, added on top of the CT18 baseline. PDFs in the various insets are represented by different colors and each solid line corresponds to different bins of the distribution under scrutiny. In both experiments, the gluon PDF appears to be strongly correlated at large x in the interval $0.05 \lesssim x \lesssim 0.4$ for all distributions. The correlation plots relative to the other experimental data considered in this work are very similar.

The results from ePump for the preliminary assessment of impact of individual experiments are reported in Table III, where they are shown for different values of the central scale in the theory predictions, and can be compared to those from the global fit which will be discussed later in Sec. III. We study the quality-of-fit in terms of $\chi^2/N_{\rm pt}$ and use it as the criterion for an initial investigation of the data.

Looking at the $\chi^2/N_{\rm pt}$ values from ePump in Table III, we note that the $y_{t\bar{t}}$ and $p_{T,t_{1,2}}$ distributions produce an acceptable fit quality (with $\mu_F = \mu_R = H_T/2$ or $H_T/4$) among the ATL13had measurements, while for the $H_T^{t\bar{t}}$ and $m_{t\bar{t}}$ distributions $\chi^2/N_{\rm pt} > 1.5$ regardless of the scale choice.

The $y_{t\bar{t}}$ and y_t distributions are the only two that can be well described (for $\mu_F = \mu_R = H_T/2$ or $H_T/4$) among the CMS13ll measurements, and the $m_{t\bar{t}}$ distribution is the only acceptable candidate in the CMS13lj measurements as $y_{t\bar{t}}$ cannot produce a good fit, regardless of the scale choice.

The ATL13lj case deserves more efforts because of the two bin resolutions and the bin-by-bin statistical uncertainties. In the ATL13lj measurements with the CMS bin resolution, we observe that the $y_{t\bar{t}}$, $m_{t\bar{t}}$ and y_t distributions produce acceptable $\chi^2/N_{\rm pt}$ values in Table III, while the $p_{T,t}$ distribution produces $\chi^2/N_{\rm pt} > 2$, independently of the scale choice.

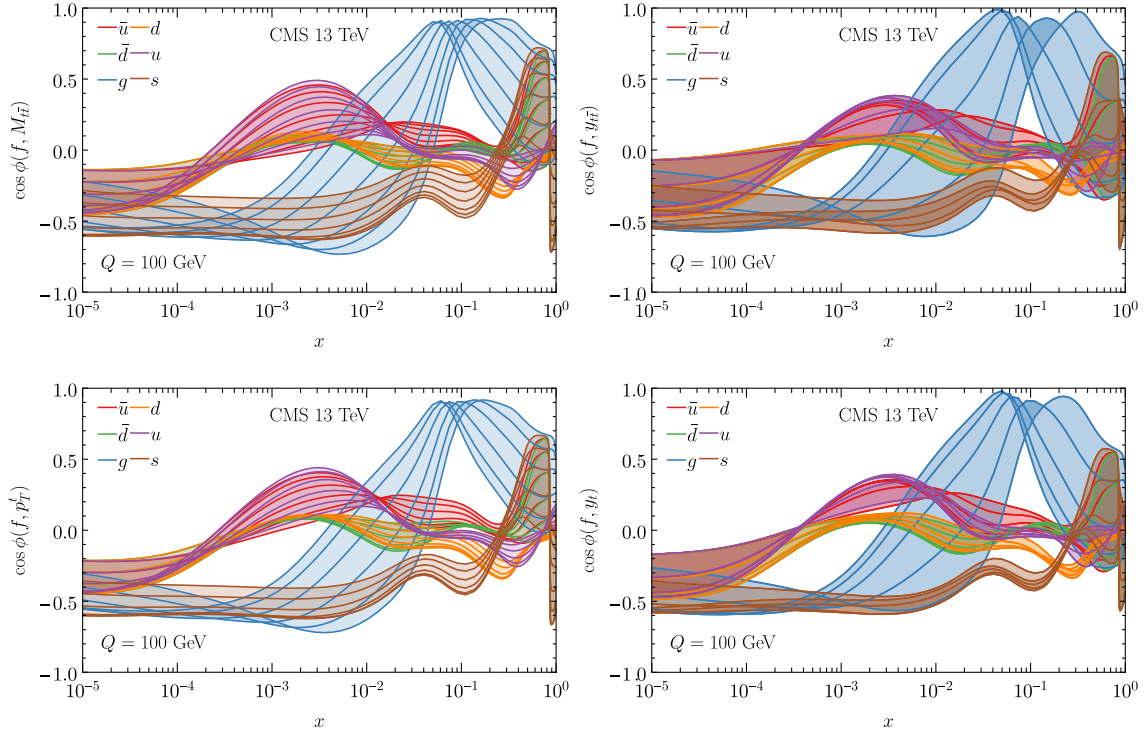We discuss the ATL13lj measurements with the ATLAS binning resolution separately in Sec. II D, where we

FIG. 3. Same as in Fig. 2, but for the CMS13ll measurements [26].

D. Statistical correlations in the ATLAS lepton + jets data

The impact of bin-by-bin statistical correlations in the ATL13lj measurements [23] available on the HEPData [194] repository is first analyzed with ePump. The theoretical predictions are obtained with the MATRIX computer program as described in Sec. II C, with top-quark mass set to $m_t = 172.5$ GeV in the pole mass approximation. The impact of central-scale dependence is analyzed for the three different scales $\mu_R = \mu_F = \{H_T, H_T/2, H_T/4\}$.

In Table II we report the χ^2/N_{pt} values from ePump obtained with and without inclusion of statistical correlations. The implementation using bin-by-bin statistical correlations is labeled as "WSC" while the one without statistical correlations is "NSC." In addition, we use the notation adopted in the ATLAS publication where $y_{t\bar{t}}^B$ denotes the boosted topology for the rapidity of the $t\bar{t}$ system, while $H_T^{t\bar{t}}$ denotes the scalar sum of the transverse momenta of the hadronic and leptonic top quarks, $H_T^{t\bar{t}} = p_{T,t}^{had} + p_{T,t}^{lep}$.

To consistently account for the statistical correlations, the WSC implementation adopts the ePump "error type 2" option. This option corresponds to a prescription of χ^2 calculation in ePump, in which nuisance parameters, statistical and systematical correlation matrices, can be read

from file as inputs. In this case, ePump treats the correlated systematic uncertainties in terms of their absolute (rather than percent) contribution to each data point, and therefore as additive errors. The NSC implementation adopts the ePump "error type 4" option, where the nuisance parameters representation is used and correlated systematic uncertainties are normalized to data.[1]

The "error type 2" and "error type 4" choices in ePump are used to facilitate the treatment and inclusion of statistical correlations in both ePump and the global fitting code (as we shall see later) which uses the nuisance parameter's χ^2 definition. Since we studied the impact of two different scale choices (i.e., two theories), "error type 2" and "error type 4" are easy to handle as there is no need to normalize to two theories. The results from the global fits presented in the next sections are obtained using the multiplicative definition for the correlated systematic uncertainties, normalized to theory predictions, in the χ^2 definition (which

---

[1]This choice can in principle introduce the D'Agostini bias [195,196], where data fluctuations can result in a bias in favor of experimental data points with lower central values. We do not discuss this here because the overall impact of the bin-by-bin statistical correlations of the ATL13lj measurements is negligible, and the same analysis is performed in the global fit where the systematic error for each multiplicative term is determined by multiplying the fractional uncertainty times the theoretical prediction for that bin. (See Appendix C2 of the CT18 study [130].)

TABLE II. χ^2/N_{pt} values for the ATLAS 13 TeV $t\bar{t}$ production in the lepton + jets channel [23] before/after the ePump updating of the CT18 PDFs, with (WSC) and without (NSC) statistical correlation, and with central scales set to $H_T/4$, $H_T/2$, and H_T, respectively.

Scale			$H_T/4$		$H_T/2$		H_T	
Observable	N_{pt}		WSC	NSC	WSC	NSC	WSC	NSC
$m_{t\bar{t}}$	9		1.150/1.146	1.105/1.101	1.309/1.300	1.260/1.251	1.249/1.231	1.234/1.214
$y_{t\bar{t}}$	7		1.027/1.022	1.002/0.997	0.713/0.706	0.718/0.710	0.671/0.622	0.687/0.677
$y_{t\bar{t}}^{B}$	9		1.153/1.147	1.087/1.081	0.861/0.851	0.855/0.845	0.853/0.838	0.862/0.847
$H_T^{t\bar{t}}$	9		0.923/0.920	0.864/0.862	0.832/0.831	0.781/0.780	1.185/1.183	1.106/1.104
$m_{t\bar{t}} + y_{t\bar{t}}$	16		1.312/1.304	1.204/1.193	1.198/1.193	1.139/1.132	1.175/1.166	1.151/1.141
$m_{t\bar{t}} + y_{t\bar{t}}^{B}$	18		1.331/1.318	1.246/1.230	1.205/1.198	1.143/1.135	1.186/1.178	1.144/1.135
$m_{t\bar{t}} + H_T^{t\bar{t}}$	18		1.610/1.594	1.462/1.449	1.696/1.640	1.206/1.187	2.364/2.257	1.511/1.472
$y_{t\bar{t}} + y_{t\bar{t}}^{B}$	16		0.994/0.985	0.906/0.900	0.703/0.692	0.633/0.625	0.663/0.651	0.581/0.571
$y_{t\bar{t}} + H_T^{t\bar{t}}$	16		0.950/0.942	0.892/0.886	0.648/0.645	0.625/0.622	0.761/0.758	0.725/0.721
$y_{t\bar{t}}^{B} + H_T^{t\bar{t}}$	18		1.531/1.521	1.255/1.247	1.152/1.147	0.967/0.962	1.130/1.127	0.961/0.956
$m_{t\bar{t}} + y_{t\bar{t}} + y_{t\bar{t}}^{B}$	25		1.098/1.085	1.002/0.991	0.925/0.919	0.864/0.858	0.906/0.900	0.855/0.848
$m_{t\bar{t}} + y_{t\bar{t}} + H_T^{t\bar{t}}$	25		1.657/1.648	1.431/1.425	1.680/1.658	1.087/1.082	2.402/2.356	1.377/1.362
$m_{t\bar{t}} + y_{t\bar{t}}^{B} + H_T^{t\bar{t}}$	27		1.694/1.687	1.424/1.417	1.780/1.766	1.162/1.157	2.480/2.447	1.475/1.464
$y_{t\bar{t}} + y_{t\bar{t}}^{B} + H_T^{t\bar{t}}$	25		1.257/1.248	1.001/0.994	0.949/0.944	0.729/0.724	0.966/0.962	0.716/0.712
$m_{t\bar{t}} + y_{t\bar{t}} + y_{t\bar{t}}^{B} + H_T^{t\bar{t}}$	34		1.571/1.564	1.233/1.227	1.672/1.654	0.954/0.949	2.466/2.429	1.220/1.210

corresponds to "error type 1" in ePump).[2] For more details about the ePump implementation, we refer the reader to the online manual [197].

On the one hand, from the values in Table II we note that regardless of the inclusion of statistical correlations and scale choice, there is some tension between various distributions, e.g., $m_{t\bar{t}}$ and $H_T^{t\bar{t}}$, $m_{t\bar{t}}$ and $y_{t\bar{t}}$, and $m_{t\bar{t}}$ and $y_{t\bar{t}}^{B}$. Taken individually, all of these single differential distributions produce acceptable χ^2/N_{pt}. In the ePump environment, tensions are amplified when statistical correlations are included. In addition, we note that $y_B + H_T^{t\bar{t}}$ produce a net effect that is similar to that of $y_{t\bar{t}} + m_{t\bar{t}} + y_B + H_T^{t\bar{t}}$ when they are statistically combined with the scale choice $H_T/4$.

On the other hand, as illustrated in Fig. 4, the impact of the bin-by-bin statistical correlations on PDFs is negligible. In addition, even if the $m_{t\bar{t}}$, $y_{t\bar{t}}$, $y_{t\bar{t}}^{B}$, and $H_T^{t\bar{t}}$ distributions are simultaneously included in ePump, the impact of these distributions on the gluon is very small. As we shall see in Sec. III, a similar behavior is observed in the true global QCD analysis where the overall impact from statistical correlations in the ATL13lj data is very small.

A general discussion on the importance of making statistical models and the associated data publicly available can be found in Ref. [198].

E. Single vs double distributions at the LHC 13 TeV

The ATLAS and CMS collaborations published several measurements of top-quark pair production double differential distributions at 13 TeV at both particle and parton level. A question then arises about the optimal strategy to be used to best exploit these new high-precision measurements. That is, whether to use statistically-combined single differential distributions, or double differential distributions directly. While a more exhaustive analysis specifically related to this point is beyond the scope of this paper and will be addressed elsewhere, here we explore the impact on CT18 PDFs in both cases by using ePump. As a case study, we consider the ATL13lj [23] measurements. In particular, we investigate the impact on the CT18 gluon PDF from adding two single-differential (1d) distributions (e.g., $d\sigma/dm_{t\bar{t}}$ and $d\sigma/dy_{t\bar{t}}$) on top of the CT18 baseline as well as the impact from the double differential (2d) distribution $d^2\sigma/dm_{t\bar{t}}dy_{t\bar{t}}$.

In Fig. 5 we illustrate the ePump results for the CT18 gluon in four different cases: (1) when the $m_{t\bar{t}} + y_{t\bar{t}}$ 1d distributions are added together on top of the CT18 baseline (orange dashed curve), (2) the $y_{t\bar{t}}$ distribution only is added (green curve), (3) the $m_{t\bar{t}}$ distribution only is added (red curve), and (4) the $d^2\sigma/dm_{t\bar{t}}dy_{t\bar{t}}$ 2d distribution is added (purple curve). The blue band represents the CT18NNLO PDF uncertainty at 90% CL. In this example, when the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ 1d distributions are individually added, or added together on top of the CT18 baseline, they generate opposite pulls on the large-x gluon as compared to the 2d distribution. At $Q = 1.3$ GeV, the 2d distribution has a preference for a softer gluon in the $x \gtrsim 0.4$ region, while at $Q = 100$ GeV this preference is in the $x \gtrsim 0.3$ region. In both cases, the 1d distributions have opposite trends at large x with a much milder effect.

Because of these opposite trends at large x, and because the gluon PDF uncertainty is large in this region, where

---

[2]The multiplicative approach for the correlated systematic uncertainties in this choice suppresses the D'Agostini bias.
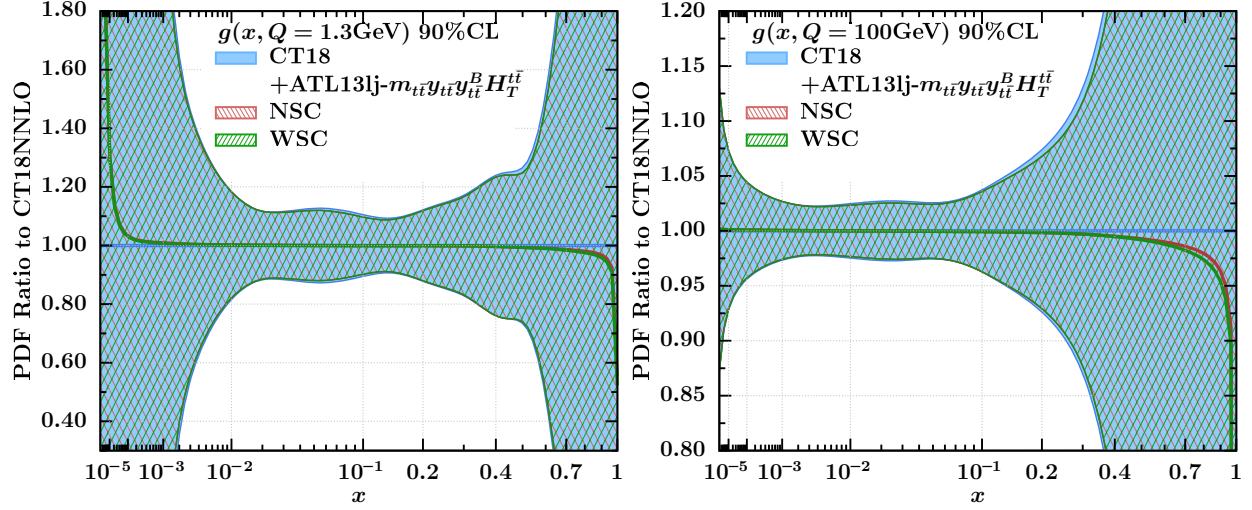
FIG. 4.   Impact of bin-by-bin statistical correlations in the ATLAS lepton + jet channel measurements [23] on the CT18NNLO PDFs with ePump. PDF uncertainties are obtained at the 90% CL.
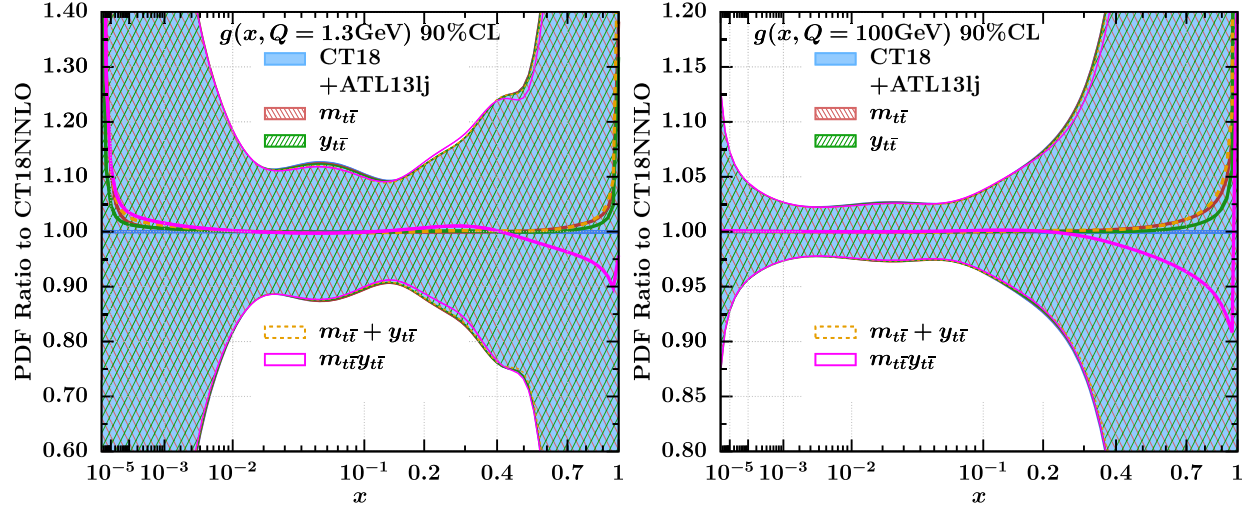

FIG. 5.   Impact on the CT18 gluon PDF from 1d and 2d distributions at ATLAS [23] at 13 TeV with ePump. Here $m_{t\bar{t}} + y_{t\bar{t}}$ indicates two 1d distributions added on top of the CT18 baseline, while $m_{t\bar{t}}y_{t\bar{t}}$ refers to the inclusion of a 2d distribution. The blue band represents the CT18NNLO PDF uncertainty at 90% CL.

there is essentially no data, we conclude that there is no obvious preference between the 1d $m_{t\bar{t}} + y_{t\bar{t}}$ combination, and the 2d $d^2\sigma/dm_{t\bar{t}}dy_{t\bar{t}}$ distribution. Ultimately, the impact from 2d and other multiple differential distributions must be assessed in the more general environment of a global PDF fit and explored against that of 1d distributions and their combinations. This will be analyzed in a future work.

III. GLOBAL QCD ANALYSIS: IMPACT FROM INDIVIDUAL SINGLE DIFFERENTIAL DISTRIBUTIONS AT THE LHC 13 TeV

In this section, we extend the analysis of the 13 TeV $t\bar{t}$ data to the global fit and extract post-CT18 PDFs at NNLO.

In particular, we illustrate the results of multiple global QCD analyses at NNLO using the 1d absolute differential cross sections from the ATL13had, ATL13lj, CMS13ll, and CMS13lj measurements individually added on top of the CT18 baseline. We analyze the impact of these high-precision measurements on the gluon PDF and other relevant PDF combinations, and compare the results to the CT18 fit.

The setup of the global fits is the default one used in the CT18 study, where we use the same tolerance criterion to estimate the post-CT18 PDF uncertainties, and the same PDF parametrizations. The inclusion of these new 13 TeV $t\bar{t}$ data that place constraints on the gluon at large x would require a further analysis of PDF parametrizations and in

TABLE III. Results of the NNLO global QCD analysis for all the 13 TeV measurements from ATLAS and CMS included one at a time, and in a combined manner with statistical correlations when these are available. Included are also the $\chi^2/N_{\rm pt}$ values from ePump. The analysis is performed by adopting different central-scale choices for the $t\bar{t}$ 13 TeV measurements.

			ePump			Global fit	
Exp	Observable	$N_{\rm pt}$	H_T	$H_T/2$	$H_T/4$	$H_T/2$	$H_T/4$
ATL13had	$m_{t\bar{t}}$	9	1.75	1.57	1.60	1.53	1.47
	$H_T^{t\bar{t}}$	11	1.98	1.77	1.59	1.50	1.74
	$y_{t\bar{t}}$	12	1.28	1.15	0.94	1.05	1.07
	p_{T,t_1}	10	1.30	1.19	1.12	1.20	1.33
	p_{T,t_2}	8	1.13	0.84	1.05	0.84	1.59
CMS13ll	$m_{t\bar{t}}$	7	3.46	3.07	3.14	3.12	3.23
	$y_{t\bar{t}}$	10	1.66	0.97	0.68	0.94	0.67
	$p_{T,t}$	6	3.60	3.70	3.68	3.56	3.05
	y_t	10	1.33	0.94	0.87	1.00	0.69
CMS13lj	$m_{t\bar{t}}$	15	1.49	1.38	1.81	1.20	1.67
	$y_{t\bar{t}}$	10	6.47	6.24	6.42	6.01	5.88
ATL13lj			CMS bins				
	$m_{t\bar{t}}$	7	2.40	1.17	0.68	0.83	0.66
	$y_{t\bar{t}}$	10	0.91	0.69	0.62	0.74	0.75
	$p_{T,t}$	6	2.34	2.01	2.47	1.35	1.43
	y_t	10	1.30	1.07	1.10	1.16	0.68
			ATLAS bins without statistical correlation (NSC)				
	$m_{t\bar{t}}$	9	1.55	1.12	0.94	1.27	0.92
	$y_{t\bar{t}}$	7	0.91	0.74	0.80	0.76	0.90
	$y_{t\bar{t}}^B$	9	1.40	1.27	1.53	0.85	0.93
	$H_T^{t\bar{t}}$	9	1.35	0.91	0.93	0.81	0.80
	$m_{t\bar{t}} + y_{t\bar{t}} + y_{t\bar{t}}^B + H_T^{t\bar{t}}$	34	1.87	1.28	1.46	0.93	1.06
			ATLAS bins with statistical correlations (WSC)				
	$m_{t\bar{t}}$	9	1.68	1.35	0.98	1.29	0.96
	$y_{t\bar{t}}$	7	0.88	0.75	0.92	0.75	0.92
	$y_{t\bar{t}}^B$	9	1.06	0.87	1.01	0.86	0.99
	$H_T^{t\bar{t}}$	9	1.40	0.85	0.85	0.86	0.86
	$m_{t\bar{t}} + y_{t\bar{t}} + y_{t\bar{t}}^B + H_T^{t\bar{t}}$	34	3.10	1.61	1.32	1.59	1.32

particular the study of a more flexible parametrization for the gluon PDF. This requires substantial effort and an extended analysis using more flexible PDF parametrizations will be presented in a forthcoming study.

The value of the top-quark mass is set to $m_t^{(\rm pole)} = 172.5$ GeV, and the strong coupling constant at the Z-boson mass is set to $\alpha_s(M_Z) = 0.118$.

To find optimal combinations of measurements that allow us to include as much information as possible and, at the same time, minimize tension between observables, we first perform analyses including the 13 TeV $t\bar{t}$ datasets from every single channel one at a time, in individual global fits. Distributions are included in a combined manner only for the ATL13lj measurements where bin-by-bin statistical

correlations are studied in advance in ePump preliminary assessments.

In addition, we consider different choices for the central scale in the 13 TeV $t\bar{t}$ theory predictions and investigate their impact. The quality-of-fit in terms of $\chi^2/N_{\rm pt}$ (as well as the same values previously obtained with ePump), are reported in Table III.

The results of individual global fits using the 13 TeV ATLAS and CMS $t\bar{t}$ measurements in the various channels are discussed below.

A. Impact from the ATLAS lepton + jets channel

We start by analyzing the impact from the ATL13lj measurements [23] using the two bin resolutions, i.e., the

ATLAS and CMS binning, published by the ATLAS collaboration and available on the HEPData repository [194].

ATL13lj resolved with CMS bins. The ATLAS measurements resolved in terms of CMS bins share the bin size with the CMS13ll data that have equal IL [26]. Bin-by-bin statistical correlations are not available for this specific binning resolution. Looking at the individual $\chi^2/N_{\rm pt}$ values in Table III, the 1d distributions that are better described by the theory are $m_{t\bar{t}}$ and $y_{t\bar{t}}$, while the description of the $p_{T,t}$ and y_t distributions is more challenging due, in part, to the fact that they refer to reconstructed hadronically decayed top quarks. We note that both the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions have $\chi^2/N_{\rm pt} \lesssim 1$ independently of the choice of the central scale in their theory predictions. The $y_{t\bar{t}}$ spectrum appears to be more stable when the central scale is varied, with $\chi^2/N_{\rm pt} \approx 0.75$ in both cases. The description of the y_t distribution is largely affected by the central-scale choice.

The $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions provide us with most of the information and introduce the least tension when they are independently added on top of the CT18 baseline.

In Fig. 6 we compare the individual impact on the large-x gluon obtained using the two bin resolutions in a global fit at NNLO. Error bands with different hatching represent PDF uncertainties at 90% CL, and the central-scale choice in the 13 TeV $t\bar{t}$ theory predictions is set to $H_T/4$ (results are very similar for $H_T/2$). We observe that the constraints placed by the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions resolved with the CMS13ll bins have pulls in opposite directions above $x \gtrsim 0.3$. In particular, the $m_{t\bar{t}}$ distribution prefers a harder gluon while the $y_{t\bar{t}}$ one prefers a softer gluon as compared to CT18. Overall, the impact on the gluon PDF error is small. As reflected in Table III, both distributions exhibit $\chi^2/N_{\rm pt} \lesssim 1$ with small variations when the central-scale choice in their theory prediction is varied.

ATL13lj resolved with ATLAS bins. The impact from both the individual and cumulative (in terms of bin-by-bin statistically combined) 1d ATL13lj distributions, resolved with ATLAS bins, is analyzed together with the impact of statistical correlations. The most relevant information is obtained from the $m_{t\bar{t}}$, $y_{t\bar{t}}$, $H_T^{t\bar{t}}$, and $y_{t\bar{t}}^B$ distributions for which bin-by-bin statistical correlations are available. Their $\chi^2/N_{\rm pt}$ values from the NNLO global fit in Table III is of order 1 or less for all distributions, regardless of the central-scale choice in their theory predictions, with some deterioration when the four spectra are statistically combined and fitted together. This is expected as statistical correlations impose further constraints in the fit. Correlated systematic uncertainties are given in terms of the nuisance parameter representation.

By looking at the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions in Fig. 6, we observe that the pulls on the gluon at large x are in the same direction, i.e., both $m_{t\bar{t}}$ and $y_{t\bar{t}}$ prefer a harder gluon at $x \gtrsim 0.4$ and distortions are more pronounced for $m_{t\bar{t}}$. However, these distortions are milder as compared to those obtained from the same distributions resolved with the CMS13ll bins. The behavior of the gluon at large-x constrained by the inclusion of the $y_{t\bar{t}}^B$ and $H_T^{t\bar{t}}$ distributions in the fit, is very similar to that of the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions, respectively, and it is not shown here. The individual distribution impact on the PDF uncertainty is negligible also in this case.

Impact of bin-by-bin statistical correlations in the global fit. In Fig. 7 we illustrate the impact on the gluon PDF from the $m_{t\bar{t}}$, $y_{t\bar{t}}$, $H_T^{t\bar{t}}$, and $y_{t\bar{t}}^B$ distributions of the ATL13lj data added together on top of the CT18 baseline, and added together including statistical correlations. The gluon PDF is shown at a scale of $Q = 100$ GeV to emphasize the impact, while error bands with different hatching represent PDF
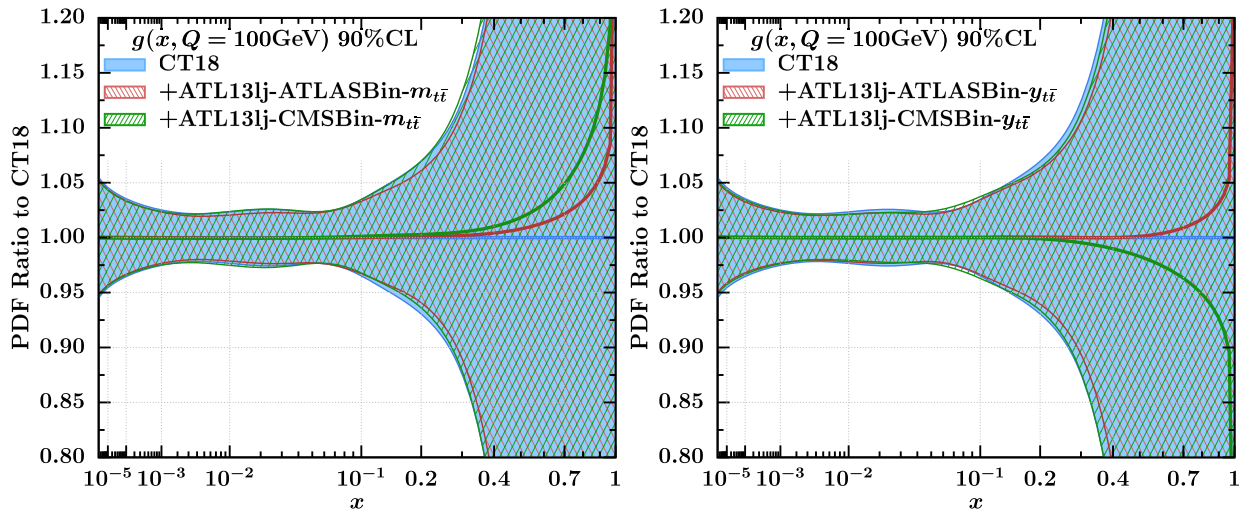

FIG. 6. *ATLAS* lepton + jet *channel*. Impact of the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions on the gluon PDF in a NNLO global fit. Distributions are shown for two different binning resolutions labeled by "ATLASBin" and "CMSBin" respectively. Error bands with different hatching represent PDF uncertainties at 90% CL. The central scale in their theory predictions is set to $H_T/4$.
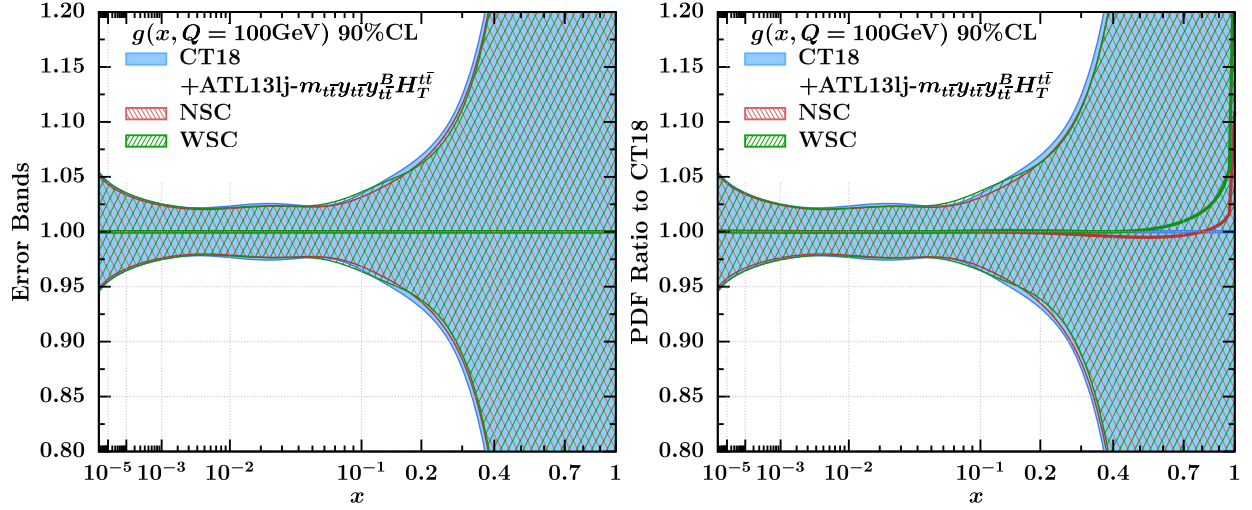
FIG. 7. *PDF impact of statistical correlations in the ATL13lj data.* Impact of statistical correlations on the gluon PDF in a NNLO global QCD analysis. Left: impact on the error bands. Right: PDF ratios to CT18 PDFs. The $y_{t\bar{t}}$, $y_{t\bar{t}}^B$, $m_{t\bar{t}}$, and $H_T^{t\bar{t}}$ distributions are added with (green) and without (red) bin-by-bin statistical correlations. The error bands with different hatching represent PDF uncertainties at 90% CL. The central-scale choice in the 13 TeV $t\bar{t}$ theory predictions is set to $H_T/4$.

uncertainties at 90% CL. Also here, the central-scale choice in the 13 TeV $t\bar{t}$ theory predictions is set to $H_T/4$. Considerations are similar for the scale choice $H_T/2$. The left panel in Fig. 7 shows that statistical correlations have a negligible impact on the gluon PDF errors. In the right panel, we show distortions in the gluon central value at large x in fits with and without statistical correlations. The impact of statistical correlations is very small and mostly localized at $x \gtrsim 0.6$ where there is weak or essentially no constraint from data. In addition, we observe that when the four distributions are added together, their effect on the gluon central value is diluted in the fit, and the quality-of-fit improves ($\chi^2/N_{\rm pt} \approx 1.3$ in Table III) when the central scale is set to $H_T/4$ as compared to $H_T/2$, in presence of statistical correlations. Without statistical correlations, a good description of the combined spectra is obtained by using either the $H_T/2$ ($\chi^2/N_{\rm pt} \approx 0.93$), or the $H_T/4$ ($\chi^2/N_{\rm pt} \approx 1.06$) scale choice, with a mild improvement in the $H_T/2$ case.

As discussed in Sec. II D, the small impact of statistical correlations from the ATL13lj data is confirmed by ePump. However, the observed trend of pulls on the gluon from ePump in Fig. 4 is in the opposite direction as compared to the global fit in Fig. 7. This is plausible, because the statistical procedure to update PDFs and their errors in ePump poses more restrictions as compared to a global PDF fit. In addition, $\chi^2/N_{\rm pt}$ may differ depending on the treatment of the errors (i.e., the "error type" option) in ePump (see for instance the $\chi^2/N_{\rm pt}$ values in Tables II and III). In Table III, the ePump $\chi^2/N_{\rm pt}$ values are obtained by using the "error type 1" option, where uncorrelated statistical and systematic errors are given, and correlated systematic errors (with percent correlation to each data point) are also given.

The correlated systematic errors are treated multiplicatively, i.e., by multiplying the percentages by the original best-fit theory predictions for each data point.

The interplay between the two bin resolutions in the ATL13lj measurements is further discussed in Sec. IV where optimal data combinations are selected. The two bin resolutions differ in number of bins and bin size. We have cross checked the theory predictions in both cases, with independent calculations from MATRIX [61,62] and the NNLO version of fastNLO tables [180], and we find consistency. Overall, we find that the gluon impact of the ATL13lj measurements in the global analysis is negligible.

B. Impact from the ATLAS all-hadronic channel

The impact on the gluon when the ATL13had measurements [24] are added on top of the CT18 baseline is illustrated in Fig. 8. The most relevant information in the global fit is obtained from the $y_{t\bar{t}}$, $m_{t\bar{t}}$, $H_T^{t\bar{t}}$, p_{T,t_1}, and p_{T,t_2} distributions which we study here. The $y_{t\bar{t}}$, $m_{t\bar{t}}$, and $H_T^{t\bar{t}}$ distributions produce visible impact on the large-x gluon at $x \gtrsim 0.5$. Pulls are all in the same direction and there is a preference for a softer gluon at large x, with a more pronounced effect from $H_T^{t\bar{t}}$ at $x \gtrsim 0.5$.

A milder impact is observed for $y_{t\bar{t}}$ and $m_{t\bar{t}}$ at $x \gtrsim 0.6$. The p_{T,t_1} and p_{T,t_2} distributions produce an almost identical behavior, with most of the impact located at $x \gtrsim 0.5$. The constraining power of these data in this kinematic region is limited by large experimental uncertainties that affect these distributions.

For all distributions, the impact on PDF uncertainties is negligible. The $\chi^2/N_{\rm pt}$ in Table III is of order 1 for the $y_{t\bar{t}}$ distribution regardless of the central scale choice in the theory predictions. Moreover, the $m_{t\bar{t}}$ and $H_T^{t\bar{t}}$ distributions
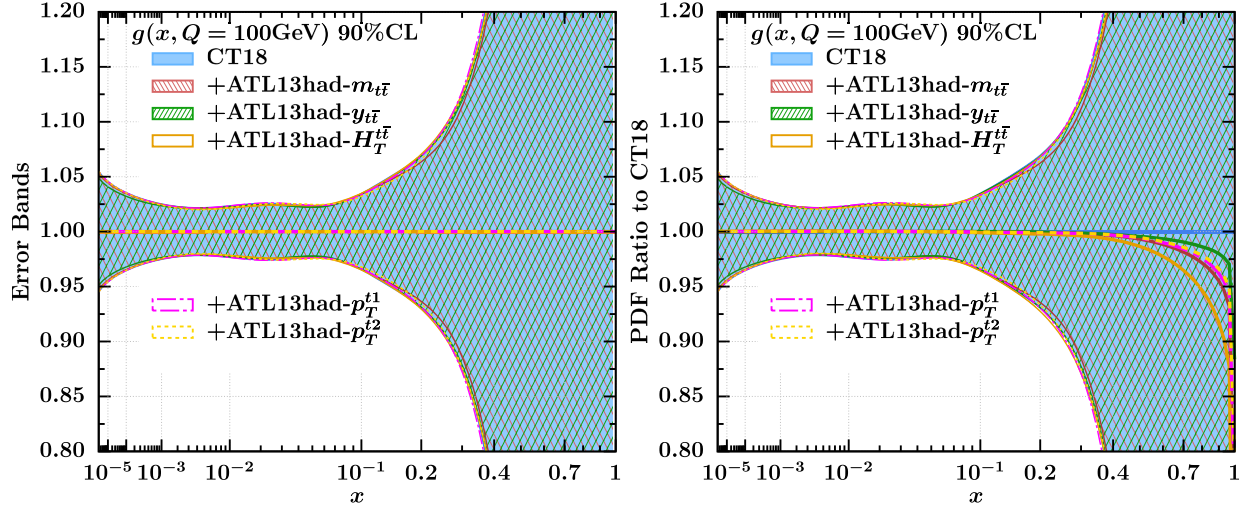
FIG. 8. *ATLAS all-hadronic channel.* Impact of the $y_{t\bar{t}}$, $m_{t\bar{t}}$, $H_T^{t\bar{t}}$, p_{T,t_1} and p_{T,t_2} distributions, added one at a time in the NNLO global QCD analysis. Error bands with different hatching represent PDF uncertainties at 90% CL. The central-scale choice in the theory predictions is set to $H_T/4$.

are in general not well described, while the χ^2/N_{pt} for individual fits using p_{T,t_1} and p_{T,t_2} depends on the central-scale choice. As before, correlated systematic uncertainties are given in terms of nuisance parameters. As for the ATL13lj data, the overall individual impact of the ATL13had measurements is negligible.

C. Impact from the CMS dilepton channel

In Fig. 9 we illustrate the individual PDF impact of the CMS13ll measurements [26]. The subset from which we gather the most relevant information to constrain the gluon is obtained by considering the $y_{t\bar{t}}$, $m_{t\bar{t}}$, y_t, and $p_{T,t}$

distributions. As discussed in Sec. II B, the correlated systematic uncertainties for these measurements are published in terms of the covariance matrix representation. Therefore, we perform conversion to nuisance parameters as discussed in Appendix D.

From the χ^2/N_{pt} values in Table III, the $m_{t\bar{t}}$ and $p_{T,t}$ distributions do not produce a good fit, regardless of the central-scale choice in their theory predictions. This is also confirmed by the results from ePump. The $y_{t\bar{t}}$ and y_t spectra are the distributions for which we obtain the best description with $\chi^2/N_{\text{pt}} \approx 1$ using the central-scale choice $H_T/2$, and with $\chi^2/N_{\text{pt}} \approx 0.7$ with central-scale set to $H_T/4$. We note in particular that the fit quality for the y_t distribution
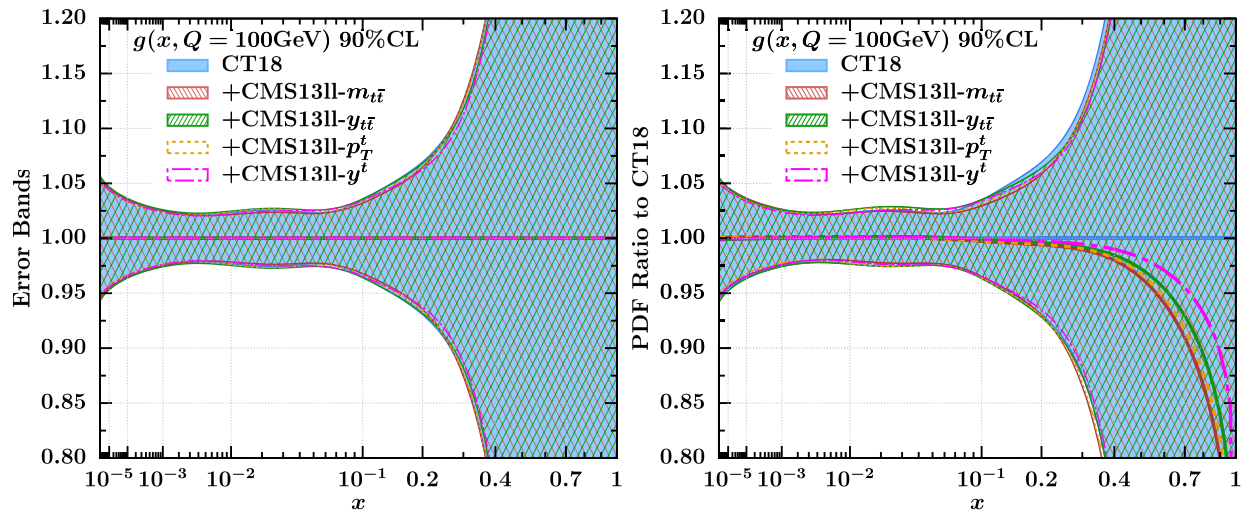

FIG. 9. *CMS dilepton channel.* Impact of the $y_{t\bar{t}}$, $p_{T,t}$, y_t, and $m_{t\bar{t}}$ distributions added one at a time on top of the CT18 baseline in the global fit. Error bands with different hatching represent PDF uncertainties at 90% CL. The central-scale choice in the 13 TeV $t\bar{t}$ theory predictions is $H_T/4$.
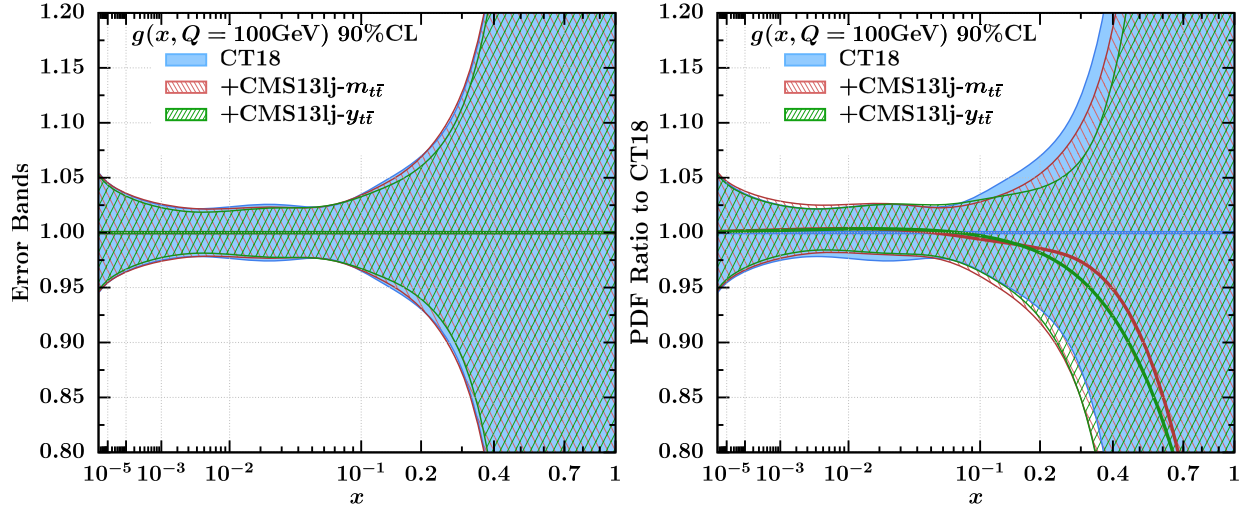
FIG. 10. *CMS* lepton + jets *channel.* Impact of the $y_{t\bar{t}}$, and $m_{t\bar{t}}$ distributions added one at a time in the NNLO global fit. Error bands with different hatching represent PDF uncertainties at 90% CL.

is similar to that observed in the NNPDF4.0 global fit ($\chi^2/N_{\mathrm{pt}} = 0.52$).

The gluon central value with central-scale choice $H_T/4$ in Fig. 9 is affected by the CMS13ll data in the $x \gtrsim 0.25$ region with a strong preference for a much softer large-x gluon as compared to the ATL13lj and ATL13had data. However, the impact on the gluon PDF uncertainty is very modest and comparable to that of the ATL13lj data.

In general, the y_t and $y_{t\bar{t}}$ distributions are expected to show less sensitivity to radiative corrections and m_t variations [199]. The χ^2/N_{pt} description of these measurements in the global fit from Table III is overall good for ATL13had, CMS13ll, and ATL13lj, with a very mild dependence on the scale choice for ATL13had and ATL13lj resolved with CMS bins. However, the constraining power of these datasets is modest due to larger errors as compared to the CMS13lj [28] data which are obtained with higher IL.

D. Impact from the CMS lepton + jets channel

The CMS collaboration released two sets of measurements in the lepton + jets channel, obtained with 35.8 fb^{-1} of IL [27] and with 137 fb^{-1} IL [28], respectively. We focus on the 137 fb^{-1} measurements as they extend the previous ones [27] and their precision is significantly improved. In addition, we observe that they place stronger constraints on the gluon in the global analysis, due to their improved systematic uncertainties which results in higher precision. Bin-by-bin statistical correlations are not available for these measurements.

We study the individual impact of the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions, which provide us with the most relevant information among the CMS13lj measurements in the

global fit. In Fig. 10, we illustrate the impact from these two measurements on the gluon PDF.

From the χ^2/N_{pt} values in Table III, we observe that the $y_{t\bar{t}}$ distribution cannot be well described in either the global QCD analysis or with ePump, with $\chi^2/N_{\mathrm{pt}} \approx 6$ or more, regardless of the scale choice. This can be in part ascribed to tensions with the jet production data and the $t\bar{t}$ pair production data at 8 TeV previously included in the CT18 analysis. However, another cause of disagreement may be related to the correlated systematic uncertainties associated to these measurements. We observe that even by varying the scale in the theory prediction, the $y_{t\bar{t}}$ distribution seems to have a preference for a too soft gluon as compared to jet data. A thorough study of the fit quality of the $y_{t\bar{t}}$ distribution would also require an investigation of amended versions of the CT18 gluon parametrizations. This will be presented in a forthcoming work. The $m_{t\bar{t}}$ distribution instead produces a better fit with a preference for a slightly harder gluon in the $0.2 \lesssim x \lesssim 0.65$ range as compared to the $y_{t\bar{t}}$ spectrum. Moreover, we note that the description of $m_{t\bar{t}}$ improves when the scale in the theory prediction is set to $H_T/2$.

We note that this dynamical scale choice is different from the default $H_T/4$ discussed in Ref. [199]. Throughout this work, the criterion according to which the optimal central scale for the theory prediction is chosen, is based on improvements in the quality-of-fit.

In Fig. 10, we observe that these measurements have stronger impact on the gluon uncertainty as compared to those previously analyzed. This is mainly ascribed to a better control of experimental uncertainties which in turn, enhances the constraining power of these data. It would be useful to examine the impact of bin-by-bin statistical correlations, should they become available.

IV. OPTIMAL COMBINATIONS OF 13 TeV TOP-QUARK PAIR PRODUCTION MEASUREMENTS IN THE GLOBAL ANALYSIS

The analysis of impact from individual 13 TeV $t\bar{t}$ differential cross section measurements in various channels at ATLAS and CMS, previously discussed in Sec. III, allows us to select optimal combinations of measurements from which we can extract maximum information to constrain the gluon PDF, and minimize tensions among data in the extended baseline. The guiding principle for this selection is essentially based on the quality of impact and quality of description in the global fit, in terms of χ^2. That is, we select those measurements that place effective constraints and produce a good quality fit, and that do not deteriorate the description of data in the preexisting baseline.

Looking at the χ^2/N_{pt} values and central-scale choice for the 13 TeV $t\bar{t}$ theory predictions in Table III, we observe that the $y_{t\bar{t}}$ distribution is well described (i.e., $\chi^2/N_{\text{pt}} \approx 1$) in both the ATL13had and CMS13ll cases. We therefore select the $y_{t\bar{t}}$ distribution from these measurements.

We select the $m_{t\bar{t}}$ distribution from the CMS13lj measurements, because the $y_{t\bar{t}}$ one does not produce a good fit ($\chi^2/N_{\text{pt}} \geq 5$ regardless of the central-scale choice).

Because of the differences observed in the two bin resolutions of the ATL13lj measurements, we consider two separate cases: in one case we include the $y_{t\bar{t}}$ distribution with the CMS bin resolution, whose χ^2 appears to be more stable against central-scale changes in the theory prediction as compared to $m_{t\bar{t}}$ (see Table III); in the other, we include the $m_{t\bar{t}}, y_{t\bar{t}}, y_{t\bar{t}}^B$, and H_T^T distributions added together without statistical correlations as the latter produce negligible impact on PDF uncertainties. In addition, this greatly simplifies the global analysis.

Therefore, we identify two optimal combinations which we refer to as CT18 + nTT1 and CT18 + nTT2,

respectively. The CT18 + nTT1 combination includes the ATL13had-$y_{t\bar{t}}$, CMS13ll-$y_{t\bar{t}}$, CMS13lj-$m_{t\bar{t}}$, and ATL13lj-$y_{t\bar{t}}$ distributions (resolved in terms of the CMS bins), while the CT18 + nTT2 combination includes the same distributions from the ATL13had, CMS13ll, CMS13lj measurements, and the $y_{t\bar{t}} + y_{t\bar{t}}^B + m_{t\bar{t}} + H_T^{t\bar{t}}$ combination without statistical correlations from ATL13lj, resolved with ATLAS bins.

A. Main results from the post-CT18 global analysis with extended baseline

The results of the NNLO global analysis obtained with the nTT1 and nTT2 combinations are summarized in Table IV, and illustrated in Figs. 14 and 15 where we show the gluon PDF ratio to CT18, and the $R_s = (s + \bar{s})/(\bar{u} + \bar{d})$ ratio. $g(x, Q)$ and $R_s(x, Q)$ are selected as representative cases as they show the most visible impact from the inclusion of the new data in the global fit.

In Table IV we compare the χ^2/N_{pt} values obtained from the CT18 + nTT1 and CT18 + nTT2 to CT18. There, we only report the CT18 datasets that exhibit a noticeable change in χ^2/N_{pt}. The remaining CT18 datasets have the same χ^2/N_{pt} as in Tables I and II in Ref. [130].

The CT18 NNLO fit [130] has $N_{\text{pt}} = 3681$ and $\chi^2 = 4293$ with $\chi^2/N_{\text{dof}} = 1.16$. In the CT18 + nTT1 fit, the total number of points is $N_{\text{pt}} = 3728$ and the resulting total χ^2 from the global fit is $\chi^2 = 4341$ when the central scale of the 13 TeV $t\bar{t}$ theory is set to $H_T/2$, while we obtain $\chi^2 = 4346$ when the same scale is set to $H_T/4$. In the CT18 + nTT2 fit, the total number of points is $N_{\text{pt}} = 3752$, $\chi^2 = 4366$ for the central-scale choice $H_T/2$, and $\chi^2 = 4376$ for $H_T/4$.

Comparing the χ^2 values of the individual experiments to those of CT18 in Table IV, we observe that the most noticeable quality-of-fit deterioration happens for both nTT1 and nTT2, regardless of the scale choice, in the case

TABLE IV. Data sets of the extended NNLO global QCD analysis including the optimal combinations CT18 + nTT1 and CT18 + nTT2. Here we directly compare the quality-of-fit found for CT18 + nTT1 and CT18 + nTT2 vs CT18 NNLO on the basis of χ^2/N_{pt} and scale choices $\{H_T/2, H_T/4\}$ for the central theory predictions of the $t\bar{t}$ data at 13 TeV. We only report the CT18 datasets for which the χ^2/N_{pt} exhibits a noticeable change.

| ID#	Experimental dataset		N_{pt}	CT18	$H_T/2$ nTT1	$H_T/2$ nTT2	$H_T/4$ nTT1	$H_T/4$ nTT2
110	CCFR F_2^p	[200]	69	1.1	1.1	1.1	1.2	1.2
545	CMS 8 TeV 19.7 fb^{-1}, single incl. jet cross section, $R = 0.7$, (extended in y)	[201]	185	1.1	1.2	1.2	1.2	1.2
573	CMS 8 TeV 19.7 fb^{-1}, $t\bar{t}$ norm. double-diff. top p_T and y cross section	[22]	16	1.2	1.2	1.2	1.1	1.1
580	ATLAS 8 TeV 20.3 fb^{-1}, $t\bar{t}$ $p_{T,t}$ and $m_{t\bar{t}}$ abs. spectrum	[19]	15	0.6	0.7	0.7	0.7	0.7
521	ATLAS 13 TeV 36.1 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section all-hadronic	[24]	12	$\cdots$	1.0	1.0	1.1	1.1
528	CMS 13 TeV 35.9 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section dilepton ch.	[26]	10	$\cdots$	0.8	0.8	0.5	0.7
532	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section l + j ch. CMS-bin	[23]	10	$\cdots$	0.7	$\cdots$	0.8	$\cdots$
587	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}, m_{t\bar{t}}, y_{t\bar{t}}^B, H_T^{t\bar{t}}$ cross sections l + j ch.	[23]	34	$\cdots$	$\cdots$	1.0	$\cdots$	1.1
581	CMS 13 TeV 137 fb^{-1}, $t\bar{t}$ abs. $m_{t\bar{t}}$ cross section l + j ch.	[28]	15	$\cdots$	1.1	1.1	1.6	1.7
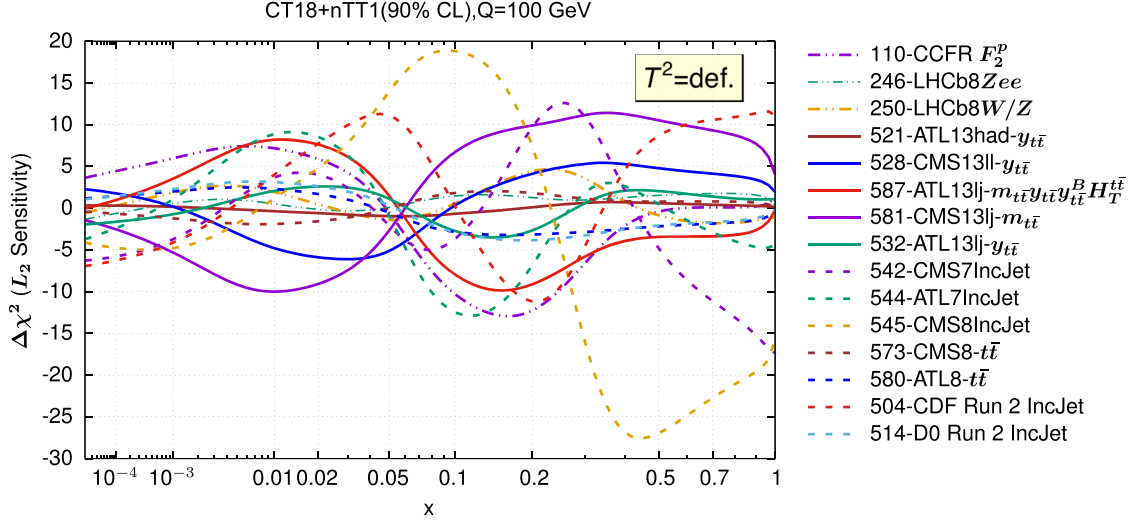
FIG. 11. L_2 sensitivity study for the gluon PDF: $t\bar{t}$ and jet production in the CT18 + nTT1 fit.

of the CMS 8 TeV single inclusive jet cross section [201] measurements, and for the CMS13lj $m_{t\bar{t}}$ distribution when the scale in the 13 TeV $t\bar{t}$ theory is set to $H_T/4$.

In addition, a different PDF suppression preference over a wide range of x for both nTT1 and nTT2 is observed between the LHCb 8 TeV $Z \to e^-e^+$ forward rapidity cross section [202] and the CMS13lj $m_{t\bar{t}}$ distribution. The LHCb 8 TeV measurements have an impact on the strange PDF, anti-quarks (e.g., $\bar{u}, \bar{d}$), and their errors at both small and large x. In particular, the $\bar{u}$ and $\bar{d}$ PDFs are impacted in the $10^{-4} \lesssim x \lesssim 10^{-2}$ range. This interplay as well as that between the top-quark and the jet data can be further understood in terms of the L_2 sensitivity [203], that is a statistical tool to explore the pulls from individual measurements on the best-fit PDFs, and to identify tensions between competing datasets. It is important to point out that the use of the L_2 sensitivity in this study follows a different approach because it is used to study PDF suppression preferences using PDFs that are obtained after a global fit.

To illustrate the different suppression-pattern preference between the new $t\bar{t}$ and jet production data, in Fig. 11 we present the L_2 sensitivity plot for the gluon PDF in the CT18 + nTT1 fit. The plot relative to the CT18 + nTT2 gluon is very similar and we omit it here. The $T^2 = $ def in the inset label refers to the default tolerance criteria adopted in the CT18 analysis. From the gluon suppression patterns in Fig. 11, we observe for example that the CMS13lj-$m_{t\bar{t}}$ and the CMS13ll-$y_{t\bar{t}}$ data have opposite trend with respect to the CMS 8 TeV inclusive jet production data in the $x > 0.3$ range, where CMS13lj-$m_{t\bar{t}}$ and CMS13ll-$y_{t\bar{t}}$ prefer a softer gluon. We note that a positive value of $\Delta\chi^2$ in the L_2 sensitivity plot, cf. Fig. 11, indicates the preference of decreasing the gluon PDF. Namely, a positive sensitivity indicates a negative pull on the corresponding PDF (see further discussion in Ref. [203]). Another different suppression-pattern trend is observed for the ATLAS 7 TeV

inclusive jet data, which prefer a softer gluon in the $10^{-3} \lesssim x \lesssim 0.05$ range as compared to CMS13lj-$m_{t\bar{t}}$ and CMS13ll-$y_{t\bar{t}}$.

In Figs. 12 and 13, we illustrate the L_2 sensitivity for the CMS13lj-$m_{t\bar{t}}$ distribution and the LHCb 8 TeV 2.0 fb^{-1} $Z \to e^-e^+$ forward rapidity cross section measurements in the CT18 + nTT1 global fit. As before, the same figures for the CT18 + nTT2 global fit are very similar and we do not show them here.

In Fig. 12, we observe that the $m_{t\bar{t}}$ distribution in CMS13lj has a strong preference for a softer gluon in the $0.1 \lesssim x \lesssim 0.7$ range, as compared to the LHCb 8 TeV $Z \to e^-e^+$ large rapidity data, which prefer a harder gluon in the same range. In addition, we observe a different suppression preference for the u, d, $\bar{u}$, $\bar{d}$, and s quarks across the entire x range, where in particular we note a higher preference for softer u- and $\bar{u}$-quark PDFs in the small-x ($10^{-4} \lesssim x \lesssim 10^{-2}$) range by the LHCb 8 TeV Z rapidity measurements.

These different suppression preferences have an impact on quark ratios, in particular on R_s. The L_2 sensitivity for the R_s ratio is illustrated in Fig. 13 where u_V, d_V and other PDF ratios are also shown.[3]

The increase in the χ^2 values in Table IV and the different suppression preferences discussed above are in part reflected by an increase in the uncertainty of the R_s ratio, as compared to CT18. This is illustrated in Figs. 14 and 15. A more flexible choice for the strange-quark parametrization allowing $s \ne \bar{s}$ in the global fit, such as CT18As or CT18As_Lat NNLO analysis with lattice QCD constraints [204], may in general be beneficial to the description of the data. However, the different suppression preferences for the u, d, $\bar{u}$, and $\bar{d}$ quarks across the entire range of x also play

---

[3]In the CT18 analysis $s = \bar{s}$, therefore $R_s = 2s/(\bar{u} + \bar{d})$.
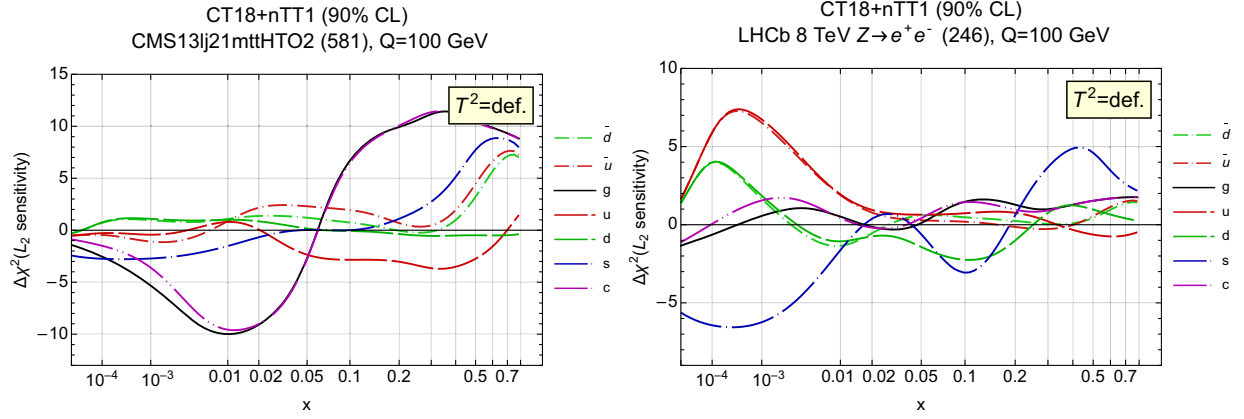
FIG. 12. L_2 sensitivity study to explore the interplay between the CMS13lj (left) and LHCb 8 TeV 2.0 fb^{-1} $Z \to e^- e^+$ forward rapidity cross section measurements (right).
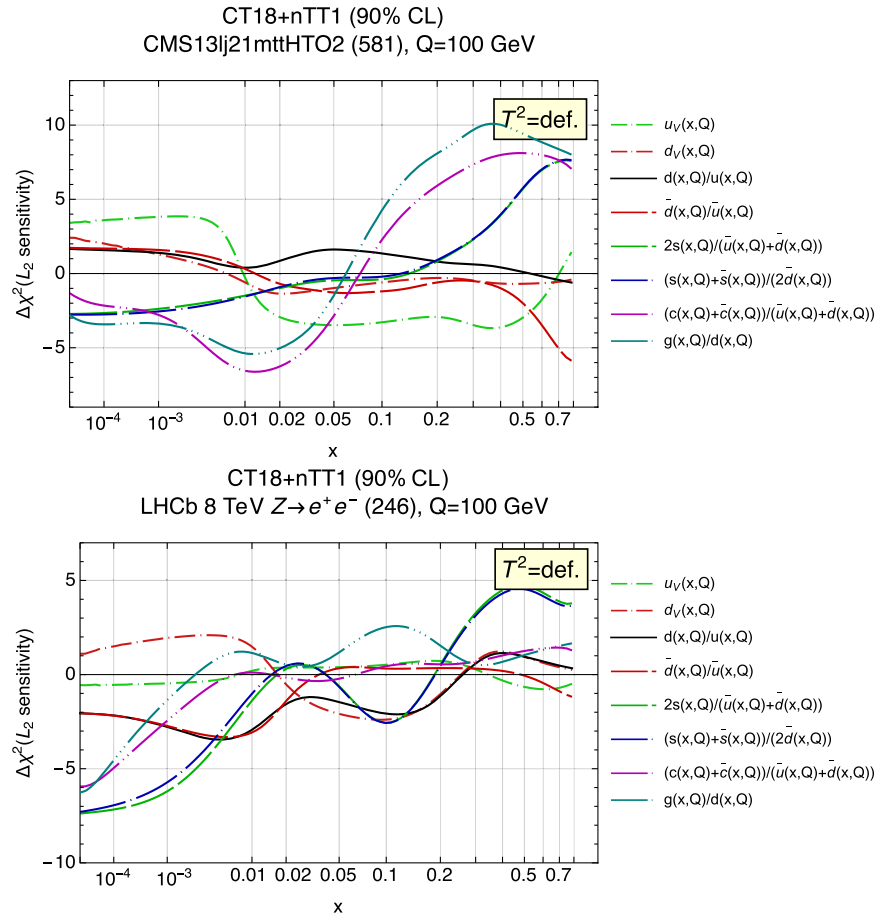

FIG. 13.   Same as in Fig. 12, but for different PDF ratios.

an important role in the determination of the uncertainty associated to R_s.

The R_s uncertainty is shown in the right plot of Figs. 14 and 15 for CT18 + nTT1 and CT18 + nTT2, respectively. We note that the uncertainty increase in the $0.3 \lesssim x \lesssim 0.8$ region has a different shape in both the CT18 + nTT1 and

CT18 + nTT2 fits for the two different scale choices $H_T/2$ and $H_T/4$. The left plots in Figs. 14 and 15 illustrate the resulting changes in the gluon PDF and its uncertainty, as well as (small) variations induced by selecting the two central scales $H_T/2$, and $H_T/4$, in the $t\bar{t}$ theory predictions at 13 TeV. For both scale choices we obtain a reduction in
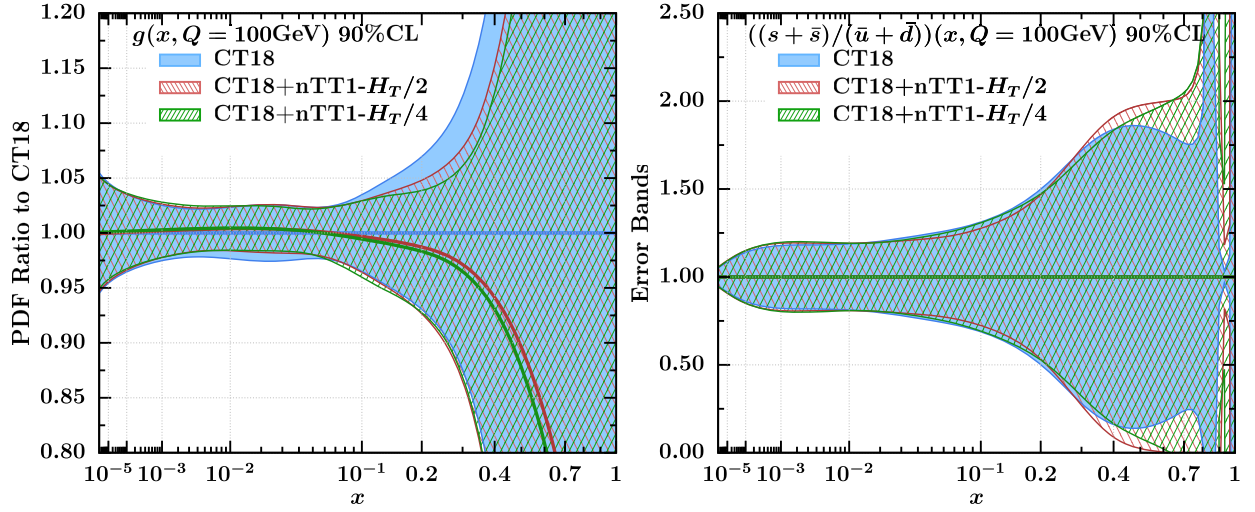
FIG. 14. Impact of $CT18 + nTT1$ on the global QCD analysis at NNLO. Left: gluon PDF ratio to CT18 NNLO. Right: $R_s(x, Q) = (s + \bar{s})/(\bar{u} + \bar{d})$. Error bands with different hatching represent different choices for the central scale in the 13 TeV $t\bar{t}$ theory predictions: $H_T/4$ (green), and $H_T/2$ (red). PDF uncertainties are evaluated at the 90% CL.
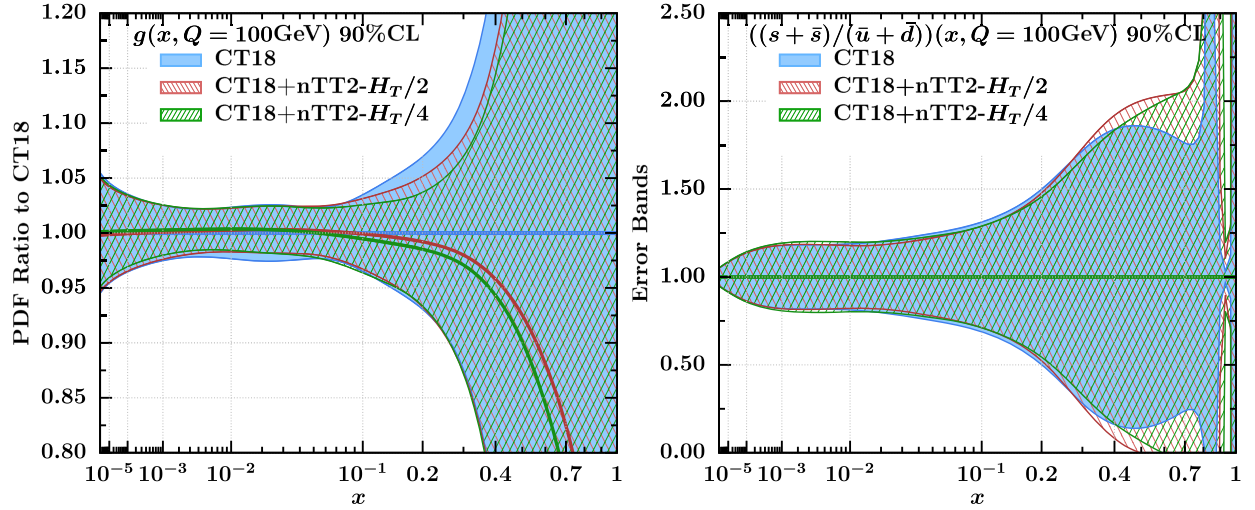
FIG. 15. Same as in Fig. 14 but for $CT18 + nTT2$.

the gluon uncertainty in the $0.2 \lesssim x \lesssim 0.5$ range and above in the extrapolation region, and a much less pronounced reduction in the $2 \times 10^{-3} \lesssim x \lesssim 5 \times 10^{-2}$ range.

We observe a preference for a softer gluon as compared to CT18 in the $10^{-1} \lesssim x \lesssim 0.6$ region. This constraint is mainly driven by the $m_{t\bar{t}}$ distribution from the CMS13lj data, with an improved description when the central scale is set to $H_T/2$. A data vs theory comparison for the CMS13lj $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions, obtained with CT18nTT1 and $CT18 + nTT2$ PDFs, respectively, is shown in Fig. 25 in Appendix B.

While the quality-of-fit of the two combinations $CT18 + nTT1$ and $CT18 + nTT2$ is essentially the same as that of CT18 with $\chi^2/N_{\text{dof}} \approx 1.16$, small differences between the

$CT18 + nTT1$ and $CT18 + nTT2$ global fits are noticeable in Figs. 14 and 15 as compared to CT18. These are discussed in more detail in Sec. IV B.

B. Interplay between top-quark and jet production data

Top-quark pair production and inclusive jet production at the LHC place complementary constraints on the gluon and other PDFs as these two processes overlap in the Q-x kinematic plane. It is therefore interesting to further examine the interplay between top-quark pair and jet production data in NNLO global analyses with the $CT18 + nTT1$ and $CT18 + nTT2$ combinations. In particular, we analyze the individual constrains placed on the gluon from these processes separately.

To this purpose, we perform three main fits in each of which we study the impact of different central scales in the 13 TeV $t\bar{t}$ theory predictions. These global fits are performed by using the CT18 baseline as the starting point, where we remove either the jet or the 8 TeV $t\bar{t}$ data, or both, and add the nTT1 combination with scale choice $H_T/2$ or $H_T/4$. The same fits performed with the nTT2 combination produce similar results.

The complementarity and tensions of jet and top data has also been analyzed before for 8 TeV top-quark pair production measurements in the context of the PDF4LHC21 studies [203,205,206] as well as in the recent PDF analyses in Refs. [130,131,134,207].

Global fits without inclusive jet data. In a first global fit, which we label CT18 mJet, PDFs are obtained from the CT18 fit by removing the inclusive jet production data. To explore the impact of the new 13 TeV $t\bar{t}$ combinations we consider two variants of this fit including the nTT1 datasets and where the theory predictions for the latter are included with different central scales. These fit variants are: (i) the CT18 baseline without QCD jets and with nTT1 datasets, with central scale $H_T/2$, labeled as CT18mJet + nTT1-$H_T/2$; and (ii) the CT18 baseline without QCD jets and with nTT1 datasets, with central scale $H_T/4$, labeled as CT18mJet + nTT1-$H_T/4$.

We focus our attention on the gluon which is the PDF receiving most of the impact. The results for the gluon PDF and its uncertainties resulting from these global analyses are illustrated in Fig. 16 where they are compared to the CT18 gluon at $Q = 100$ GeV.

Looking at the PDF ratio plot in Fig. 16 (left), we observe that removing the QCD jet data from the CT18 baseline (CT18mJet) results in an increase of the PDF uncertainty at large x in the $0.1 \lesssim x \lesssim 0.5$ range where there is a preference for a harder gluon with a bump in the $0.2 \lesssim x \lesssim 0.6$ region. Upon inclusion of the nTT1 combination in

the CT18 mJet+nTT1-$H_T/2$ and CT18 mJet+nTT1-$H_T/4$ global fits, we observe a softer gluon in the $x \gtrsim 0.1$ range, regardless of the central-scale choice.

Comparing the gluon error bands in Fig. 16 (right), we note that uncertainties are only marginally reduced when nTT1 is included, with a small increase in the $0.06 \lesssim x \lesssim 0.15$ region for scale choice $H_T/4$. This is due to small tensions with several datasets in the baseline, in particular the D0 run II 9.7 fb^{-1} electron charge asymmetry A_{ch}, with $p_{Tl} > 25$ GeV measurements [208], and the structure function measurements F_2^p [209] and F_2^d [210] from the BCDMS collaboration. From an L_2 sensitivity study, one can see that the latter have a preference for a harder gluon PDF at large x as compared to the nTT1 data, while the D0 run II electron charge asymmetry A_{ch} data are sensitive to the u- and $\bar{u}$-quark PDFs with opposite trend as compared to BCDMS F_2^d.

Global fits without $t\bar{t}$ data. In a second global fit, labeled as CT18mTop8, PDFs are obtained from the CT18 fit by removing top-quark pair production data (at $\sqrt{S} = 8$ TeV only in CT18). As before, we consider two fit variants to assess the impact of the 13 TeV $t\bar{t}$ data, where the nTT1 combination is included with scale choices $H_T/2$ and $H_T/4$ in the theoretical predictions for the nTT1 data subset. These fit variants are labeled as CT18mTop8 + nTT1-$H_T/2$ and CT18mTop8 + nTT1-$H_T/4$, respectively. The results for the gluon PDF are shown in Fig. 17 where we illustrate the impact. This is complementary to that in Fig. 16.

We note that when the 8 TeV $t\bar{t}$ measurements are removed from the CT18 baseline, the impact is negligible on both the central value and uncertainty of the gluon. When the nTT1 combination is included, there is a preference for a softer gluon at large x in the $0.15 \lesssim x \lesssim 0.9$ region. However, the impact on the gluon uncertainty is negligible, as shown in Fig. 17 (right), regardless of the
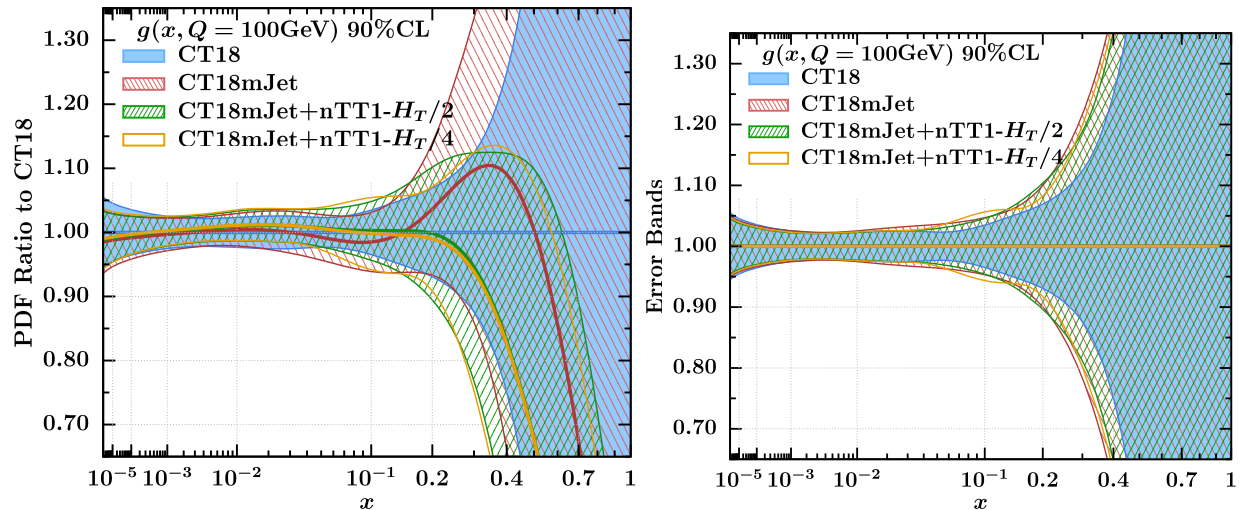

FIG. 16.   Global fit without jet data vs CT18NNLO. Left: PDF ratio to CT18. Right: error bands comparison.
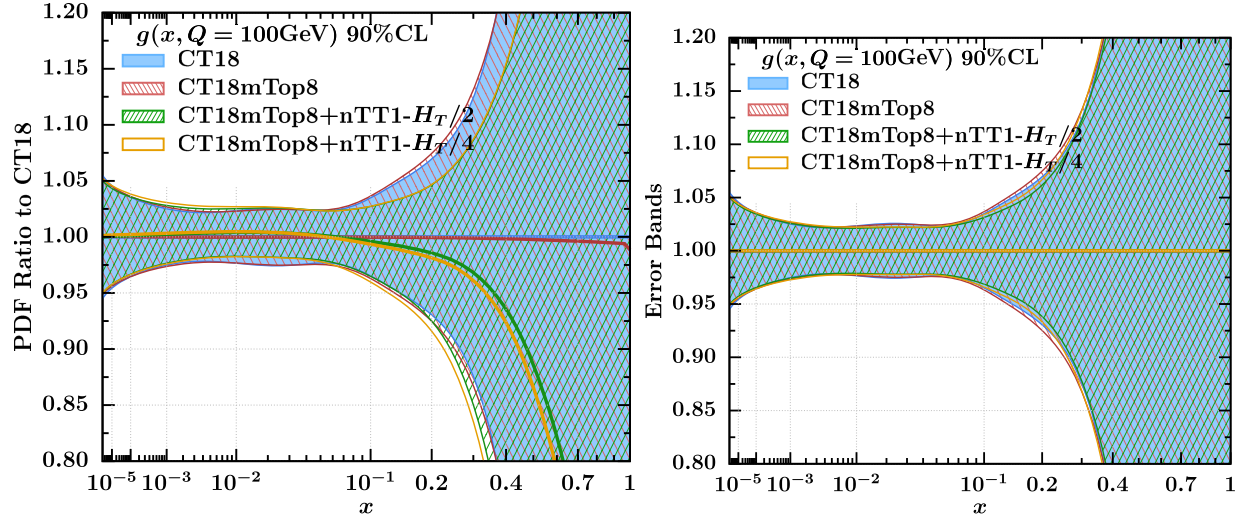
FIG. 17.   Global fit without $t\bar{t}$ data vs CT18NNLO. Left: PDF ratio to CT18. Right: error bands comparison.

scale choice. Moreover, we observe that the central value of the gluon in the nTT1 fits is harder as compared to the CT18mJet + nTT1-$H_T/2$ and CT18mJet + nTT1-$H_T/4$ fits. This is due to the impact of inclusive jet production data which prefer a softer gluon PDF in the large-x region.

Global fits without QCD jets and $t\bar{t}$ production. Finally, in a third fit labeled as CT18 mJet&Top8, PDFs are obtained by removing both the inclusive jet and $t\bar{t}$ data from the CT18 baseline. The results of this fit are then compared to the fit variants obtained by including the nTT1 data subset with central scale choice $H_T/2$ for the nTT1 theory predictions labeled as CT18mJet & Top8 + nTT1-$H_T/2$, and to that obtained with $H_T/4$, labeled as CT18mJet & Top8 + nTT1-$H_T/4$.

The results are shown in Fig. 18 where we observe that the general trend is similar to that in Fig. 16. However, we note that the 8 TeV $t\bar{t}$ data lead to an approximately 8%

reduction in the gluon central value in the $0.2 \lesssim x \lesssim 0.6$ region [compare the red solid line in Fig. 16 (left) to that in Fig. 18 (left)].

Looking at the PDF uncertainties in Fig. 18 (right), we note only a small increase in the $0.15 \lesssim x \lesssim 0.4$ range as compared to the CT18mJet fit, which reflects the small impact of the 8 TeV $t\bar{t}$ data in the CT18 fit. When the nTT1 combination is included, the gluon central value becomes softer in the $0.2 \lesssim x \lesssim 0.6$ range similar to that of the CT18mJet + nTT1-$H_T/2$ and CT18mJet + nTT1-$H_T/4$ fits with a reduction of the uncertainty in the $0.2 \lesssim x \lesssim 0.5$ range as compared to the CT18mJet&Top8 fit [see Fig. 18 (right)].

The conclusions we draw by comparing the results of the three main fits discussed above is that inclusive jet data place stronger constraints on the gluon as compared to top-quark pair production. This is mainly ascribed to the much
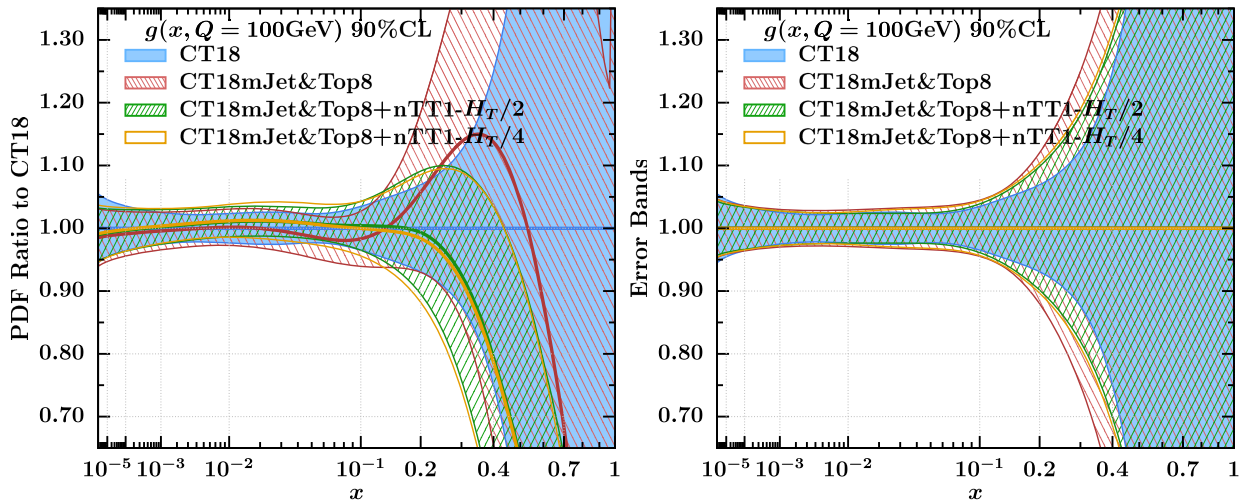

FIG. 18.   Global fit without $t\bar{t}$ and jet data vs CT18NNLO. Left: PDF ratio to CT18. Right: error bands comparison.

TABLE V. Same as in Table IV, but for the CT18mTop, CT18mJet, and CT18mJet & Top8 global fits.

				$H_T/2$	$H_T/2$	$H_T/4$	$H_T/4$
ID#	Experimental dataset	N_{pt}	CT18mTop	nTT1	nTT2	nTT1	nTT2
110	CCFR F_2^p	[200] 69	1.1	1.1	1.1	1.2	1.2
545	CMS 8 TeV 19.7 fb^{-1}, single incl. jet cross section, $R = 0.7$, (extended in y)	[201] 185	1.1	1.2	1.2	1.2	1.2
521	ATLAS 13 TeV 36.1 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section all-hadronic	[24] 12	$\cdots$	1.0	1.0	1.1	1.1
528	CMS 13 TeV 35.9 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section dilepton ch.	[26] 10	$\cdots$	0.8	0.9	0.5	0.6
532	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section l + j ch. CMS-bin	[23] 10	$\cdots$	0.7	$\cdots$	0.7	$\cdots$
587	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}, m_{t\bar{t}}, y_{t\bar{t}}^B, H_T^{t\bar{t}}$ cross sections l + j ch.	[23] 34	$\cdots$	$\cdots$	1.0	$\cdots$	1.1
581	CMS 13 TeV 137 fb^{-1}, $t\bar{t}$ abs. $m_{t\bar{t}}$ cross section l + j ch.	[28] 15	$\cdots$	1.1	1.2	1.6	1.6
	CT18mJet		CT18mJet				
110	CCFR F_2^p	[200] 69	1.2	1.2	1.2	1.2	1.2
573	CMS 8 TeV 19.7 fb^{-1}, $t\bar{t}$ norm. double-diff. top p_T and y cross section	[22] 16	1.2	1.2	1.2	1.2	1.1
580	ATLAS 8 TeV 20.3 fb^{-1}, $t\bar{t}$ $p_{T,t}$ and $m_{t\bar{t}}$ abs. spectrum	[19] 15	0.7	0.7	0.7	0.7	0.7
521	ATLAS 13 TeV 36.1 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section all-hadronic	[24] 12	$\cdots$	1.0	1.0	1.0	1.0
528	CMS 13 TeV 35.9 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section dilepton ch.	[26] 10	$\cdots$	0.6	0.7	0.4	0.5
532	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section l + j ch. cms-bin	[23] 10	$\cdots$	0.5	$\cdots$	0.6	$\cdots$
587	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}, m_{t\bar{t}}, y_{t\bar{t}}^B, H_T^{t\bar{t}}$ cross section l + j ch.	[23] 34	$\cdots$	$\cdots$	0.9	$\cdots$	1.1
581	CMS 13 TeV 137 fb^{-1}, $t\bar{t}$ abs. $m_{t\bar{t}}$ cross section l + j ch.	[28] 15	$\cdots$	0.9	1.0	1.3	1.5
	CT18mJet & Top		CT18mJet & Top				
110	CCFR F_2^p	[200] 69	1.2	1.2	1.2	1.2	1.2
521	ATLAS 13 TeV 36.1 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section all-hadronic	[24] 12	$\cdots$	1.0	1.0	1.0	1.0
528	CMS 13 TeV 35.9 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section dilepton ch.	[26] 10	$\cdots$	0.5	0.7	0.3	0.5
532	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}$ cross section l + j ch. CMS-bin	[23] 10	$\cdots$	0.4	$\cdots$	0.5	$\cdots$
587	ATLAS 13 TeV 36 fb^{-1}, $t\bar{t}$ abs. $y_{t\bar{t}}, m_{t\bar{t}}, y_{t\bar{t}}^B, H_T^{t\bar{t}}$ cross section l + j ch.	[23] 34	$\cdots$	$\cdots$	1.0	$\cdots$	1.1
581	CMS 13 TeV 137 fb^{-1}, $t\bar{t}$ abs. $m_{t\bar{t}}$ cross section l + j ch.	[28] 15	$\cdots$	0.8	1.0	1.2	1.4

larger number of data points in the inclusive jet measurements which tend to dilute the impact from $t\bar{t}$ production.

However, the impact from the 13 TeV $t\bar{t}$ production data results in a softer gluon at large x, similar to that of the LHC jet data, but with a very different degree of suppression. In fact, the phase-space suppression in the hard scattering contributions for these two processes is different and generates different shapes and suppression in Figs. 16 (left) and 17 (left). Moreover, we note that most of the gluon suppression at large x is driven by the CMS13lj data with 137 fb^{-1} of IL.

In Table V we summarize the most noticeable changes in the χ^2/N_{pt} values, as described in Table IV, but for the CT18mTop, CT18mJet, and CT18mJet & Top8 global fits. Overall, the description of the new 13 TeV $t\bar{t}$ data improves after removing the 8 TeV $t\bar{t}$ and/or jet data from the CT18 baseline. In particular, one can consider the CT18 + nTT1-$H_T/2$ global fit where the $\chi^2/N_{pt} = 1.1$ for the CMS13lj-$m_{t\bar{t}}$ data. In the CT18mJet fit the χ^2/N_{pt} decreases to 0.9 and in CT18mJet & Top8 + nTT1-$HT/2$ it further reduces to 0.8. Similarly, for the CMS13ll-$y_{t\bar{t}}$ data, one finds $\chi^2/N_{pt} = 0.8$ in the CT18 + ntt1-$HT/2$ fit and this reduces to $\chi^2/N_{pt} = 0.6$ in the CT18mJet & Top8 + ntt1-$HT/2$ fit. In addition, an L_2 sensitivity study

is shown in Sec. IV, where Fig. 11 displays the impact on the gluon PDF from both $t\bar{t}$ and jet production data as well as the different suppression preferences among the experiments.

Finally, we observe that all of the changes in shape and magnitude in the gluon central value from the impact of the new measurements are within the 90% uncertainty band obtained from the CT18 NNLO fit.

Differences between the nTT1 and nTT2 data subsets. To conclude this section, we wish to discuss differences between the nTT1 and nTT2 data subsets and their interplay with the 8 TeV top-quark data in the CT18 baseline. To facilitate this analysis and better emphasize the differences, we study the impact of the nTT1 and nTT2 data subsets on the gluon PDF uncertainty in the CT18mJet global fit, where inclusive jet data are removed. The results of this study are illustrated in Fig. 19.

In Fig. 19 (left) we observe an increase in the gluon uncertainty in the $0.06 \lesssim x \lesssim 0.15$ region when the central scale in the nTT1 theory predictions is set to $H_T/4$. The same increase is not present for the nTT2 combination in Fig. 19 (right). This is one of the main differences which emerges in the nTT1 combination containing the $y_{t\bar{t}}$ distribution from the ATL13lj measurements resolved in
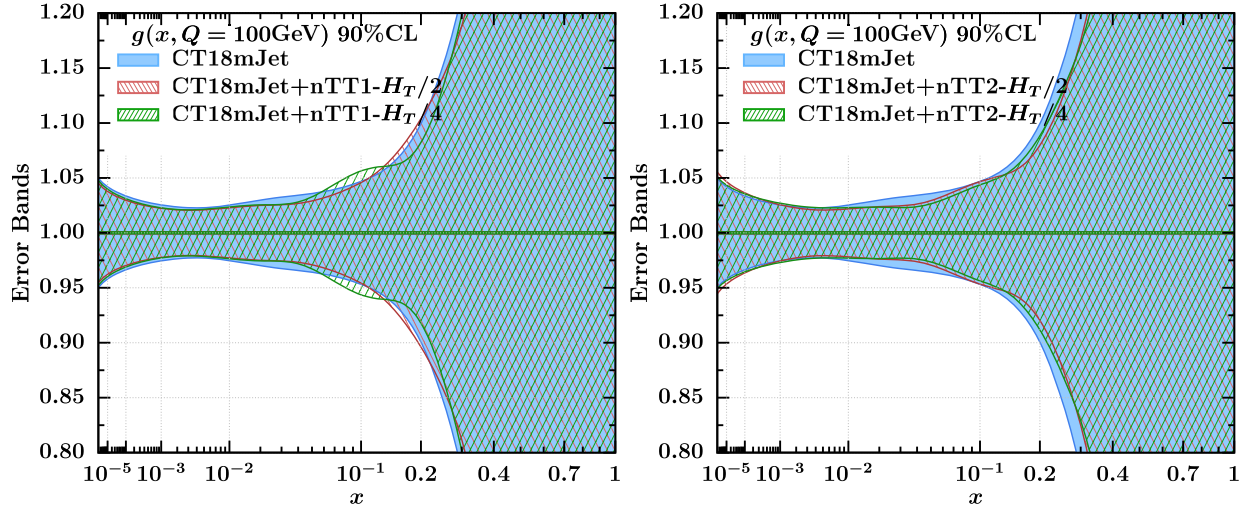
FIG. 19. Interplay between 13 TeV and 8 TeV top-quark data in a fit without QCD jets. Left: CT18 without QCD jets, and with nTT1. Right: CT18 without QCD jets, and with nTT2.

terms of CMS bins that uses $H_T/4$ as the central scale in the nTT1 theory predictions. This tension disappears when the $H_T/2$ central scale choice is used.

By looking at the $\chi^2/N_{\rm pt}$ values in Table IV, we argue that this mild tension is generated in part by the CMS 8 TeV 19.7 fb^{-1} single inclusive jet cross section, whose $\chi^2/N_{\rm pt}$ increases from 1.1 to 1.2 (regardless of scale choice); in part by the ATLAS 8 TeV 20.3 fb^{-1} top-quark $p_{T,t}$ and $m_{t\bar{t}}$ absolute distributions, with a $\chi^2/N_{\rm pt}$ increase from 0.6 to 0.7 (regardless of scale choice), and finally, by the ATL13lj data in nTT2 with a $\chi^2/N_{\rm pt}$ increase from 0.7 to 1.1 when the central scale goes from $H_T/2$ to $H_T/4$, and by the CMS13lj data in nTT1 and nTT2 with a $\chi^2/N_{\rm pt}$ increase from 1.1 to 1.6/1.7 when the central scale is reduced to $H_T/4$. As already pointed out in Sec. III D, this indicates a preference for the $H_T/2$ scale choice in the 13 TeV $t\bar{t}$ theory predictions in contrast to the suggested $H_T/4$ scale choice discussed in Ref. [199]. In our global fits, the optimal central scale for the 13 TeV $t\bar{t}$ theory predictions is chosen according to the improvements it produces in the quality-of-fit.

It is also interesting to look at the L_2 sensitivity plots for the ATL13lj measurements in the full CT18 + nTT1 and CT18 + nTT2 gloabl fits which we illustrate in Fig. 20. There, in the left panel we show the ALT13lj $y_{t\bar{t}}$ distribution resolved in terms of CMS bins. In the right panel, we show the $y_{t\bar{t}}$ resolved with ATLAS bins, combined with the $y_{t\bar{t}}^B$, $H_T^{t\bar{t}}$, and $m_{t\bar{t}}$ distributions without bin-by-bin statistical correlations. As already argued in previous discussions, the ATL13lj bin-by-bin statistical correlations have negligible impact and can essentially be ignored. We note that the two different treatments of the ATL13lj measurements lead to different preferences for the gluon and strange PDFs in the $0.2 \lesssim x \lesssim 0.7$ range: the ATL13lj $y_{t\bar{t}}$ distribution resolved

with CMS bins prefers softer gluon and softer strange PDFs, while the ATL13lj combination has opposite behavior.

ePump *optimization.* To further investigate differences in the CT18 + nTT1 and CT18 + nTT2 combinations, it is interesting to identify eigenvector (EV) directions that have maximal PDF sensitivity. The ePump framework allows us to identify a reduced set of error PDFs that contain the majority of the PDF dependence of the observables under consideration. The procedure is entirely based on the Hessian method and is documented in Ref. [146]. The new eigenvectors contain exactly the same information as the original eigenvectors, but are optimized so that a smaller set of error PDFs can be chosen for use with the set of observables to any required PDF-sensitivity. In turn, this allows us to assess and validate the data that place the strongest constraints on PDF errors.

In Fig. 21, we illustrate the behavior of the optimized eigenvector directions for the CT18 + nTT1 and CT18 + nTT2 combinations, while fractional contributions to the PDF error from the leading eigenvectors for each individual dataset are reported in Tables VI and VII. With ePump, we find six optimized PDF error sets for both CT18 + nTT1 and CT18 + nTT2.

In the case of the CT18 + nTT1 optimized directions, we observe that the three leading eigenvectors approximately account for 99% of the PDF error band. Among them, the largest contribution is from CMS13lj $m_{t\bar{t}}$, with fractional uncertainty of approximately 31%. The second largest contribution is from ATL13had $y_{t\bar{t}}$, with roughly 25%, and the smallest contributions are from CMS13ll $y_{t\bar{t}}$ and ATL13lj $y_{t\bar{t}}$ resolved in terms of CMS bins, both with fractional uncertainty of approximately 22%.

In the case of CT18 + nTT2, the data with the largest contribution is the ATL13lj combination of $m_{t\bar{t}}\,y_{t\bar{t}}$, $y_{t\bar{t}}^B$, and $H_T^{t\bar{t}}$, which account for approximately 47% of fractional
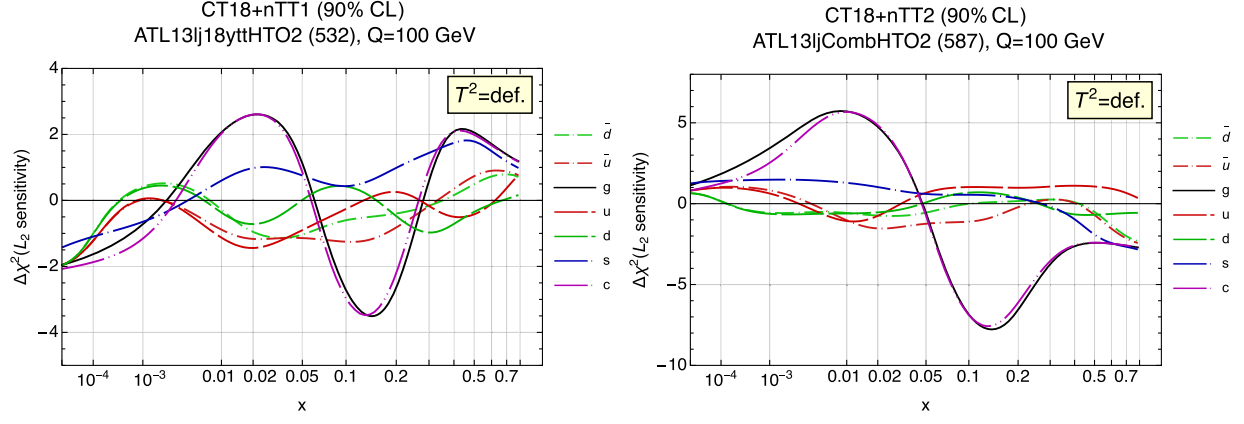
FIG. 20. L_2 sensitivity of the various PDFs for the ATL13lj measurements in the CT18 + nTT1 (left) and CT18 + nTT2 (right) global fits. Here, the central scale in the 13 TeV $t\bar{t}$ theory predictions is set to $H_T/2$.
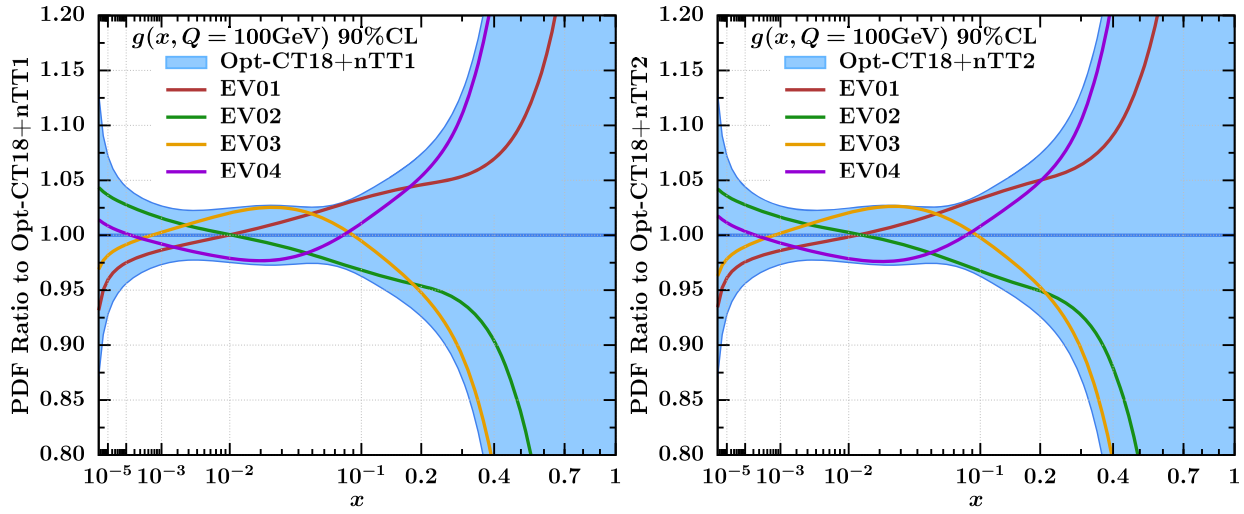

FIG. 21. Optimized eigenvector directions with ePump for the CT18 + nTT1 (left) and CT18 + nTT2 (right) combinations.

uncertainty. This is due to the larger number of data points. The second largest contribution is from CMS13lj $m_{t\bar{t}}$ with approximately 21%, and ATL13had $y_{t\bar{t}}$ with 16%. The smallest contribution is from CMS13ll $y_{t\bar{t}}$, with fractional uncertainty of 13%.

These differences in the treatment of the ATL13lj measurements prompted us to identify the two optimal combinations CT18 + nTT1 and CT18 + nTT2, but overall they have mild impact in the global fits we have analyzed.

TABLE VI. Fractional contribution of leading eigenvectors (EV) from ePump optimization for each dataset in the CT18 + nTT1 combination. The second column shows the sum of fractional contributions from individual datasets.

	CT18 + nTT1 (nTT1-N_{pt} = 47)				
Leading EV No.	Leading EV (%)	ATL13had	CMS13ll	CMS13lj	ATL13lj
1	65.80	16.96	13.42	22.00	13.42
2	31.64	7.96	7.24	9.20	7.24
3	1.69	0.49	0.47	0.27	0.47
4	0.79	0.10	0.14	0.42	0.14
5	0.07	0.02	0.01	0.02	0.01

TABLE VII. Same as in Table VII, but for the CT18 + nTT2 combination.

	CT18 + nTT2 (nTT2-N_{pt} = 71)				
Leading EV No.	Leading EV (%)	ATL13had	CMS13ll	CMS13lj	ATL13lj
1	68.43	10.53	8.64	15.53	33.73
2	28.92	5.91	4.99	5.24	12.77
3	1.52	0.36	0.33	0.12	0.71
4	1.00	0.08	0.12	0.22	0.58
5	0.11	0.02	0.01	0.01	0.07

V. CONCLUSIONS

We presented a comprehensive study of the impact of recent high-precision LHC top-quark pair production measurements at $\sqrt{S} = 13$ TeV of collision energy on PDFs of the proton, in particular the gluon, in global analyses at NNLO in QCD. This extensive analysis of post-CT18 PDFs is relevant for the next release of CTEQ-TEA PDFs which will be challenged by the inclusion of high-precision forthcoming measurements at the LHC for a multitude of standard candle processes like top-quark pair production.

Besides the PDF impact of the 13 TeV $t\bar{t}$ differential cross section measurements from ATLAS and CMS, we studied their interplay with inclusive jet production measurements in global PDF fits.

Due to differences in the binning resolution of the $t\bar{t}$ 13 TeV ATLAS lepton + jet data, we identified two optimal combinations of measurements that maximize the information to constrain the gluon, and minimize conflict with the other datasets in the extended baseline.

Overall, the impact of these measurements in reducing the uncertainty in the gluon PDF is found to be mild. However, their role is important as they constrain the behavior of the gluon PDF at large momentum fraction x in a way that complements that of inclusive jet production data. In fact, the $t\bar{t}$ and inclusive jet production processes overlap in the Q-x plane, but their matrix elements and phase-space suppression are different, and constraints on the gluon and other PDFs are placed at different values of x.

We analyzed the impact of bin-by-bin statistical correlations whenever possible, as well as that of central-scale variations in the theory prediction for the 13 TeV top-quark data. The criterion according to which a particular scale is chosen, is based on the improvements produced in the quality-of-fit description (χ^2 description) in the global analysis. We observed that the $\mu_F = \mu_R = H_T/2$ choice of the central scale improves the description of the 13 TeV lepton + jet data at CMS with 137 fb^{-1} of integrated luminosity. These are the most precise data included in this work and place stronger constrains on PDFs as compared to the other measurements. Future analyses and extensions of this work will include measurements from ATLAS with similar or higher integrated luminosity as well as implications due to novel PDF parametrizations.

The global analyses performed in this work have been challenged by the interpretation of the correlated systematic uncertainties published by the ATLAS and CMS collaborations. The default treatment of correlated systematic errors in the CTEQ-framework is in terms of nuisance parameters. The top-quark pair production measurements from the CMS collaborations have been recently published in terms of the covariance matrix representation. We performed conversion between the covariance matrix and nuisance parameter representation using a similar strategy to that used in the CT18 study. This allowed us to obtain identical χ^2 values in both representations. However, this conversion is not unique. Detailed information on both the covariance and nuisance parameter representations for experimental errors is critical to fully exploit constraints from the data in global QCD analyses for PDF determinations, and is critical to perform a simultaneous determination of m_t, α_s, and the PDFs as well as their correlations in future analyses.

APPENDIX A: DETAILS OF THE THEORETICAL CALCULATIONS

In this section, we provide more details about the calculation of the theory predictions used in this work. We point out again that global QCD analyses to determine PDFs of the proton necessitate fast, precise, and accurate theory predictions to be confronted to the experimental data in the minimization procedure to reduce the CPU turn-around time. For this reason, frameworks like `fastNLO` and `APPLGrid` for the generation of fast tables are indispensable tools.

1. NNLO QCD prediction benchmarks

As discussed in Sec. II C, two independent theory calculations for the top-quark pair production differential cross section at NNLO in QCD have been used in this work. One is obtained with fast tables [180,181] based on the STRIPPER [57,58] subtraction method, that are produced at NNLO with `fastNLO`. The other one is obtained by using NNLO/NLO K-factors, where fast tables for the NLO cross section in the denominator are generated with `APPLGrid` [179] using the MCFM [35,184] program, while the NNLO cross section in the numerator is computed with the MATRIX program [60–62], based on the q_T-subtraction method [182]. Currently, there are no fast tables available for MATRIX.

Comparisons between these two calculations are shown in Figs. 22 and 23, where the theory is compared to the $m_{t\bar{t}}$, $y_{t\bar{t}}$, y_t, and $p_{T,t}$ distributions measured at CMS at 13 TeV in the dilepton channel [26], and to the $m_{t\bar{t}}$, $y_{t\bar{t}}$, p_{T,t_1}, and $H_T^{t\bar{t}}$ distributions measured at ATLAS at 13 TeV in the all-hadronic channel [24], respectively.

For this case study, we use CT18 NNLO PDFs [130], the central scale is set to $\mu_F = \mu_R = H_T/4$, and the top-quark mass is $m_t^{(\text{pole})} = 172.5$ GeV. These are our default parameters. In general, we find agreement between the two calculations within 1% accuracy, which is sufficient for all of the analyses in this work.

In the upper insets of each panel of Fig. 22, theory predictions obtained with `fastNLO` at LO, NLO, and NNLO perturbative orders for the absolute differential distributions are compared to the CMS measurements. Statistical and systematical uncertainties are shown separately using error bars with different colors. In the lower insets, we show NNLO/NLO K-factors for the two calculations. We note that for the y_t, $y_{t\bar{t}}$ and $p_{T,t}$ distributions the overall NNLO correction is about 7%, except for the $m_{t\bar{t}}$ where the K-factor has larger variation ranging from 5% to 12%. We observe agreement between MATRIX and `fastNLO` at the percent level for all distributions, though they agree better in the $y_{t\bar{t}}$ than in $m_{t\bar{t}}$ distributions.

In the upper insets of each panel in Fig. 23, the theory predictions that are compared to the ATLAS measurements are obtained with MATRIX using default parameters. Statistical and systematical uncertainties are displayed as in Fig. 22. In the lower insets, we show the MATRIX NNLO/NLO K-factor as well as the ratio between the AppGrid theory prediction from MCFM and MATRIX both computed at NLO. These NLO calculations agree within 1% accuracy. In addition, we note that the leading transverse momentum p_{T,t_1} and $H_T^{t\bar{t}}$ distributions are affected by large QCD perturbative corrections as their K-factors produce large variations.

2. NLO electroweak corrections

We explore the impact of the electroweak (EW) corrections on the $t\bar{t}$ differential distributions at 13 TeV and discuss their impact in our global PDF analyses.

EW corrections are computed as K-factors using the multiplicative scheme as described in Ref. [113], which are available in the repository [181]. These EW corrections include contributions of order $\mathcal{O}(\alpha_s^2\alpha)$ as well as subleading ones, of order $\mathcal{O}(\alpha_s\alpha^2)$ and $\mathcal{O}(\alpha^3)$. Meanwhile, EW corrections are also incorporated in MadGraph5_aMC@NLO [111,211] which performs differential cross section calculations in an automated fashion, up to NLO in both couplings. Recently, MadGraph5_aMC@NLO has been interfaced with the PineAPPL library [114] to obtain fast interpolation grids which include EW corrections up to $\mathcal{O}(\alpha_s^2\alpha)$. In addition, EW corrections for top-quark pair production at hadron colliders are also included in MCFM [115].

In Fig. 24, we illustrate EW corrections from Czakon *et al.* [113,181], PineAPPL [111,114], and MCFM [115] which are defined as multiplicative K-factors $K_{\text{EW}} = \text{QCD} \times \text{EW}/\text{QCD}$. The 13 TeV $t\bar{t}$ theory predictions for the differential distributions are calculated using the CT18NNLO PDFs, and use the same bins shared by the
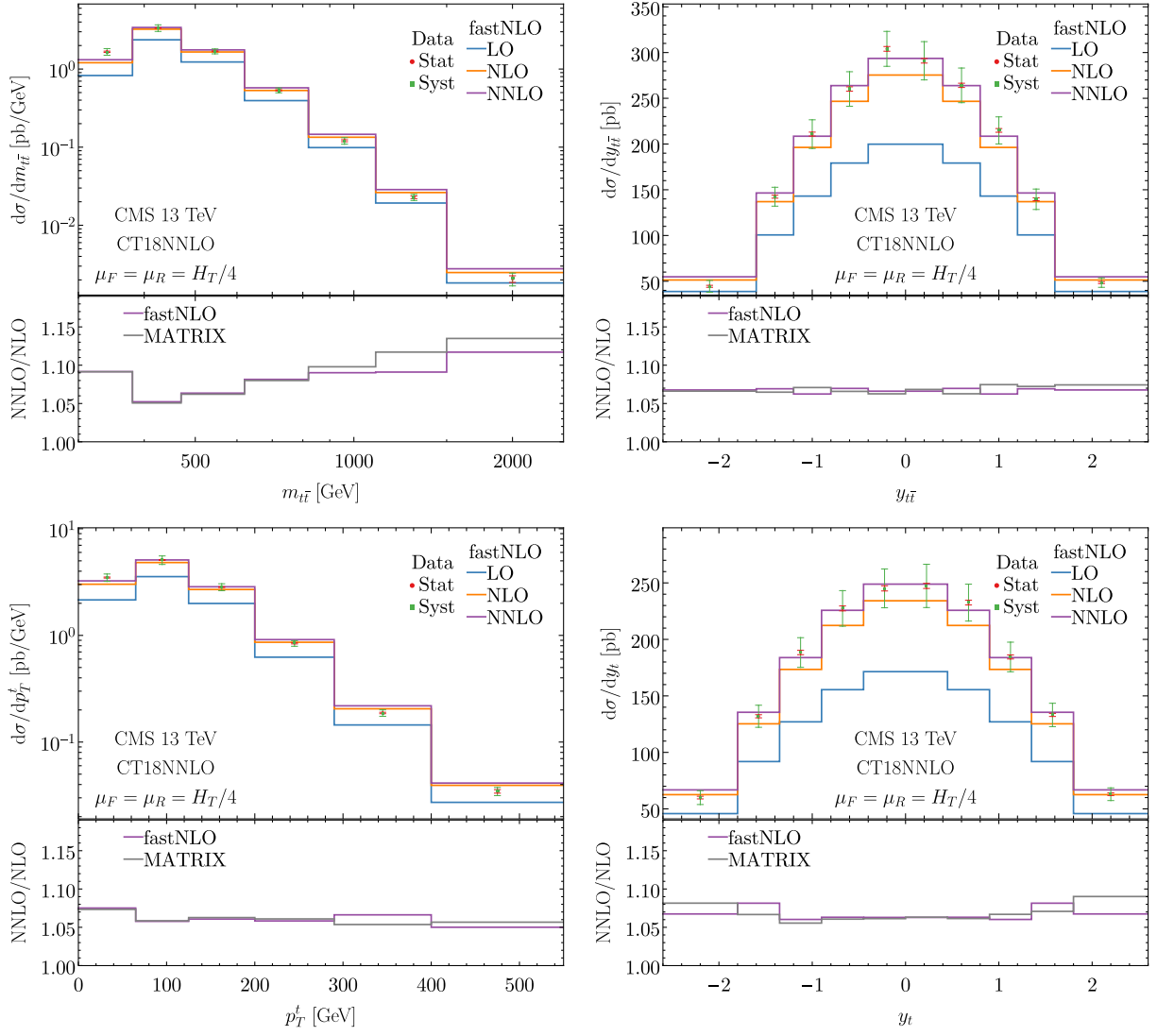
FIG. 22. Theoretical predictions for the $m_{t\bar{t}}$, $y_{t\bar{t}}$, y_t, and $p_{T,t}$ distributions compared to the 13 TeV CMS measurements in the dilepton channel [26]. The CT18 NNLO PDFs [130], scale choice $\mu_F = \mu_R = H_T/4$, and pole mass $m_t^{(\text{pole})} = 172.5$ GeV are selected here as default parameters.

ATLAS distributions in the lepton + jets [23] channel and the CMS distributions in the dilepton channel [26].

While the two calculations produce almost identical shapes in the various distributions, differences within 1% or smaller are found, mainly driven by the subleading EW corrections. In addition, we observe that the rapidity distributions y_t and $y_{t\bar{t}}$ appear to be more stable against EW corrections as compared to $m_{t\bar{t}}$ and $p_{T,t}$, with most of the impact affecting bins at large rapidity.

EW corrections calculated using multiplicative EW K-factors are explored in our global PDF analyses. We essentially observe no sizable impact on either the central-value PDFs or their uncertainties. The impact of EW corrections in our PDF analyses is negligible given the current size of the experimental errors and other theoretical uncertainties affecting the calculation in the global fit.

APPENDIX B: CMSlj13 $m_{t\bar{t}}$ AND $y_{t\bar{t}}$ WITH CT18 + nTT1 AND CT18 + nTT2 PDFs

In this section, we illustrate a data vs theory comparison for the $m_{t\bar{t}}$ and $y_{t\bar{t}}$ distributions at CMS 13 TeV in the lepton + jet channel (CMS13lj) which we show in Fig. 25. Here, the theory predictions are obtained with the CT18, CT18 + nTT1 and CT18 + nTT2 PDFs after the global fit. In both cases we observe a mild improvement when using the updated PDFs, as compared to Fig. 1 As discussed in this work, cf. Table IV, the χ^2/N_{pt} for the $m_{t\bar{t}}$ distribution in the CT18 + nTT1 and CT18 + nTT2 fits goes from 1.4 down to 1.1 for the scale chosen to be $H_T/2$. This is mostly due to an improvement in bin 8, 10, 12, and 13. However, the theory for both distributions continues to overestimate the data in the higher bins. As discussed in the main text,
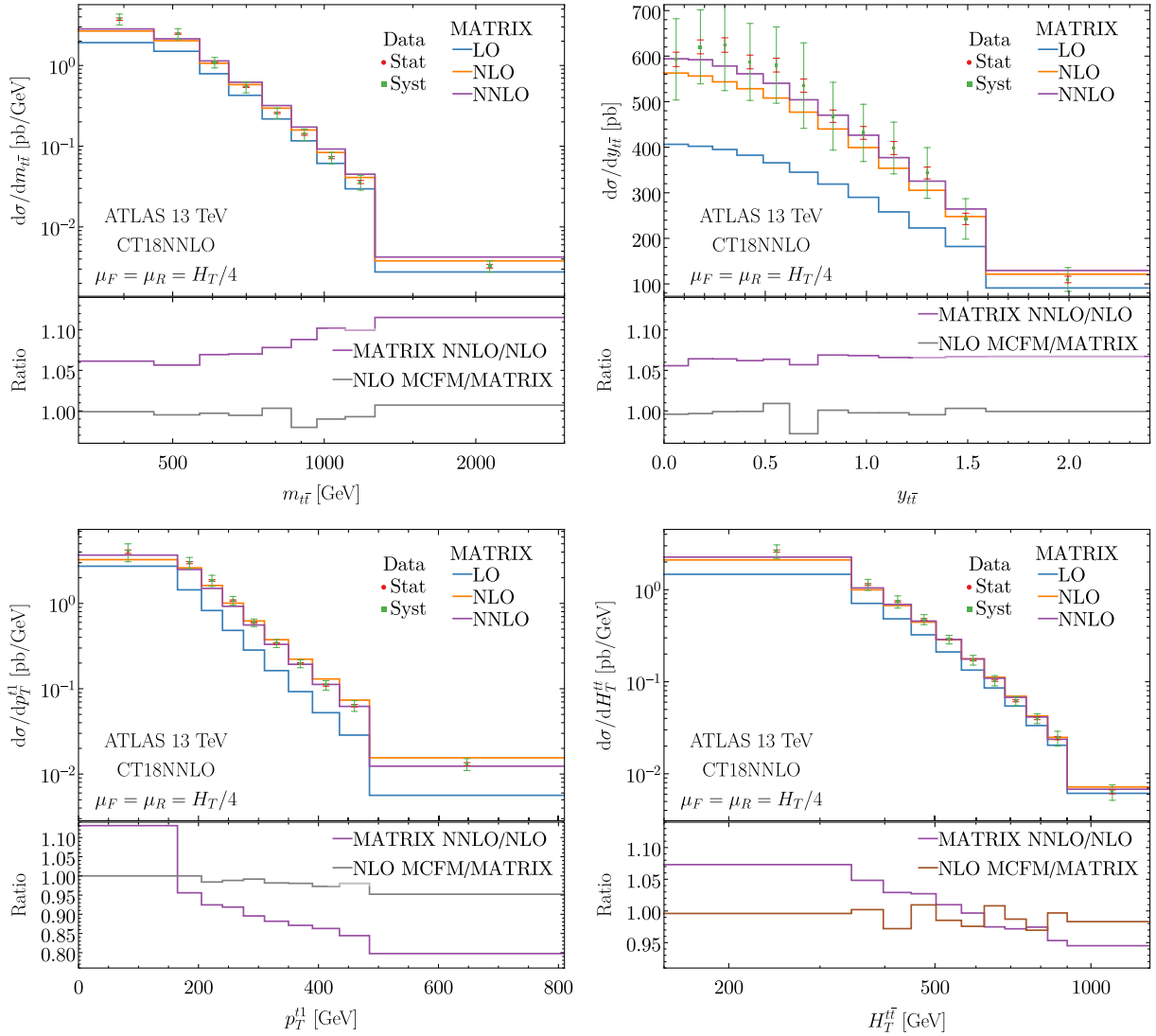
FIG. 23.   Theoretical predictions for the $m_{t\bar{t}}$, $y_{t\bar{t}}$, p_{T,t_1}, and $H_T^{t\bar{t}}$ distributions compared to the 13 TeV ATLAS measurements in the all-hadronic channel [24]. Here, default parameters are used as in Fig. 22.

the $y_{t\bar{t}}$ distribution cannot produce a good fit and therefore is not included in this analysis.

APPENDIX C: ePump VS GLOBAL FIT: GLUON PDF FROM THE CT18 + CMSlj13-$y_{t\bar{t}}$ BASELINE FIT

As pointed out in the original papers [146,147], the ePump profiling method may not be reliable when the updated central PDF deviates significantly from the baseline. In this work, we have only used ePump as an auxiliary tool for the selection of the optimal combinations of the datasets to be analyzed via global fits. Therefore, in this work we do not report the full updated results of ePump as compared to a true global fit. The global fit is the main result of our PDF global analysis. With that said, we have checked that the ePump

updated results agree well with the global fit for the optimal combinations of the datasets.

As shown in Table III, the CT18 + CMSlj13-$y_{t\bar{t}}$ data cannot be fit well either by ePump-updating or a global fit, which yields a large value of χ^2/N_{pt} for describing the CMSlj13-$y_{t\bar{t}}$ data. Hence, the CMSlj13-$y_{t\bar{t}}$ dataset is not included in the optimal combinations nTT1 and nTT2, as introduced in the main text. Here, we update the CT18 PDFs by including this data to the original CT18 dataset, as an example, to compare the post-CT18 gluon PDF obtained from ePump-updating to that from a global fit. In Fig. 26, we show the impact of the CMSlj13-$y_{t\bar{t}}$ measurements on both the central value and error band, and we also show the nominal CT18 results. The ePump result for the gluon central value in Fig. 26 (left) is very close to the result from the global fit and the same is true for the error band in
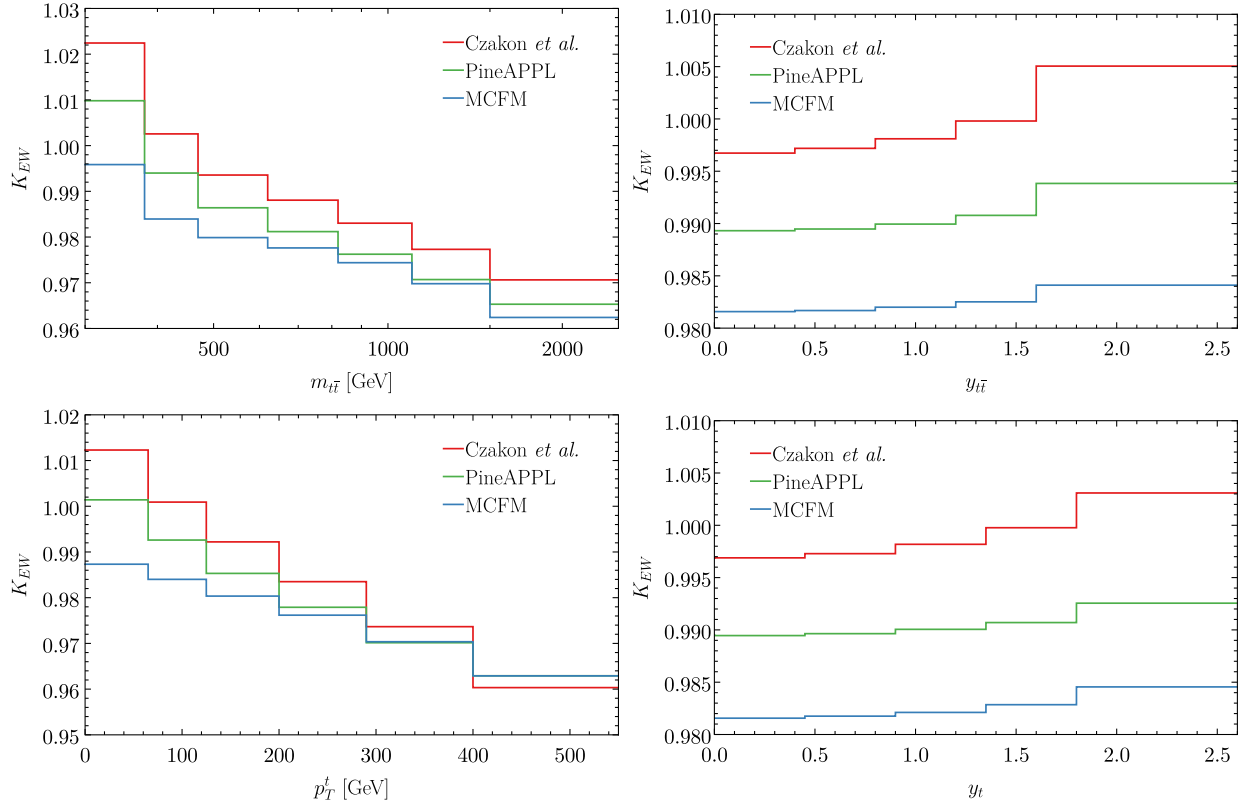
FIG. 24.   Comparison of the EW corrections from the Czakon *et al.* [113], PineAPPL [114], and MCFM [115] for the $t\bar{t}$ production at the LHC 13 TeV measured by the ATLAS collaboration in the lepton + jets channel [23] and CMS in the dilepton one [26].
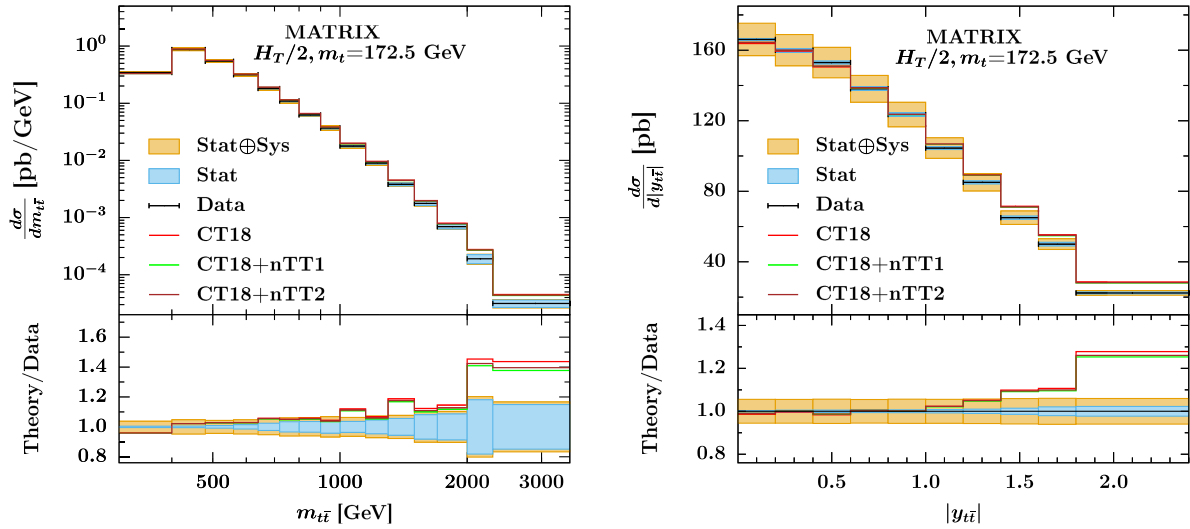

FIG. 25.   Theory predictions for the $m_{t\bar{t}}$ (left) and $|y_{t\bar{t}}|$ (right) distributions obtained with CT18 PDF, and with CT18 + nTT1 and CT18 + nTT2 PDFs after the global fit. The theory is compared to the CMS13lj measurements.

Fig. 26 (right). Namely, even for this dataset, with its χ^2/N_{pt} value much larger than 1 the ePump update is still found to be in good agreement with the result from a global fit. A similar (or better) level of agreement is found for the other datasets in the new nTT1 and nTT2 combinations.

APPENDIX D: TREATMENT OF EXPERIMENTAL UNCERTAINTIES

In global QCD analyses, the agreement between theory and experiment can be investigated by quantifying the agreement between individual data points and shifted
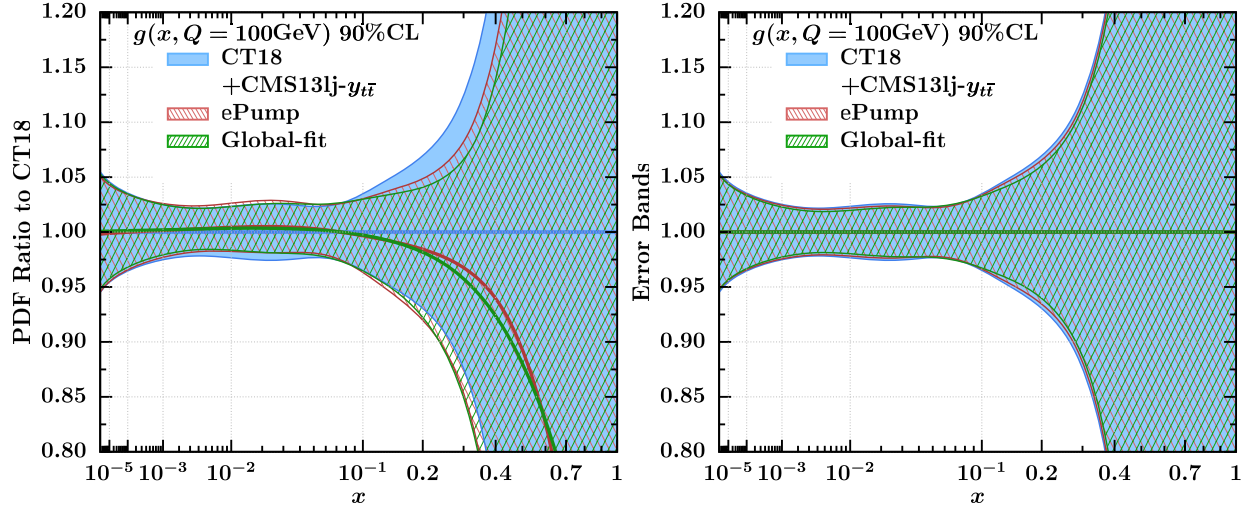
FIG. 26.   Post-CT18 gluon PDF and its uncertainty in the CT18 + CMS13lj-$y_{t\bar{t}}$ fit. Bands with different hatching represent the ePump update(red), and global fit results(green). The blue band is CT18 at 90% CL.

residuals (which will be defined later) with the corresponding nuisance parameters. Under the assumption that both these quantities are distributed according to their normal distribution, which we refer to as $\mathcal{N}(0,1)$, this can be used to test the goodness-of-fit (see Refs. [212] and references therein for more details). The experimental collaborations at the LHC often present experimental uncertainties using the covariance matrix representation (see the discussion in Sec. II B). Therefore, to examine the normal distribution of shifted residuals a conversion to the nuisance parameters representation is required. This conversion is generally not unique, especially when the statistical and uncorrelated systematical errors are not fully specified, and when correlated systematic uncertainties and their sources are not explicitly known. In this case, a question arises about finding the optimal strategy to factor the covariance matrix and perform the conversion to nuisance parameters, and select the optimal number N_{corr} of correlated sources that captures most of the features of the true correlated sources of uncertainty and does not introduce artificial fluctuations in the χ^2 calculation (see the discussion in Ref. [213]).

In this work, when the information on the experimental uncertainties is not fully provided with the measurements used in our analysis (e.g., the CMS13ll, CMS13lj measurements), we express the covariance matrix in terms of nuisance parameters representation by using a version the $\Sigma + K$ decomposition, which is an iterative procedure introduced in the CT18 study [130] to obtain the nuisance parameter representation from the covariance matrix. This allows us to numerically match the χ^2 in the two representations and obtain identical values. An extended discussion of this problem, in which we consider alternative methods of performing the conversion from the covariance matrix to the nuisance parameter representation, and study the problem of finding an optimal representation for the independent experimental errors that limits artificial fluctuations in the calculation of the χ^2 function, will be addressed in a separate work. Independent discussions on the treatment of the experimental uncertainty correlations in global PDF analyses can also be found in Refs. [212,213].

[1] F. Abe *et al.* (CDF Collaboration), Observation of top quark production in $\bar{p}p$ collisions, Phys. Rev. Lett. **74**, 2626 (1995).

[2] S. Abachi *et al.* (D0 Collaboration), Observation of the top quark, Phys. Rev. Lett. **74**, 2632 (1995).

[3] R. L. Workman *et al.* (Particle Data Group), Review of particle physics, Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[4] G. Aad *et al.* (ATLAS Collaboration), Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716**, 1 (2012).

[5] S. Chatrchyan *et al.* (CMS Collaboration), Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. B **716**, 30 (2012).

[6] V. Branchina, F. Contino, and A. Pilaftsis, Protecting the stability of the electroweak vacuum from Planck-scale gravitational effects, Phys. Rev. D **98**, 075001 (2018).

[7] E. Bentivegna, V. Branchina, F. Contino, and D. Zappalà, Impact of new physics on the EW vacuum stability in a curved spacetime background, J. High Energy Phys. 12 (2017) 100.

[8] S. Alekhin, A. Djouadi, and S. Moch, The top quark and Higgs boson masses and the stability of the electroweak vacuum, Phys. Lett. B **716**, 214 (2012).

[9] F. Bezrukov, M. Y. Kalmykov, B. A. Kniehl, and M. Shaposhnikov, Higgs boson mass and new physics, J. High Energy Phys. 10 (2012) 140.

[10] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, Higgs mass and vacuum stability in the standard model at NNLO, J. High Energy Phys. 08 (2012) 098.

[11] J. Ellis, J. R. Espinosa, G. F. Giudice, A. Hoecker, and A. Riotto, The probable fate of the Standard Model, Phys. Lett. B **679**, 369 (2009).

[12] G. Isidori, G. Ridolfi, and A. Strumia, On the metastability of the standard model vacuum, Nucl. Phys. **B609**, 387 (2001).

[13] S. R. Coleman, The fate of the false vacuum. 1. Semi-classical theory, Phys. Rev. D **15**, 2929 (1977); **16**, 1248 (E) (1977).

[14] C. G. Callan, Jr. and S. R. Coleman, The fate of the false vacuum. 2. First quantum corrections, Phys. Rev. D **16**, 1762 (1977).

[15] R. Frederix and F. Maltoni, Top pair invariant mass distribution: A window on new physics, J. High Energy Phys. 01 (2009) 047.

[16] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, O new physics, where art thou? A global search in the top sector, J. High Energy Phys. 02 (2020) 131.

[17] G. Aad et al. (ATLAS Collaboration), Measurements of top quark pair relative differential cross-sections with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV, Eur. Phys. J. C **73**, 2261 (2013).

[18] G. Aad et al. (ATLAS Collaboration), Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, Phys. Rev. D **90**, 072004 (2014).

[19] G. Aad et al. (ATLAS Collaboration), Measurements of top-quark pair differential cross-sections in the lepton + jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector, Eur. Phys. J. C **76**, 538 (2016).

[20] S. Chatrchyan et al. (CMS Collaboration), Measurement of differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV, Eur. Phys. J. C **73**, 2339 (2013).

[21] S. Chatrchyan et al. (CMS Collaboration), Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 02 (2014) 024; 02 (2014) 102(E).

[22] A. M. Sirunyan et al. (CMS Collaboration), Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions, Eur. Phys. J. C **77**, 459 (2017).

[23] G. Aad et al. (ATLAS Collaboration), Measurements of top-quark pair differential and double-differential cross-sections in the ℓ + jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector, Eur. Phys. J. C **79**, 1028 (2019); **80**, 1092(E) (2020).

[24] G. Aad et al. (ATLAS Collaboration), Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector, J. High Energy Phys. 01 (2021) 033.

[25] G. Aad et al. (ATLAS Collaboration), Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector, J. High Energy Phys. 07 (2023) 141.

[26] A. M. Sirunyan et al. (CMS Collaboration), Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons, J. High Energy Phys. 02 (2019) 149.

[27] A. M. Sirunyan et al. (CMS Collaboration), Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton + jets events from pp collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. D **97**, 112003 (2018).

[28] A. Tumasyan et al. (CMS Collaboration), Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton + jets events from proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. D **104**, 092013 (2021).

[29] P. Nason, S. Dawson, and R. K. Ellis, The total cross-section for the production of heavy quarks in hadronic collisions, Nucl. Phys. **B303**, 607 (1988).

[30] P. Nason, S. Dawson, and R. K. Ellis, The one particle inclusive differential cross-section for heavy quark production in hadronic collisions, Nucl. Phys. **B327**, 49 (1989); **B335**, 260(E) (1990).

[31] W. Beenakker, H. Kuijf, W. L. van Neerven, and J. Smith, QCD corrections to heavy quark production in $p\bar{p}$ collisions, Phys. Rev. D **40**, 54 (1989).

[32] R. Meng, G. A. Schuler, J. Smith, and W. L. van Neerven, Simple formulae for the order α_s^3 QCD corrections to the reaction $p\bar{p} \rightarrow Q\bar{Q}$ X, Nucl. Phys. **B339**, 325 (1990).

[33] W. Beenakker, W. L. van Neerven, R. Meng, G. A. Schuler, and J. Smith, QCD corrections to heavy quark production in hadron hadron collisions, Nucl. Phys. **B351**, 507 (1991).

[34] M. L. Mangano, P. Nason, and G. Ridolfi, Heavy quark correlations in hadron collisions at next-to-leading order, Nucl. Phys. **B373**, 295 (1992).

[35] J. M. Campbell, R. K. Ellis, and W. T. Giele, A multi-threaded version of MCFM, Eur. Phys. J. C **75**, 246 (2015).

[36] S. Frixione, P. Nason, and B. R. Webber, Matching NLO QCD and parton showers in heavy flavor production, J. High Energy Phys. 08 (2003) 007.

[37] S. Frixione, P. Nason, and G. Ridolfi, A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction, J. High Energy Phys. 09 (2007) 126.

[38] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, MadGraph/MadEvent v4: The new web generation, J. High Energy Phys. 09 (2007) 028.

[39] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer, Automation of next-to-leading order computations in QCD: The FKS subtraction, J. High Energy Phys. 10 (2009) 003.

[40] A. G. Grozin, Heavy quark effective theory, Springer Tracts Mod. Phys. **201**, 1 (2004).

[41] M. Czakon, A. Mitov, and S. Moch, Heavy-quark production in gluon fusion at two loops in QCD, Nucl. Phys. **B798**, 210 (2008).

[42] M. Czakon, A. Mitov, and S. Moch, Heavy-quark production in massless quark scattering at two loops in QCD, Phys. Lett. B **651**, 147 (2007).

[43] A. Mitov and S. Moch, The singular behavior of massive QCD amplitudes, J. High Energy Phys. 05 (2007) 001.

[44] A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, Two-loop divergences of scattering amplitudes with massive partons, Phys. Rev. Lett. **103**, 201601 (2009).

[45] A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, Two-loop divergences of massive scattering amplitudes in non-Abelian gauge theories, J. High Energy Phys. 11 (2009) 062.

[46] N. Kidonakis, Two-loop soft anomalous dimensions and NNLL resummation for heavy quark production, Phys. Rev. Lett. **102**, 232003 (2009).

[47] I. Bierenbaum, M. Czakon, and A. Mitov, The singular behavior of one-loop massive QCD amplitudes with one external soft gluon, Nucl. Phys. **B856**, 228 (2012).

[48] P. Bärnreuther, M. Czakon, and P. Fiedler, Virtual amplitudes and threshold behaviour of hadronic top-quark pair-production cross sections, J. High Energy Phys. 02 (2014) 078.

[49] A. G. Grozin, R. N. Lee, and A. F. Pikelner, Four-loop QCD cusp anomalous dimension at small angle, J. High Energy Phys. 11 (2022) 094.

[50] N. Kidonakis, Four-loop massive cusp anomalous dimension in QCD: A calculation from asymptotics, Phys. Rev. D **107**, 054006 (2023).

[51] M. Czakon, P. Fiedler, and A. Mitov, Total top-quark pair-production cross section at hadron colliders through $O(\alpha_S^4)$, Phys. Rev. Lett. **110**, 252004 (2013).

[52] M. Czakon and A. Mitov, NNLO corrections to top pair production at hadron colliders: The quark-gluon reaction, J. High Energy Phys. 01 (2013) 080.

[53] M. Czakon and A. Mitov, NNLO corrections to top-pair production at hadron colliders: The all-fermionic scattering channels, J. High Energy Phys. 12 (2012) 054.

[54] P. Bärnreuther, M. Czakon, and A. Mitov, Percent level precision physics at the Tevatron: First genuine NNLO QCD corrections to $q\bar{q} \to t\bar{t} + X$, Phys. Rev. Lett. **109**, 132001 (2012).

[55] M. Czakon, D. Heymes, and A. Mitov, High-precision differential predictions for top-quark pairs at the LHC, Phys. Rev. Lett. **116**, 082003 (2016).

[56] M. Czakon, P. Fiedler, D. Heymes, and A. Mitov, NNLO QCD predictions for fully-differential top-quark pair production at the Tevatron, J. High Energy Phys. 05 (2016) 034.

[57] M. Czakon, A novel subtraction scheme for double-real radiation at NNLO, Phys. Lett. B **693**, 259 (2010).

[58] M. Czakon, Double-real radiation in hadronic top quark pair production as a proof of a certain concept, Nucl. Phys. **B849**, 250 (2011).

[59] M. Czakon and D. Heymes, Four-dimensional formulation of the sector-improved residue subtraction scheme, Nucl. Phys. **B890**, 152 (2014).

[60] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, J. Mazzitelli, and H. Sargsyan, Top-quark pair hadroproduction at next-to-next-to-leading order in QCD, Phys. Rev. D **99**, 051501 (2019).

[61] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, and J. Mazzitelli, Top-quark pair production at the LHC: Fully differential QCD predictions at NNLO, J. High Energy Phys. 07 (2019) 100.

[62] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, and J. Mazzitelli, Top-quark pair hadroproduction at NNLO: Differential predictions with the $\overline{\text{MS}}$ mass, J. High Energy Phys. 08 (2020) 027.

[63] M. Grazzini, S. Kallweit, and M. Wiesemann, Fully differential NNLO computations with MATRIX, Eur. Phys. J. C **78**, 537 (2018).

[64] A. Behring, M. Czakon, A. Mitov, A. S. Papanastasiou, and R. Poncelet, Higher order corrections to spin correlations in top quark pair production at the LHC, Phys. Rev. Lett. **123**, 082001 (2019).

[65] M. Czakon, A. Mitov, and R. Poncelet, NNLO QCD corrections to leptonic observables in top-quark pair production and decay, J. High Energy Phys. 05 (2021) 212.

[66] G. F. Sterman, Summation of large corrections to short distance hadronic cross-sections, Nucl. Phys. **B281**, 310 (1987).

[67] S. Catani and L. Trentadue, Resummation of the QCD perturbative series for hard processes, Nucl. Phys. **B327**, 323 (1989).

[68] N. Kidonakis and G. F. Sterman, Resummation for QCD hard scattering, Nucl. Phys. **B505**, 321 (1997).

[69] E. Laenen, G. Oderda, and G. F. Sterman, Resummation of threshold corrections for single particle inclusive cross-sections, Phys. Lett. B **438**, 173 (1998).

[70] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, NLL resummation of the heavy quark hadroproduction cross-section, Nucl. Phys. **B529**, 424 (1998); **B803**, 234(E) (2008).

[71] M. Czakon, A. Mitov, and G. F. Sterman, Threshold resummation for top-pair hadroproduction to next-to-next-to-leading log, Phys. Rev. D **80**, 074017 (2009).

[72] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, HATHOR: Hadronic top and heavy quarks cross section calculator, Comput. Phys. Commun. **182**, 1034 (2011).

[73] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation, Phys. Lett. B **710**, 612 (2012).

[74] M. Czakon and A. Mitov, Top++: A program for the calculation of the top-pair cross-section at hadron colliders, Comput. Phys. Commun. **185**, 2930 (2014).

[75] N. Kidonakis, High order corrections and subleading logarithms for top quark production, Phys. Rev. D **64**, 014009 (2001).

[76] N. Kidonakis, E. Laenen, S. Moch, and R. Vogt, Sudakov resummation and finite order expansions of heavy quark hadroproduction cross-sections, Phys. Rev. D **64,** 114001 (2001).

[77] N. Kidonakis and R. Vogt, Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction, Phys. Rev. D **68,** 114014 (2003).

[78] N. Kidonakis and R. Vogt, The theoretical top quark cross section at the Tevatron and the LHC, Phys. Rev. D **78,** 074005 (2008).

[79] N. Kidonakis, Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution, Phys. Rev. D **82,** 114030 (2010).

[80] N. Kidonakis, NNNLO soft-gluon corrections for the top-quark p_T and rapidity distributions, Phys. Rev. D **91,** 031501 (2015).

[81] N. Kidonakis, NNNLO soft-gluon corrections for the top-antitop pair production cross section, Phys. Rev. D **90,** 014006 (2014).

[82] N. Kidonakis, Top-quark double-differential distributions at approximate N^{3}LO, Phys. Rev. D **101,** 074006 (2020).

[83] N. Kidonakis, M. Guzzi, and A. Tonero, Top-quark cross sections and distributions at aN^{3}LO, Phys. Rev. D **108,** 054012 (2023).

[84] M. Neubert, Heavy quark symmetry, Phys. Rep. **245,** 259 (1994).

[85] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, An effective field theory for collinear and soft gluons: Heavy to light decays, Phys. Rev. D **63,** 114020 (2001).

[86] C. W. Bauer, D. Pirjol, and I. W. Stewart, Soft collinear factorization in effective field theory, Phys. Rev. D **65,** 054022 (2002).

[87] M. Beneke, A. P. Chapovsky, M. Diehl, and T. Feldmann, Soft collinear effective theory and heavy to light currents beyond leading power, Nucl. Phys. **B643,** 431 (2002).

[88] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, Threshold expansion at order α_s^4 for the $t\bar{t}$ invariant mass distribution at hadron colliders, Phys. Lett. B **687,** 331 (2010).

[89] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, Renormalization-group improved predictions for top-quark pair production at hadron colliders, J. High Energy Phys. 09 (2010) 097.

[90] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L.-L. Yang, RG-improved single-particle inclusive cross sections and forward-backward asymmetry in $t\bar{t}$ production at hadron colliders, J. High Energy Phys. 09 (2011) 070.

[91] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, Precision predictions for the $t\bar{t}$ production cross section at hadron colliders, Phys. Lett. B **703,** 135 (2011).

[92] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, Hadronic top-quark pair production with NNLL threshold resummation, Nucl. Phys. **B855,** 695 (2012).

[93] M. Beneke, P. Falgari, S. Klein, J. Piclum, C. Schwinn, M. Ubiali, and F. Yan, Inclusive top-pair production phenomenology with TOPIXS, J. High Energy Phys. 07 (2012) 194.

[94] A. Ferroglia, B. D. Pecjak, L. L. Yang, B. D. Pecjak, and L. L. Yang, The NNLO soft function for the pair invariant mass distribution of boosted top quarks, J. High Energy Phys. 10 (2012) 180.

[95] A. Ferroglia, B. D. Pecjak, and L. L. Yang, Soft-gluon resummation for boosted top-quark production at hadron colliders, Phys. Rev. D **86,** 034010 (2012).

[96] A. Ferroglia, S. Marzani, B. D. Pecjak, and L. L. Yang, Boosted top production: Factorization and resummation for single-particle inclusive distributions, J. High Energy Phys. 01 (2014) 028.

[97] W. Beenakker, A. Denner, W. Hollik, R. Mertig, T. Sack, and D. Wackeroth, Electroweak one loop contributions to top pair production in hadron colliders, Nucl. Phys. **B411,** 343 (1994).

[98] J. H. Kuhn, A. Scharf, and P. Uwer, Electroweak corrections to top-quark pair production in quark-antiquark annihilation, Eur. Phys. J. C **45,** 139 (2006).

[99] W. Bernreuther, M. Fücker, and Z. G. Si, Mixed QCD and weak corrections to top quark pair production at hadron colliders, Phys. Lett. B **633,** 54 (2006); **644,** 386(E) (2007).

[100] J. H. Kuhn, A. Scharf, and P. Uwer, Electroweak effects in top-quark pair production at hadron colliders, Eur. Phys. J. C **51,** 37 (2007).

[101] W. Bernreuther, M. Fuecker, and Z.-G. Si, Weak interaction corrections to hadronic top quark pair production, Phys. Rev. D **74,** 113005 (2006).

[102] W. Bernreuther, M. Fucker, and Z.-G. Si, Weak interaction corrections to hadronic top quark pair production: Contributions from quark-gluon and $b\bar{b}$ induced reactions, Phys. Rev. D **78,** 017503 (2008).

[103] A. V. Manohar and M. Trott, Electroweak Sudakov corrections and the top quark forward-backward asymmetry, Phys. Lett. B **711,** 313 (2012).

[104] J. H. Kühn, A. Scharf, and P. Uwer, Weak interactions in top-quark pair production at hadron colliders: An update, Phys. Rev. D **91,** 014020 (2015).

[105] J. M. Campbell, D. Wackeroth, and J. Zhou, Electroweak corrections at the LHC with MCFM, Proc. Sci. DIS2015 (**2015**) 130 [arXiv:1508.06247].

[106] W. Hollik and M. Kollar, NLO QED contributions to top-pair production at hadron collider, Phys. Rev. D **77,** 014008 (2008).

[107] W. Bernreuther and Z.-G. Si, Distributions and correlations for top quark pair production and decay at the Tevatron and LHC, Nucl. Phys. **B837,** 90 (2010).

[108] W. Hollik and D. Pagani, The electroweak contribution to the top quark forward-backward asymmetry at the Tevatron, Phys. Rev. D **84,** 093003 (2011).

[109] J. H. Kuhn and G. Rodrigo, Charge asymmetries of top quarks at hadron colliders revisited, J. High Energy Phys. 01 (2012) 063.

[110] W. Bernreuther and Z.-G. Si, Top quark and leptonic charge asymmetries for the Tevatron and LHC, Phys. Rev. D **86,** 034026 (2012).

[111] D. Pagani, I. Tsinikos, and M. Zaro, The impact of the photon PDF and electroweak corrections on $t\bar{t}$ distributions, Eur. Phys. J. C **76,** 479 (2016).

[112] A. Denner and M. Pellen, NLO electroweak corrections to off-shell top-antitop production with leptonic decays at the LHC, J. High Energy Phys. 08 (2016) 155.

[113] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos, and M. Zaro, Top-pair production at the LHC through NNLO QCD and NLO EW, J. High Energy Phys. 10 (2017) 186.

[114] S. Carrazza, E. R. Nocera, C. Schwan, and M. Zaro, PineAPPL: Combining EW and QCD corrections for fast evaluation of LHC processes, J. High Energy Phys. 12 (2020) 108.

[115] J. M. Campbell, D. Wackeroth, and J. Zhou, Study of weak corrections to Drell-Yan, top-quark pair, and dijet production at high energies with MCFM, Phys. Rev. D **94**, 093009 (2016).

[116] S. Catani, M. Grazzini, and H. Sargsyan, Transverse-momentum resummation for top-quark pair production at the LHC, J. High Energy Phys. 11 (2018) 061.

[117] R. Angeles-Martinez, M. Czakon, and S. Sapeta, NNLO soft function for top quark pair production at small transverse momentum, J. High Energy Phys. 10 (2018) 201.

[118] R. Bonciani, S. Catani, M. Grazzini, H. Sargsyan, and A. Torre, The q_T subtraction method for top quark production at hadron colliders, Eur. Phys. J. C **75**, 581 (2015).

[119] S. Catani, M. Grazzini, and A. Torre, Transverse-momentum resummation for heavy-quark hadroproduction, Nucl. Phys. **B890**, 518 (2014).

[120] H. T. Li, C. S. Li, D. Y. Shao, L. L. Yang, and H. X. Zhu, Top quark pair production at small transverse momentum in hadronic collisions, Phys. Rev. D **88**, 074004 (2013).

[121] H. X. Zhu, C. S. Li, H. T. Li, D. Y. Shao, and L. L. Yang, Transverse-momentum resummation for top-quark pairs at hadron colliders, Phys. Rev. Lett. **110**, 082001 (2013).

[122] W.-L. Ju and M. Schönherr, Projected transverse momentum resummation in top-antitop pair production at LHC, J. High Energy Phys. 02 (2023) 075.

[123] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann, and G. Zanderighi, Next-to-next-to-leading order event generation for top-quark pair production, Phys. Rev. Lett. **127**, 062001 (2021).

[124] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann, and G. Zanderighi, Top-pair production at the LHC with MINNLO$_{PS}$, J. High Energy Phys. 04 (2022) 079.

[125] S. Alioli, A. Broggio, and M. A. Lim, Zero-jettiness resummation for top-quark pair production at the LHC, J. High Energy Phys. 01 (2022) 066.

[126] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, J. High Energy Phys. 11 (2004) 040.

[127] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: The POW-HEG method, J. High Energy Phys. 11 (2007) 070.

[128] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: The POWHEG BOX, J. High Energy Phys. 06 (2010) 043.

[129] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, Parton distribution functions, α_s, and heavy-quark masses for LHC Run II, Phys. Rev. D **96**, 014011 (2017).

[130] T.-J. Hou *et al.*, New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, Phys. Rev. D **103**, 014013 (2021).

[131] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs, Eur. Phys. J. C **81**, 341 (2021).

[132] J. McGowan, T. Cridge, L. A. Harland-Lang, and R. S. Thorne, Approximate N³LO parton distribution functions with theoretical uncertainties: MSHT20aN³LO PDFs, Eur. Phys. J. C **83**, 185 (2023); **83**, 302(E) (2023).

[133] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).

[134] R. D. Ball *et al.* (NNPDF Collaboration), The path to proton structure at 1% accuracy, Eur. Phys. J. C **82**, 428 (2022).

[135] T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, An investigation of the α_S and heavy quark mass dependence in the MSHT20 global PDF analysis, Eur. Phys. J. C **81**, 744 (2021).

[136] A. M. Cooper-Sarkar, M. Czakon, M. A. Lim, A. Mitov, and A. S. Papanastasiou, Simultaneous extraction of α_s and m_t from LHC $t\bar{t}$ differential distributions, arXiv:2010.04171.

[137] T. Cridge and M. A. Lim, Constraining the top-quark mass within the global MSHT PDF fit, Eur. Phys. J. C **83**, 805 (2023).

[138] Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. M. Alvarado, J. Rojo, and M. Ubiali, The top quark legacy of the LHC Run II for PDF and SMEFT analyses, J. High Energy Phys. 05 (2023) 205.

[139] M. Kadir, A. Ablat, S. Dulat, T.-J. Hou, and I. Sitiwaldi, The impact of ATLAS and CMS single differential top-quark pair measurements at $\sqrt{s} = 8$ TeV on CTEQ-TEA PDFs, Chin. Phys. C **45**, 023111 (2021).

[140] M. Czakon, S. Dulat, T.-J. Hou, J. Huston, A. Mitov, A. S. Papanastasiou, I. Sitiwaldi, Z. Yu, and C. P. Yuan, A study of the impact of double-differential top distributions from CMS on parton distribution functions, J. Phys. G **48**, 015003 (2020).

[141] S. Bailey and L. Harland-Lang, Differential top quark pair production at the LHC: Challenges for PDF fits, Eur. Phys. J. C **80**, 60 (2020).

[142] M. Czakon, N. P. Hartland, A. Mitov, E. R. Nocera, and J. Rojo, Pinning down the large-x gluon with NNLO top-quark pair differential distributions, J. High Energy Phys. 04 (2017) 044.

[143] M. Guzzi, K. Lipka, and S.-O. Moch, Top-quark pair production at hadron colliders: Differential cross section and phenomenological applications with DiffTop, J. High Energy Phys. 01 (2015) 082.

[144] R. Gauld, Feasibility of top quark measurements at LHCb and constraints on the large-x gluon PDF, J. High Energy Phys. 02 (2014) 126.

[145] M. Czakon, M. L. Mangano, A. Mitov, and J. Rojo, Constraints on the gluon PDF from top quark pair production at hadron colliders, J. High Energy Phys. 07 (2013) 167.

[146] C. Schmidt, J. Pumplin, C. P. Yuan, and P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach, Phys. Rev. D **98**, 094005 (2018).

[147] T.-J. Hou, Z. Yu, S. Dulat, C. Schmidt, and C. P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach. II, Phys. Rev. D **100**, 114024 (2019).

[148] E. R. Nocera, M. Ubiali, and C. Voisey, Single top production in PDF fits, J. High Energy Phys. 05 (2020) 067.

[149] M. Aaboud *et al.* (ATLAS Collaboration), Measurement of top quark pair differential cross-sections in the dilepton channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV with ATLAS, Phys. Rev. D **94**, 092003 (2016); **101**, 119901 (A) (2020).

[150] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV, Eur. Phys. J. C **75**, 542 (2015).

[151] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the inclusive t$\bar{\text{t}}$ cross section in pp collisions at $\sqrt{s} = 5.02$ TeV using final states with at least one charged lepton, J. High Energy Phys. 03 (2018) 115.

[152] CMS Collaboration, First measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 5.02$ TeV, Report No. CMS-PAS-TOP-16-015, 2016.

[153] G. Aad *et al.* (ATLAS Collaboration), Simultaneous measurements of the t$\bar{\text{t}}$, $W^{+}W^{-}$, and $Z/\gamma^{*} \rightarrow \tau\tau$ production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Rev. D **91**, 052005 (2015).

[154] ATLAS Collaboration, Measurement of the top quark pair production cross-section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV in the single-lepton channel using semi-leptonic b decays, Report No. ATLAS-CONF-2012-131, 2012.

[155] G. Aad *et al.* (ATLAS Collaboration), Measurements of the top quark branching ratios into channels with leptons and quarks with the ATLAS detector, Phys. Rev. D **92**, 072005 (2015).

[156] G. Aad *et al.* (ATLAS Collaboration), Measurement of the t$\bar{\text{t}}$ production cross section in the tau + jets channel using the ATLAS detector, Eur. Phys. J. C **73**, 2328 (2013).

[157] ATLAS Collaboration, Measurement of the t$\bar{t}$ production cross section in the all-hadronic channel in ATLAS with $\sqrt{s} = 7$ TeV data, Report No. ATLAS-CONF-2012-031, 2012.

[158] G. Aad *et al.* (ATLAS Collaboration), Measurement of the t$\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector, Eur. Phys. J. C **74**, 3109 (2014); **76**, 642(A) (2016).

[159] G. Aad *et al.* (ATLAS Collaboration), Measurement of the top pair production cross section in 8 TeV proton-proton collisions using kinematic information in the lepton + jets final state with ATLAS, Phys. Rev. D **91**, 112013 (2015).

[160] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section in the $e\mu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV, J. High Energy Phys. 08 (2016) 029.

[161] V. Khachatryan *et al.* (CMS Collaboration), Measurements of the t$\bar{t}$ production cross section in lepton + jets final states in pp collisions at 8 TeV and ratio of 8 to 7 TeV cross sections, Eur. Phys. J. C **77**, 15 (2017).

[162] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s} = 7$ TeV in dilepton final states containing a τ, Phys. Rev. D **85**, 112007 (2012).

[163] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section in the τ + Jets channel in pp collisions at $\sqrt{s} = 7$ TeV, Eur. Phys. J. C **73**, 2386 (2013).

[164] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section in the all-jet final state in pp collisions at $\sqrt{s} = 7$ TeV, J. High Energy Phys. 05 (2013) 065.

[165] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV in dilepton final states containing one τ lepton, Phys. Lett. B **739**, 23 (2014).

[166] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section in the all-jets final state in pp collisions at $\sqrt{s} = 8$ TeV, Eur. Phys. J. C **76**, 128 (2016).

[167] M. Aaboud *et al.* (ATLAS Collaboration), Measurement of the t$\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, Phys. Lett. B **761**, 136 (2016); **772**, 879 (E) (2017).

[168] G. Aad *et al.* (ATLAS Collaboration), Measurement of the t$\bar{t}$ production cross-section in the lepton + jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment, Phys. Lett. B **810**, 135797 (2020).

[169] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **116**, 052002 (2016).

[170] CMS Collaboration, Measurement of the inclusive and differential t$\bar{t}$ production cross sections in lepton + jets final states at 13 TeV, Report No. CMS-PAS-TOP-15-005, 2015.

[171] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the t$\bar{t}$ production cross section using events in the $e\mu$ final state in pp collisions at $\sqrt{s} = 13$ TeV, Eur. Phys. J. C **77**, 172 (2017).

[172] CMS Collaboration, Measurement of the t$\bar{t}$ production cross section at 13 TeV in the all-jets final state, Report No. CMS-PAS-TOP-16-013, 2016.

[173] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV, Phys. Rev. D **93**, 072004 (2016).

[174] A. Mitov and G. Sterman, Final state interactions in single- and multi-particle inclusive cross sections for hadronic collisions, Phys. Rev. D **86**, 114038 (2012).

[175] T. Kluge, K. Rabbertz, and M. Wobisch, `fastNLO`: Fast pQCD calculations for PDF fits, in *Proceedings of the 14th International Workshop on Deep Inelastic Scattering* (2006), pp. 483–486, arXiv:hep-ph/0609285.

[176] M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, and F. Stober (fastNLO Collaboration), Theory-data comparisons

for jet measurements in hadron-induced processes, arXiv:1109.1310.

[177] D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch (fastNLO Collaboration), New features in version 2 of the fastNLO project, in *Proceedings of the 20th International Workshop on Deep-Inelastic Scattering and Related Subjects* (2012), pp. 217–221, arXiv:1208.3641.

[178] D. Britzger, G. S. Klaus Rabbertz, F. Stober, and M. Wobisch, Recent developments of the fastNLO toolkit, Proc. Sci. DIS2015 (**2015**) 055.

[179] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov, and M. Sutton, A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGrid Project, Eur. Phys. J. C **66**, 503 (2010).

[180] M. Czakon, D. Heymes, and A. Mitov, fastNLO tables for NNLO top-quark pair differential distributions, arXiv:1704.08551.

[181] Repository with $t\bar{t}$ fastNNLO tables, https://www.precision.hep.phy.cam.ac.uk/results/ttbar-fastnlo/.

[182] S. Catani and M. Grazzini, An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC, Phys. Rev. Lett. **98**, 222002 (2007).

[183] Repository with MATRIX code, https://matrix.hepforge.org/.

[184] J. M. Campbell and R. K. Ellis, Top-quark processes at NLO in production and decay, J. Phys. G **42**, 015005 (2015).

[185] H. Paukkunen and P. Zurita, PDF reweighting in the Hessian matrix approach, J. High Energy Phys. 12 (2014) 100.

[186] W. T. Giele and S. Keller, Implications of hadron collider observables on parton distribution function uncertainties, Phys. Rev. D **58**, 094023 (1998).

[187] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, Reweighting and unweighting of parton distributions and the LHC W lepton asymmetry data, Nucl. Phys. **B855**, 608 (2012).

[188] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali (NNPDF Collaboration), Reweighting NNPDFs: The W lepton asymmetry, Nucl. Phys. **B849**, 112 (2011); **B854**, 926(E) (2012); **B855**, 927(E) (2012).

[189] B.-T. Wang, T. J. Hobbs, S. Doyle, J. Gao, T.-J. Hou, P. M. Nadolsky, and F. I. Olness, Mapping the sensitivity of hadronic experiments to nucleon structure, Phys. Rev. D **98**, 094030 (2018).

[190] C. Willis, R. Brock, D. Hayden, T.-J. Hou, J. Isaacson, C. Schmidt, and C.-P. Yuan, New method for reducing parton distribution function uncertainties in the high-mass Drell-Yan spectrum, Phys. Rev. D **99**, 054004 (2019).

[191] N. Yalkun and S. Dulat, QCD analysis of CMS W + charm measurements at LHC with $\sqrt{s} = 7$ TeV and implications for the strange PDF, Chin. Phys. C **43**, 123101 (2019).

[192] A. Ablat, S. Dulat, R. Rashidin, A. Ruzi, and N. Yalkun, The impact of single top data on CT14nnlo PDFs, Int. J. Theor. Phys. **59**, 3023 (2020).

[193] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C. P. Yuan, Implications of CTEQ global analysis for collider observables, Phys. Rev. D **78**, 013004 (2008).

[194] The ATLAS Collaboration, HEPData repository for ATLAS lep + jet 13 TeV, Eur. Phys. J. C **79**, 12 (2019).

[195] G. D'Agostini, On the use of the covariance matrix to fit correlated data, Nucl. Instrum. Methods Phys. Res., Sect. A **346**, 306 (1994).

[196] G. D'Agostini, Bayesian reasoning in high-energy physics: Principles and applications, Report No. CERN-99-03, CERN-YELLOW-99-03, CERN Yellow Reports: Monographs.

[197] HEPforge repository for ePump, https://epump.hepforge.org/.

[198] K. Cranmer *et al.*, Publishing statistical models: Getting the most out of particle physics experiments, SciPost Phys. **12**, 037 (2022).

[199] M. Czakon, D. Heymes, and A. Mitov, Dynamical scales for multi-TeV top-pair production at the LHC, J. High Energy Phys. 04 (2017) 071.

[200] U.-K. Yang *et al.* (CCFR/NuTeV Collaboration), Measurements of F_2 and $xF_3^\nu - xF_3^{\bar{\nu}}$ from CCFR $\nu_\mu -$ Fe and $\bar{\nu}_\mu -$ Fe data in a physics model independent way, Phys. Rev. Lett. **86**, 2742 (2001).

[201] V. Khachatryan *et al.* (CMS Collaboration), Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV, J. High Energy Phys. 03 (2017) 156.

[202] R. Aaij *et al.* (LHCb Collaboration), Measurement of forward $Z \to e^+e^-$ production at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 05 (2015) 109.

[203] X. Jing *et al.*, Quantifying the interplay of experimental constraints in analyses of parton distributions, Phys. Rev. D **108**, 034029 (2023).

[204] T. J. Hou, H. W. Lin, M. Yan, and C. P. Yuan, Impact of lattice strangeness asymmetry data in the CTEQ-TEA global analysis, Phys. Rev. D **107**, 076018 (2023).

[205] T. Cridge (PDF4LHC21 Combination Group), PDF4LHC21: Update on the benchmarking of the CT, MSHT and NNPDF global PDF fits, SciPost Phys. Proc. **8**, 101 (2022).

[206] R. D. Ball *et al.*, The PDF4LHC21 combination of global PDF fits for the LHC Run III, J. Phys. G **49**, 080501 (2022).

[207] R. A. Khalek *et al.*, Phenomenology of NNLO jet production at the LHC and its impact on parton distributions, Eur. Phys. J. C **80**, 797 (2020).

[208] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the electron charge asymmetry in $p\bar{p} \to W + X \to e\nu + X$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, Phys. Rev. D **91**, 032007 (2015); **91**, 079901(E) (2015).

[209] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the proton structure functions $F_2(X, Q^2)$ and R from deep inelastic muon scattering at high Q^2, Phys. Lett. B **223**, 485 (1989).

[210] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the deuteron structure functions

$F_2(X, Q^2)$ and R from deep inelastic muon scattering at high Q^2, Phys. Lett. B **237**, 592 (1990).

[211] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. High Energy Phys. 07 (2014) 079.

[212] K. Kovařík, P. M. Nadolsky, and D. E. Soper, Hadronic structure in high-energy collisions, Rev. Mod. Phys. **92**, 045003 (2020).

[213] Z. Kassabov, E. R. Nocera, and M. Wilson, Regularising experimental correlations in LHC data: Theory and application to a global analysis of parton distributions, Eur. Phys. J. C **82**, 956 (2022).