Developments in the parton distribution functions of the proton

Marco Guzzi[1,a]

[1]*Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany*

Abstract. A general overview on the parton distribution functions of the proton extracted by QCD global analyses of experimental data is presented. In particular, their impact on physics observables of interest for the LHC and general results of QCD analyses at the next-to-next-to-leading order approximation are discussed.

1 Introduction

The current inflow of new experimental measurements with unprecedented energies and accuracy from hadron colliders requires theory predictions at high order in perturbation theory, which are needed for a consistent description the data. Furthermore, the development of efficient tools incorporating the state-of-the-art of QCD calculations and new methods for the analyses is necessary to fully exploit the constraining power of the present and forthcoming measurements.

The recent discovery of the Higgs boson [1, 2] and extensive searches for signals of new physics in hadron-hadron collisions at the Large Hadron Collider (LHC) and Tevatron, demand high-precision predictions to test the validity of the Standard Model (SM) and factorization theorems in Quantum Chromodynamics (QCD). QCD factorization requires the knowledge of the perturbative calculable hard scattering contributions and determination of nonperturbative parton distribution functions (PDFs) of the nucleon, which are indispensable ingredients of all hard-scattering processes involving initial-state hadrons. Investigations on the structure of the nucleon are crucial for a multitude of current and future high-energy physics programs and the progress is based on combined efforts of theorists and experimentalists.

The interpretation of experimental measurements at hadron colliders relies to a large extent on the precise knowledge of the fundamental QCD parameters and PDFs. According to QCD factorization, the generic structure of the inclusive cross section for the high-energy collision of two hadrons (h_1, h_2) is written as a convolution product where the distributions of elementary partons inside the hadrons (PDFs) and the infrared safe short-distance contributions (hard scatterings), are separated. The factorization formula for the unpolarized Drell-Yan (DY) machanism is chosen here as an illustrative case, because this formulation has been extensively studied in the past and is valid for a large class of processes [3–8].

The hadronic reaction for the DY process is written as $h_1 + h_2 \rightarrow (V^* \rightarrow l^+l^-) + X$, where the initial state partons annihilate into a virtual boson $(\gamma, Z, W^\pm)$ of virtuality Q^2, which subsequently

[a]e-mail: marco.guzzi@desy.de

converts into a pair of final-state leptons that are detected. The hadronic state X denotes any remnants of fragmented protons in the final state. The factorized inclusive cross section is given by

$$\sigma(h_1 h_2 \to l^+ l^- + X) = \sum_{a,b} \int_0^1 dx_1 \int_0^1 dx_2 \, f_{h_1 \to a}(x_1, \alpha_s(\mu_r), \mu_f) f_{h_2 \to b}(x_2, \alpha_s(\mu_r), \mu_f)$$

$$\times \hat{\sigma}^{ab}(x_1, x_2; \alpha_s(\mu_r), Q, \mu_f, \mu_r) + O\left(\frac{\Lambda^2}{Q^2}\right), \tag{1}$$

where the summation indices a, b run over all possible intial-state partons (gluons, quarks and anti-quarks of different flavours), whose fractions of longitudinal momentum are indicated by x_1 and x_2 respectively. PDFs are denoted by $f_{h_i \to j}$ with $i = 1, 2$ and $j = a, b$, and represent the probability of finding parton j carrying a fraction x_i of hadron momentum, in hadron i, at the factorization scale μ_f. The hard-scattering contribution for the process $a + b \to l^+ + l^- + X$, calculable in perturbation theory, is denoted by $\hat{\sigma}^{ab}$. The range of validity of the factorized formula is for large virtualities when $Q^2 \sim \sqrt{\hat{s}}$ (where $\sqrt{\hat{s}} = \sqrt{x_1 x_2 S}$). The term $O\left(\Lambda^2/Q^2\right)$ represents suppressed corrections which are not treated here. In general, expressions in perturbative calculations are obtained by truncating the perturbative series at a certain fixed order of the coupling constant α_s. This introduces a dependence on arbitrary scales denoted as renormalization μ_r and factorization μ_f scales respectively. In particular, μ_f is the scale at which the separation between long-distance and short-distance effects is realized. Gluons, quarks and antiquarks are the known constituents of the proton and their distributions as a function of the momentum fraction x and generic scale μ, at which partons are probed, are universal quantities that do not depend on the specific hard process under consideration. Differently from the hard-scattering cross section, the analytic structure of the PDFs cannot be predicted by perturbative QCD (pQCD), but is determined by comparing sets of physical cross sections, such as in Eq. 1, to experimental measurements. For this reason PDFs are considered as "data-driven" quantities. Furthermore, similarly to the case of renormalization scheme, a set of rules has to be provided in order to define the PDFs when a cross section calculation is performed, because the separation of PDFs from hard-scatterings is not unique in the factorized formula. These rules are known as "factorization scheme" and are necessary to calculate consistently the cross sections that have to be compared to the experimental measurements. The most widely used factorization scheme is the modified minimal subtraction scheme $\overline{\text{MS}}$. Formally, a PDF is defined as the expectation value of a certain operator in a hadron state and it is constructed such that it contains an ultra-violet (UV) regulator in dimensional regularization (DR) with $n = 4 - 2\varepsilon$ space-time dimensions and is gauge invariant (see Refs. [9–12] for more details). For the case of quarks one has,

$$f_{(0) \, j/h}(\xi) = \int \frac{dw^-}{2\pi} \, e^{-i\xi P^+ w^-} \langle P | \overline{\psi}_j^{(0)}(0, w^-, \mathbf{0}_\text{T}) W(w^-, 0) \frac{\gamma^+}{2} \psi_j^{(0)}(0) | P \rangle_\text{c}, \tag{2}$$

where the Wilson-line factor for gauge invariance is

$$W(w^-, 0) = P\left[e^{-ig_0 \int_0^{w^-} dy^- A_{(0)\alpha}^+(0, y^-, \mathbf{0}_\text{T}) t_\alpha} \right]. \tag{3}$$

A direct consequence of factorization is that the scale dependence or "evolution" of PDFs can be predicted by the renormalisation group equations (RGE's). By imposing that physical observables are independent on μ_f, it leads to a representation of parton evolution in terms of integro-differential equations known as Dokshitzer-Gribov-Lipatov-Altarelli-Parisi [13–17] (DGLAP) equations

$$\frac{d f_i(x, \mu_r, \mu_f)}{d \log \mu_f^2} = \sum_{j=q\bar{q},g} \int_x^1 \frac{dy}{y} P_{ij}\left(\frac{x}{y}; \alpha_s, \mu_r, \mu_f\right) f_j(y, \mu_r, \mu_f), \tag{4}$$

where the functions P_{ij} are the evolution kernels or splitting functions [18–20], which represent the probability of finding parton i in parton j, and have perturbative expansion in α_s. Therefore, once PDFs are determined by a direct comparison with the experiments at the initial scale Q_0, their evolution at the scale μ is entirely determined by DGLAP equations. In the past few years, the precision of several experimental measurements noticeably increased and the relative uncertainties became compatible with those of the theory predictions. Efforts to reduce PDF uncertainties are therefore crucial for future high-energy physics programs at hadron colliders, in particular for discriminating signals of new physics from SM. The goal of pinning down PDF errors and their mutual correlations is achieved by using combined constraints from measurements of different processes. This relies on the ongoing progress in theory calculations and on the possibility of including a larger number of experimental data sets in new efficient tools for PDF determination.

2 Global QCD analyses in the NNLO era

In the past decade methods adopted in perturbative calculations of the hard-scattering cross sections have experienced a tremendous progress. Numerous techniques to solve complicated multidimensional integrals, also in the presence of mass, have been introduced as well as methodologies to reduce their multiplicities. New sophisticated ways to calculate loops [21, 22] based on generalized unitarity methods [23–27] and more efficient ways to calculate real emission corrections have been developed [28, 29] together with other regularization schemes [30–33]. These techniques are crucial for the extractions of the pole singularities in the NLO and NNLO amplitudes. Improved tensor reduction schemes [34, 35] to manage stability at numerical level are employed in the presence of a large number of Feynman diagrams, and in disparate cases, numeric approaches present more advantages with respect to analytic calculations. Several NLO QCD calculations for multi-particle final state processes $2 \rightarrow 3, 4, ..$ of high complexity and requiring great numerical stability have been recently accomplished (see for example Refs. [36, 37]). The most recent progress for $2 \rightarrow 2$ processes at NNLO with colored initial and final state has been possible thanks to the development of all these new techniques (see for example Refs. [38–40]). Despite the high level of accuracy reached in the determination of the hard scattering cross sections, a precise determination of the partonic content of the proton continuosly requires a lot of efforts on both theoretical and experimental sides. The uncertainties associated to the PDFs represent one of the dominant contribution to the errors on hadronic cross sections. The PDFs of the proton cannot be derived from the first principles and must determined from experimental data by using a variety of sophisticated analytical and statistical methods. Unpolarized PDF sets to the NNLO accuracy are now provided by the main PDF groups [41–46] which include NNLO cross section calculations and diverse world data sets in their analyses. While these PDFs are dominated by pre-LHC measurements, work is ongoing to include new LHC data to provide unique new constraints. Furthermore, different PDF groups use different approaches and methodologies in their fits. This leads to different theoretical predictions and errors for physical observable of interest, like the Higg cross section. Most of the differences can be ascribed to several causes: data sets considered in the analyses, treatment of the heavy-flavor, input assumptions for the PDFs at Q_0, different values of $\alpha_s(M_Z)$, etc.. An accurate comparison of the theoretical settings and fitting methodologies employed in the determination of PDFs with emphasis on factors which are important for the Higgs production can be found in [47]. In Fig. 1 shown is the gluon PDF obtained in fits of different PDF groups.
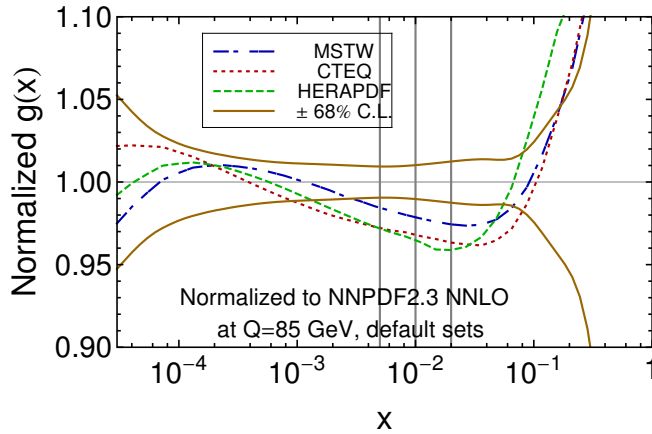
Figure 1. Central gluon PDFs from recently published NNLO PDF sets at scale $Q = 85$ GeV, normalized to the NNPDF2.3 gluon PDF and superimposed on the NNPDF2.3 68% c.l. PDF uncertainty. Figure from Ref. [47].

3 Most relevant contrains on PDFs from data

Measurements of lepton-proton deep-inelastic-scattering (DIS) at HERA [48] are the most important data sets in PDFs determination and provide the staple information about the proton structure at scales Q below a few hundred GeV, while the hadron-(anti)hadron collider data probe PDFs in a broad range of energies. Inclusive DIS is used as a standard candle measurement of the proton structure functions and serve as the backbone for the PDF determination at low and medium x. Semi-inclusive DIS processes like production of electroweak bosons, jets, and heavy quarks probe details of the flavor composition of PDFs and their dependence on the QCD coupling strength and heavy-quark masses. Measurements of electroweak boson production at the LHC are sensitive to PDFs which are poorly constrained in certain kinematic regions, like flavor composition of the quark sea at $x < 0.01$. Lepton pair production at ATLAS and CMS covers the medium x range, while LHCb is well-suited for investigating regions of low and high x. In particular these measurements at rapidities $y > 2$ can probe different PDF sets at $x > 0.1$. Jet production measurements at the LHC and Tevatron are sensitive to the gluon distribution in the proton and QCD coupling α_s. Furthermore, top-quark pair production at the LHC, being dominated by gluon-gluon fusion, can be an important probe for the gluon distribution at large x and will allow us to pin down the correlation between top-quark mass m_t and α_s. Precise measurements of total and differential cross sections for top-quark pair production at a center-of-mass energies $\sqrt{S} = 7$ and 8 TeV have been recently published by the CMS [49, 50] and ATLAS [51–54] collaborations.

4 Recent NNLO QCD analysis including $t\bar{t}$ data

An example of the possible impact of the currently available LHC measurements of $t\bar{t}$ production on PDF determination has been recently illustrated in [55], where the open source DiffTop code [56] to calculate the differential and total cross section at approximate NNLO in QCD, is interfaced to the HERAFitter [57] platform for QCD analyses to determine the PDFs. Fast theoretical calculations for the approximate NNLO p_T spectrum of the final-state top quark (see Fig.2) are obtained by using grids generated by the FastNLO package [58–60]. This allows the user to include measurements of

differential $t\bar{t}$ production cross sections into NNLO QCD fits of PDFs. The NNLO parton evolution utilized is that implemented in the QCDNUM [61] code, which is the default parton-evolution package of HERAFITTER. The combined HERA I measurements [48] of inclusive DIS are included in this analysis. Constraints on the u and d-quark distributions in the x-range not properly covered by the HERA I measurements are put by the CMS precise measurements of electron [62] and muon [63] charge asymmetry in W-boson production at $\sqrt{S} = 7$ TeV. The $t\bar{t}$ measurements included here are the total inclusive cross sections at the LHC [64–67] at $\sqrt{S}$ of 7 and 8 TeV, and CDF [68] Tevatron, as well as the normalized differential cross-sections [49, 52] of $t\bar{t}$ production at the LHC at $\sqrt{S} = 7$ TeV as a function of p_T. The results of this analysis in which the differential cross sections for $t\bar{t}$ are used in a NNLO QCD fit for the first time, are shown in Fig. 3. A moderate impact of the new LHC data sets on the reduction of the gluon uncertainty at large x is illustrated. In the future, a reduction of the statistic and systematic uncertainties in the high-energy run of the LHC will be of clear advantage and will allow us for a simultaneous determination of the gluon distribution and the top-quark mass.
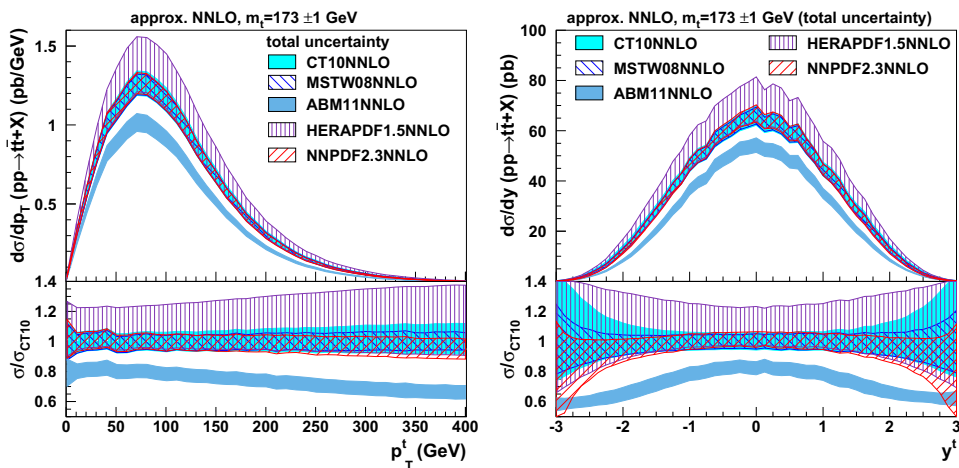
Figure 2. The approximate NNLO predictions for top-quark pair production cross sections at the LHC at $\sqrt{S} = 7$ TeV as functions of p_T^t (left) and y^t (right). Predictions obtained by using different PDF sets are presented by bands of different hatches. The total uncertainty is obtained by summing the uncertainties due to PDFs, α_s, m_t and scale variations in quadrature. Figs. from [55]

5 Conclusions

In this brief paper few important aspects of the determination of the PDFs of the proton were outlined together with an example of recent progress in NNLO PDF fits. LHC unprecedented energies brought us in a new precision era. Huge efforts are ongoing among the various PDF groups to reduce the PDFs uncertainties, which still remain among the major sources of systematical uncertainties in theoretical predictions for crucial observables at the LHC. The high luminosity that will be reached at LHC run II will shed light on many aspects and details of the PDFs in kinematic ranges which are currently poorly constrained. This will play a fundamental role in the signal/background discrimation and in searches of any possible signal of Physics beyond the Standard Model.
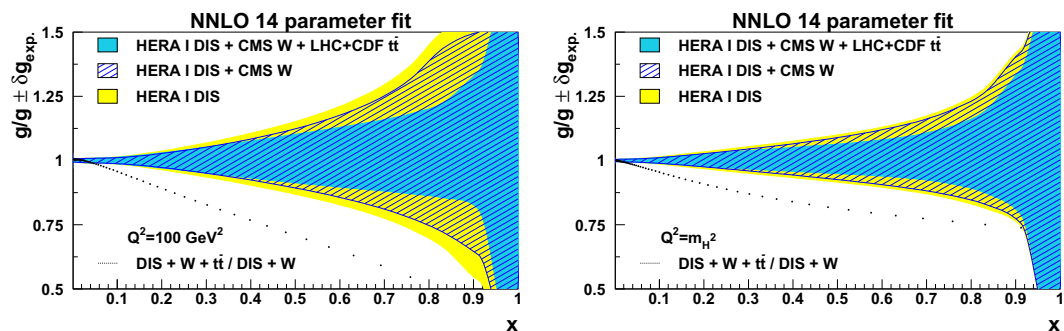
Figure 3. Uncertainties of the gluon distribution as a function of *x*, as obtained in our NNLO fit by using: inclusive DIS measurements only (light shaded band), DIS and *W* lepton charge asymmetry data (hatched band), and DIS, lepton charge asymmetry and the $t\bar{t}$ production measurements (dark shaded band), shown at the scales of $Q^2 = 100$ GeV^2 (left) and $Q^2 = m_H^2$ (right). The ratio of $g(x)$ obtained in the fit including $t\bar{t}$ data to that obtained by using DIS and lepton charge asymmetry, is represented by a dotted line. Figs. from [55]

Acknowledgements

References

[1] G. Aad et al. (ATLAS Collaboration), Phys.Lett. **B716**, 1 (2012), `1207.7214`

[2] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. **B716**, 30 (2012), `1207.7235`

[3] J.C. Collins, D.E. Soper, G.F. Sterman, Phys.Lett. **B109**, 388 (1982)

[4] J.C. Collins, D.E. Soper, G.F. Sterman, Nucl.Phys. **B223**, 381 (1983)

[5] J.C. Collins, D.E. Soper, G.F. Sterman, Phys.Lett. **B134**, 263 (1984)

[6] J.C. Collins, D.E. Soper, G.F. Sterman, Nucl.Phys. **B250**, 199 (1985)

[7] J.C. Collins, D.E. Soper, G.F. Sterman, Nucl.Phys. **B261**, 104 (1985)

[8] J.C. Collins, D.E. Soper, G.F. Sterman, Adv.Ser.Direct.High Energy Phys. **5**, 1 (1988), `hep-ph/0409313`

[9] J. Collins, *Foundations of perturbative QCD*, Vol. 32 (Cambridge monographs on particle physics, nuclear physics and cosmology., 2011)

[10] J. Collins, Int.J.Mod.Phys.Conf.Ser. **4**, 85 (2011), `1107.4123`

[11] J. Collins, PoS **LC2008**, 028 (2008), `0808.2665`

[12] J.C. Collins, Acta Phys.Polon. **B34**, 3103 (2003), `hep-ph/0304122`

[13] V.N. Gribov, L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972)

[14] V.N. Gribov, L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 675 (1972)

[15] L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975)

[16] Y.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977)

[17] G. Altarelli, G. Parisi, Nucl. Phys. B **126**, 298 (1977)

[18] D. Gross, F. Wilczek, Phys.Rev. **D8**, 3633 (1973)

[19] G. Curci, W. Furmanski, R. Petronzio, Nucl.Phys. **B175**, 27 (1980)

[20] A. Vogt, S. Moch, J. Vermaseren, Nucl.Phys. **B691**, 129 (2004), `hep-ph/0404111`

[21] Z. Bern, L.J. Dixon, D.C. Dunbar, D.A. Kosower, Nucl.Phys. **B435**, 59 (1995), `hep-ph/9409265`

[22] Z. Bern, L.J. Dixon, D.C. Dunbar, D.A. Kosower, Nucl.Phys. **B425**, 217 (1994), `hep-ph/9403226`

[23] R. Britto, F. Cachazo, B. Feng, E. Witten, Phys.Rev.Lett. **94**, 181602 (2005), `hep-th/0501052`

[24] R. Britto, F. Cachazo, B. Feng, Nucl.Phys. **B715**, 499 (2005), `hep-th/0412308`

[25] R. Britto, F. Cachazo, B. Feng, Nucl.Phys. **B725**, 275 (2005), `hep-th/0412103`

[26] R.K. Ellis, W. Giele, Z. Kunszt, JHEP **0803**, 003 (2008), `0708.2398`

[27] G. Ossola, C.G. Papadopoulos, R. Pittau, Nucl.Phys. **B763**, 147 (2007), `hep-ph/0609007`

[28] S. Catani, S. Dittmaier, M.H. Seymour, Z. Trocsanyi, Nucl.Phys. **B627**, 189 (2002), `hep-ph/0201036`

[29] G. Bevilacqua, M. Czakon, C. Papadopoulos, M. Worek, Phys.Rev. **D84**, 114017 (2011), `1108.2851`

[30] S. Frixione, Z. Kunszt, A. Signer, Nucl.Phys. **B467**, 399 (1996), `hep-ph/9512328`

[31] A. Gehrmann-De Ridder, T. Gehrmann, E.N. Glover, JHEP **0509**, 056 (2005), `hep-ph/0505111`

[32] B. Harris, J. Owens, Phys.Rev. **D65**, 094032 (2002), `hep-ph/0102128`

[33] W. Giele, E.N. Glover, D.A. Kosower, Nucl.Phys. **B403**, 633 (1993), `hep-ph/9302225`

[34] A.I. Davydychev, M.Y. Kalmykov, Nucl.Phys. **B699**, 3 (2004), `hep-th/0303162`

[35] T. Binoth, J.P. Guillet, G. Heinrich, E. Pilon, C. Schubert, JHEP **0510**, 015 (2005), `hep-ph/0504267`

[36] Z. Bern, L. Dixon, F. Febres Cordero, S. Höche, H. Ita et al., Phys.Rev. **D88**, 014025 (2013), `1304.1253`

[37] H. Ita, Z. Bern, L. Dixon, F. Febres Cordero, D. Kosower et al., Phys.Rev. **D85**, 031501 (2012), `1108.2229`

[38] S. Catani, L. Cieri, G. Ferrera, D. de Florian, M. Grazzini, Phys.Rev.Lett. **103**, 082001 (2009), `0903.2120`

[39] S. Catani, L. Cieri, D. de Florian, G. Ferrera, M. Grazzini, Phys.Rev.Lett. **108**, 072001 (2012), `1110.2375`

[40] M. Czakon, P. Fiedler, A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013), `1303.6254`

[41] S. Alekhin, J. Blümlein, S. Moch, Phys.Rev. **D86**, 054009 (2012), `1202.2281`

[42] J. Gao, M. Guzzi, J. Huston, H.L. Lai, Z. Li et al., Phys.Rev. **D89**, 033009 (2014), `1302.6246`

[43] A. Cooper-Sarkar (ZEUS Collaboration, H1 Collaboration), PoS **EPS-HEP2011**, 320 (2011), `1112.2107`

[44] A. Martin, W. Stirling, R. Thorne, G. Watt, Eur.Phys.J. **C63**, 189 (2009), `0901.0002`

[45] R.D. Ball, V. Bertone, S. Carrazza, C.S. Deans, L. Del Debbio et al., Nucl.Phys. **B867**, 244 (2013), `1207.1303`

[46] P. Jimenez-Delgado, E. Reya, Phys.Rev. **D79**, 074023 (2009), `0810.4274`

[47] J. Butterworth, G. Dissertori, S. Dittmaier, D. de Florian, N. Glover et al. (2014), `1405.1067`

[48] F. Aaron et al. (H1 and ZEUS Collaboration), JHEP **1001**, 109 (2010), `0911.0884`

[49] S. Chatrchyan et al. (CMS Collaboration), Eur.Phys.J. **C73**, 2339 (2013), `1211.2220`

[50] S. Chatrchyan et al. (CMS Collaboration), JHEP **1402**, 024 (2014), `1312.7582`

[51] G. Aad et al. (ATLAS Collaboration), Eur.Phys.J. **C73**, 2261 (2013), `1207.5644`

[52] G. Aad et al. (ATLAS Collaboration) (2014), `1407.0371`

[53] The Atlas Collaboration ATLAS-CONF-2013-099, ATLAS-COM-CONF-2013-114, (2013)

[54] The Atlas Collaboration ATLAS-CONF-2013-097, ATLAS-COM-CONF-2013-112, (2013)

[55] M. Guzzi, K. Lipka, S.O. Moch (2014), `1406.0386`

[56] M. Guzzi, K. Lipka, S.O. Moch http://difftop.hepforge.org/

[57] The HeraFitter Collaboration https://wiki-zeuthen.desy.de/HERAFitter

[58] T. Kluge, K. Rabbertz, M. Wobisch, pp. 483–486 (2006), `hep-ph/0609285`

[59] M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, F. Stober (fastNLO Collaboration) (2011), `arXiv:1109.1310`

[60] D. Britzger, K. Rabbertz, F. Stober, M. Wobisch (fastNLO Collaboration) (2012), `arXiv:1208.3641`

[61] M. Botje (2010), http://www.nikef.nl/h24/qcdnum/index.html, `arXiv:1005.1481`

[62] V. Khachatryan et al. (CMS Collaboration), JHEP **1101**, 080 (2011), `1012.2466`

[63] S. Chatrchyan et al. (CMS Collaboration), Phys.Rev. **D90**, 032004 (2014), `1312.6283`

[64] ATLAS-CONF-2012-024, ATLAS-COM-CONF-2012-009 (2012)

[65] ATLAS-CONF-2012-149, ATLAS-COM-CONF-2012-170 (2012)

[66] S. Chatrchyan et al. (CMS Collaboration), JHEP **1211**, 067 (2012), `1208.2671`

[67] C. Collaboration (CMS Collaboration) (2012)

[68] C. Collaboration (CDF Collaboration) (2009)