Heavy-flavor impact on CTEQ-TEA global QCD analyses

Marco Guzzi^{1,*}, *Alim* Ablat², *Sayipjamal* Dulat², *Tie-Jiun* Hou³, *Pavel* Nadolsky⁴, *Ibrahim* Sitiwaldi², *Keping* Xie⁵, and *C.-P.* Yuan⁶

Abstract. We discuss heavy-flavor production at hadron colliders in recent global QCD analyses to determine parton distribution functions (PDFs) in the proton. We discuss heavy-flavor treatments in precision theory predictions at the LHC. In particular, we discuss factorization schemes in presence of heavy flavors in proton-proton collisions, as well as the impact of heavy-flavor production at the LHC on PDFs. We show results of recent updates beyond CT18, the latest global QCD analysis from the CTEQ-TEA group.

1 Introduction

Precise determination of parton distribution functions (PDFs) in the proton is critical for all current and future precision programs at the LHC. Proton PDFs are an essential ingredient of QCD factorization theorems and are therefore ubiquitous in theory predictions for standard candle observables in hadronic collisions at high perturbative order in QCD. Precise and accurate theory predictions for such observables are necessary to investigate properties of the Higgs boson and to explore the Electroweak (EW) sector of the Standard Model (SM). In addition, they are important to scrutinize and validate SM extensions and search for signals of new physics interactions. PDFs are obtained through global QCD analyses of experimental hadronic cross section measurements by using a variety of analytical and statistical methods [1–4]. They are "data-driven" quantities and they currently challenge precision in theory predictions at the LHC. In particular, precision measurements of heavy-quark (HQ) production observables in hadronic reactions play a significant role in PDF determinations because they can constrain PDFs in kinematic regions at intermediate and large parton momentum fraction x which are currently poorly constrained by data. In addition, they provide us with complementary information besides that obtained from jet production in global QCD analyses. HQ production at the LHC at small transverse momentum p_T and large rapidity y of the HQ is sensitive to PDFs at both small and large x, where $x \approx \left(\sqrt{p_T^2 + m_Q^2}/\sqrt{S}\right)e^{\pm y}$. This allows us to probe QCD dynamics simultaneously in these two regimes. This is especially true for charm and bottom (c/b) quark production. In fact, in the 4 < |y| < 4.5 rapidity range

¹Department of Physics, Kennesaw State University, 370 Paulding Ave., 30144 Kennesaw, GA, USA

²School of Physics Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046, China

³School of Nuclear Science and Technology, Univ. of South China, Hengyang, Hunan 421001, China

⁴Department of Physics, Southern Methodist University, Dallas, TX 75275-0181, USA

⁵Pittsburgh Particle Physics, Astrophysics, and Cosmology Center, Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA

⁶Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA

^{*}e-mail: mguzzi@kennesaw.edu,presenter

at the LHC with a proton-proton collision energy of 13 TeV, it can probe $x \le 10^{-5}$. When $p_T \ge 40$ GeV, it can probe $x \ge 0.2$. On the other hand, top-quark pair production at the LHC, at the same collision energy, can probe the gluon PDF already at $x \ge 0.01$, while c/b production at HERA is sensitive to intermediate and small $x \le 10^{-5}$. Probing these regimes (and beyond, at future facilities like the Electron Ion Collider (EIC) [5–7], the Future Circular Collider (FCC) [8–10], the Forward Physics Facility at the High-Luminosity (HL) LHC [11, 12], and the Super proton-proton Collider (SppC) [13]) will allow us to explore QCD factorization with an unprecedented level of accuracy. This will shed light on open questions like the intrinsic heavy-flavor content in the proton, and small-x dynamics. In this conference proceeding contribution, we shall report preliminary results of global PDF analyses which use recent high-precision HQ measurements from the LHC and HERA which are added on top of the CT18 [1] baseline. Moreover, treatments of QCD factorization in presence of HQ multiscale processes are also discussed.

2 HQ treatments in QCD factorization and the S-ACOT-MPS scheme

HQ production dynamics is nontrivial due to the interplay of massless and massive schemes, which essentially are different ways of organizing the perturbation series. In general, one can calculate cross sections by using massive schemes, that include HQs only in the final-state and apply at masses such that $p_T \leq m_Q$. In this case, the HQ p_T -spectrum can be obtained by keeping the number of flavors as fixed, i.e., using a fixed-flavor number (FFN) scheme. That is, there is no HQ PDF in the proton; HQs are generated as massive final states, and m_Q acts as an infrared cutoff. Moreover, all power terms of the type $\left(p_T^2/m_Q^2\right)^p$, where p is a positive integer, are correctly accounted for in the perturbative series. Cross sections can also be computed by using massless schemes when $p_T \gg m_O \gg m_P$, where m_P is the mass of the proton. In this case, large logarithmic terms of the type $\log^n\left(p_T^2/m_Q^2\right)$ spoil the convergence of the fixed-order expansion. Massless schemes are also known as zero-mass (ZM) schemes; the HQ is considered essentially massless everywhere and also enters the running of the strong coupling α_s . These large logarithms need to be resummed by using DGLAP evolution, that is, initial-state logs are resummed into a HQ PDF, while final-state logarithms are resummed into a fragmentation function (FF). Modern global QCD analyses determine proton PDFs using a variety of experimental data which span wide areas of the Q-x kinematic plane. Therefore, amendments to the factorization formula are necessary to calculate key cross sections across wide ranges of energy and momentum transfer. Interpolating schemes, such as General Mass Variable Flavor Number (GMVFN) schemes, introduce modifications in QCD collinear factorization so that they appear as composite schemes which retain key mass dependence and efficiently resum collinear logarithms in order to combine the FFN and ZM schemes together. They are crucial for a correct treatment of HQ in deep inelastic scattering (DIS) processes [14–22] and proton-proton (pp) collisions, and for accurate predictions of key scattering rates at the LHC in global QCD analyses to determine proton PDFs. Moreover, they provide us with the possibility of directly accessing HQ PDFs parametrized at the initial scale. This motivated the development of a general-mass (GM) factorization scheme for pp collisions [23, 24], the Simplified ACOT scheme with Massive Phase Space (S-ACOT-MPS), which is based on an amended version of the S-ACOT scheme, developed for DIS [14, 15, 18, 19, 22]. Similar GMVFN schemes are studied in [25–27] and differences are related to the treatment of the phase space. The main idea behind a GMVFN scheme such as S-ACOT-MPS is the introduction of subtraction terms order-by-order in perturbation theory which avoid double counting and cancel enhanced collinear contributions from flavor-creation (FC) terms when $\hat{s} \gg m_Q^2$, or $p_T \gg M_Q$. In S-ACOT-MPS, the subtraction and flavor excitation (FE) contributions are evaluated in one single step. This improves stability in calculating differential cross sections and facilitates the extension of this scheme to higher orders. We applied S-ACOT-MPS to prompt c and b production at NLO in QCD at the LHC, at $\sqrt{S}=13$ TeV. In Figure 1 (left), we illustrate the c-rapidity distribution determined with the CT18 and CT18X PDFs at NLO. We explore the sensitivity of prompt c production to nonperturbative charm contributions in the proton by considering theory predictions in very forward regions with $y_c > 8$. This is relevant for future applications at the Forward Physics Facility at CERN [11] which will be able to clarify multiple aspects of QCD in such an extended forward region, in coordination with the HL-LHC and EIC. In Figure 1 (right), we compare theory predictions for the rapidity distribution of B^{\pm} meson production at NLO in QCD with measurements [28] from the LHCb collaboration at $\sqrt{S}=13$ TeV. Theory uncertainties at NLO are large (O(50%)) and mainly ascribed to scale variation. This can be improved by including higher-order corrections which require an extension of the S-ACOT-MPS scheme to NNLO.

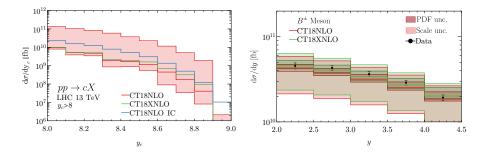


Figure 1. Left: Rapidity distributions of prompt charm at the LHC 13 TeV in the very forward region (yc > 8) [11]. The error band represents the CT18NLO induced PDF uncertainty at 68% CL. **Right**: NLO theory predictions for the rapidity distributions obtained with CT18NLO and CT18XNLO PDFs compared to B^{\pm} production data [28] from LHCb 13 TeV.

3 Impact on PDFs from high-precision $t\bar{t}$ production measurements

Top-quark pair production plays a significant role in setting constraints on PDFs, in particular the gluon, at intermediate to large x. In global PDF analyses, the impact on PDFs from differential cross section measurements of $t\bar{t}$ production at the LHC complements that from jet production measurements: $t\bar{t}$ and jet production overlap in the Q-x plane, but matrix elements and phase-space suppression are different. Therefore, constraints on the gluon may be placed at different values of x. On the other hand, challenges in the estimate of full impact from $t\bar{t}$ production on PDFs arise in presence of incompatibility of the $t\bar{t}$ measurements with other data sets in the baseline (see related discussion in [1]). More realistic estimates account for multiple PDF functional forms and some disagreements between the measurements. In this section, we explore the impact of high-precision $t\bar{t}$ measurements on proton PDFs, and illustrate preliminary results of a global QCD analysis in which $t\bar{t}$ single differential cross section measurements at $\sqrt{S} = 13$ TeV from the ATLAS [29] and CMS [30] collaborations are added on top of the CT18 data baseline. The measurements from ATLAS are obtained using events in the all-hadronic channel with 36.1 fb⁻¹ of integrated luminosity (IL), while those from CMS are obtained with events containing two leptons with 35.9 fb⁻¹ of IL. We consider the full phase space absolute measurements reported in Table 3, in which we specify the type of correlated systematic uncertainties published in the experimental papers. We added these

| data type | N_{pt} | Exp | Corr Sys. |
|---------------------------|----------|-------|------------------|
| $d\sigma/dm_{t\bar{t}}$ | 9 | ATLAS | nuisance par. |
| $d\sigma/dy_{tar{t}}$ | 12 | ATLAS | nuisance par. |
| $d\sigma/dH_{T,t\bar{t}}$ | 11 | ATLAS | nuisance par. |
| $d\sigma/dp_{T,t1}$ | 10 | ATLAS | nuisance par. |
| $d\sigma/dp_{T,t2}$ | 8 | ATLAS | nuisance par. |
| $d\sigma/dm_{t\bar{t}}$ | 7 | CMS | Covariance matr. |
| $d\sigma/dp_{T,t}$ | 6 | CMS | Covariance matr. |
| $d\sigma/dy_{tar{t}}$ | 10 | CMS | Covariance matr. |
| $d\sigma/dy_t$ | 10 | CMS | Covariance matr. |

Table 1. Details of the single differential distributions for $t\bar{t}$ production considered in our analysis.

measurements one by one to the baseline because statistical correlations were not published for all data sets. The measurements from ATLAS and CMS are delivered by using a different binning for the same distribution. The theory predictions for the CMS distributions are obtained by using the FastNNLO grids [31], that are based on the calculation published in [32]. The theory predictions for the ATLAS distributions have been obtained in two steps. We have first generated QCD theory predictions at NLO by using in-house APPLGrid fast tables [33] from the public MCFM computer code [34, 35]. Then, we calculated bin-by-bin NNLO/NLO K-factors using the computer code MATRIX [36] that is based on the theory predictions calculated in [37]. The mass of the top quark, in the pole mass approximation, has been set as $m_r^{(pole)} = 172.5$ GeV. The factorization and renormalization scales μ_F and μ_R respectively, have been chosen according to [38] where for the $m_{t\bar{t}}$, $p_{T,t\bar{t}}$, $y_{t\bar{t}}$, and y_t distributions one uses $\mu_F = \mu_R = H_T/4$ and $H_T = \sqrt{m_t^2 + p_{T,t}^2} + \sqrt{m_t^2 + p_{T,\bar{t}}^2}$, for the $p_{T,t}$ distribution one uses $M_T = \sqrt{m_t^2 + p_{T,t}^2}$, and for the $p_{T,tavg}$ one uses $\mu_F = \mu_R = M_T/2$. EW corrections [39] are also included in this analysis, but their impact on the global fit is negligible. In Figure 2, we illustrate the pulls on the gluon PDF resulting from a global fit in which the y_{tt} distributions from ATLAS and CMS are added on top of the CT18 baseline. The pulls from these distributions from both experiments are in the same direction, that is, each of them prefers a softer gluon at large x ($x \ge 0.2$). The individual impact of these measurements on the CT18 gluon is approximately the same as χ^2/N_{pt} for both distributions is of order 1. A more extensive analysis with complete results will be presented elsewhere [40].

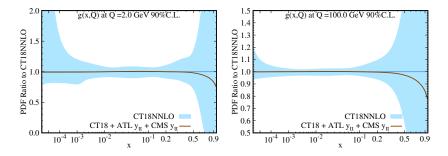


Figure 2. Left: PDF ratio to CT18NNLO at Q = 2 GeV for the gluon PDF. **Right**: same as left, but at Q = 100 GeV. The blue band represents the CT18NNLO PDF uncertainty at 90% CL.

4 CT18 and the new charm and bottom combination at HERA

To finalize our discussion, in this section we discuss the impact of the most recent c- and b-quark production combined measurements of semi-inclusive DIS at HERA from the H1 and ZEUS collaborations [41], on the CT18 gluon PDF. In particular, we illustrate the results of various fits in which these measurements replace their previous version [42, 43]. The new measurements exhibit an extended kinematic range of photon virtuality 2.5 GeV² \leq $Q^2 \le 2000 \text{ GeV}^2$ and Bjorken scaling variable $3 \cdot 10^{-5} \le x_{\text{Bi}} \le 5 \cdot 10^{-2}$, and have reduced uncertainties due to a simultaneous combination of c and b cross-section measurements with reduced correlations between them. We find that the new c and b combination at HERA [41] cannot be well described in the CT18 global analysis, and χ^2/N_{pt} is never below 1.7. For the CT18NNLO fit, we obtained $\chi^2/N_{pt} = 1.98$ for c production $(N_{pt} = 47)$, and $\chi^2/N_{pt} = 1.25$ for bottom production ($N_{pt} = 26$). For the CT18XNNLO fit, instead we obtained $\chi^2/N_{pt} =$ 1.71 for c and 1.26 for b production. We observed tensions between these new combined data and several CT18 datasets such as the LHCb 7 and 8 TeV W/Z production data [44, 45], Z-rapidity data [46] at CDF run-II, CMS 8 TeV single inclusive jet production [47], and $t\bar{t}$ double differential p_T and y cross sections [48]. For this reason, these data were not included in the CT18 global analysis [1]. However, these measurements are very important in PDF determinations because they impose direct constraints on the gluon PDF at intermediate and small x. Therefore, we have re-analyzed these data by performing a series of attempts in which we tried to vary several parameters and settings in the fits. For example, we performed fits in which we varied the weight assigned to these data. We explored the impact of consistent initial-scale Q_0 variations, and tried different parametrizations for the gluon. We tried to vary the value of the c-quark mass m_c in both the pole and MS definitions, and we also analyzed the impact of scale variations in which the factorization scale is defined as $\mu_{DIS}(x)$ = $A\sqrt{m_O^2+B^2/x^C}$ according to saturation models [49, 50]. This is expected to have similar impact to low-x resummation [51]. In addition, we explored phase-space suppression in fits in which we varied the S-ACOT- χ rescaling parameter $\chi = \zeta(1 + \zeta^{\lambda} m_O^2/Q^2)$. A more detailed description of this study can be found in [52]. Here we only provide the results relative to two of the exercises that are described above. In Figure 3, we illustrate the impact on the CT18XNNLO gluon when a scan over the MS $\overline{m}_c(m_c)$ is performed (left panel), and the impact from scale variation where parameter B is varied in the $\mu_{DIS}(x)$ scale definition, while A = 0.5 and C = 0.33(right panel). The CT18X fit is a variant of CT18 where $\mu_{F,DIS} = \mu_{DIS}(x)$. Our preliminary findings indicate that the new c/b production measurements at HERA seem to prefer a harder gluon at intermediate and small x.

5 Conclusions

In this work we have studied HQ production at hadron colliders in recent global QCD analyses and analyzed their impact in collinear proton PDF determinations. We discussed S-ACOT-MPS, a new GMVFN scheme which we applied to pp collisions to describe c/b production at central and forward rapidities. In the near future, it will be technically possible to generate predictions within the S-ACOT-MPS scheme at NNLO with suitable K-factors (NNLO/NLO) at hand. Moreover, it is easy to extend S-ACOT-MPS to other HQ production processes. We explored the impact of $t\bar{t}$ single differential cross section measurements at the LHC at \sqrt{S} =13 TeV on the CT18 PDFs. The impact of $t\bar{t}$ production at the LHC at \sqrt{S} =13 TeV complements that of jet data on the gluon PDF, in particular in the large-x region. $t\bar{t}$ and jets overlap in the Q-x plane, but being matrix elements and phase-space suppression different, constraints on the gluon PDF may be placed at different values of x. Overall, the

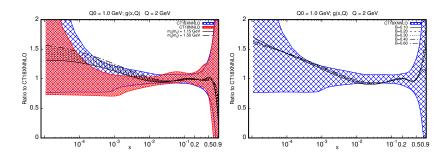


Figure 3. Ratio to the CT18XNNLO gluon PDF as a function of x at Q = 2 GeV. **Left**: scan over the c-quark mass $\overline{m}_c(m_c)$ while $\overline{m}_b(m_b) = 4.18$ GeV, and with $Q_0 = 1$ GeV. **Right**: scan over the B parameter in $\mu_{DIS}(x)$ with A = 0.5, C = 0.33, and $Q_0 = 1$ GeV. Error bands are shown at 90% confidence level for CT18NNLO (red) and CT18XNNLO (blue).

impact is found to be mild with a preference of softer gluon at large x. However, this may change when $t\bar{t}$ production in the lepton + jet channel at 13 TeV is included. We analyzed the most recent c/b combination at HERA. We find that these measurements are challenging and deserve more attention as they are very important for small-x dynamics. HQ production remains a critical process to constrain correlations between m_Q , α_s , and gluon-PDF.

Acknowledgments. The work of M.G. is supported by the NSF under Grant No. PHY-2112025. The work of K.X. at is supported in part by the U.S. DOE under Grant No. DE-SC0007914, the NSF under Grant No. PHY-1820760, and in part by PITT PACC. The work of P.N. is partially supported by the U.S. DOE under Grant No. DE-SC0010129. The work of C.P.Y. is partially supported by the NSF under Grant No. PHY-2013791. He is also grateful for the support from the Wu-Ki Tung endowed chair in particle physics.

References

- [1] T.J. Hou et al., Phys. Rev. D 103, 014013 (2021), 1912.10053
- [2] S. Bailey, T. Cridge, L.A. Harland-Lang, A.D. Martin, R.S. Thorne, Eur. Phys. J. C 81, 341 (2021), 2012.04684
- [3] R.D. Ball et al. (NNPDF), Eur. Phys. J. C 82, 428 (2022), 2109.02653
- [4] S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Phys. Rev. D 96, 014011 (2017), 1701.05838
- [5] A. Accardi et al., Eur. Phys. J. A 52, 268 (2016), 1212.1701
- [6] E.C. Aschenauer, S. Fazio, J.H. Lee, H. Mantysaari, B.S. Page, B. Schenke, T. Ullrich, R. Venugopalan, P. Zurita, Rept. Prog. Phys. 82, 024301 (2019), 1708.01527
- [7] R. Abdul Khalek et al., Nucl. Phys. A 1026, 122447 (2022), 2103.05419
- [8] A. Abada et al. (FCC), Eur. Phys. J. C 79, 474 (2019)
- [9] A. Abada et al. (FCC), Eur. Phys. J. ST **228**, 261 (2019)
- [10] A. Abada et al. (FCC), Eur. Phys. J. ST 228, 755 (2019)
- [11] L.A. Anchordoqui et al., Phys. Rept. 968, 1 (2022), 2109.10905
- [12] J.L. Feng et al. (2022), 2203.05090
- [13] J. Tang et al. (2015), 1507.03224
- [14] M.A.G. Aivazis, F.I. Olness, W.K. Tung, Phys. Rev. D **50**, 3085 (1994), hep-ph/9312318

- [15] M.A.G. Aivazis, J.C. Collins, F.I. Olness, W.K. Tung, Phys. Rev. D 50, 3102 (1994), hep-ph/9312319
- [16] M. Buza, Y. Matiounine, J. Smith, W.L. van Neerven, Eur. Phys. J. C 1, 301 (1998), hep-ph/9612398
- [17] R.S. Thorne, R.G. Roberts, Phys. Rev. D 57, 6871 (1998), hep-ph/9709442
- [18] M. Krämer, F.I. Olness, D.E. Soper, Phys. Rev. D 62, 096007 (2000), hep-ph/0003035
- [19] W.K. Tung, S. Kretzer, C. Schmidt, J. Phys. G 28, 983 (2002), hep-ph/0110247
- [20] S.Alekhin, J.Blumlein, S.Klein, S.Moch, Phys. Rev. D 81, 014032 (2010), 0908.2766
- [21] S. Forte, E. Laenen, P. Nason, J. Rojo, Nucl. Phys. B 834, 116 (2010), 1001.2312
- [22] M.Guzzi, P. Nadolsky, H.L. Lai, C.P. Yuan, Phys.Rev. D 86, 053005 (2012), 1108.5112
- [23] K. Xie, Ph.D. thesis, Southern Methodist U. (2019)
- [24] K. Xie, J.M. Campbell, P.M. Nadolsky, SciPost Phys. Proc. 8, 084 (2022), 2108.03741
- [25] I. Helenius, H. Paukkunen, JHEP 05, 196 (2018), 1804.03557
- [26] B.A. Kniehl, G. Kramer, I. Schienbein, H. Spiesberger, Phys. Rev. D 71, 014018 (2005), hep-ph/0410289
- [27] B.A. Kniehl, G. Kramer, I. Schienbein, H. Spiesberger, Eur. Phys. J. C 41, 199 (2005), hep-ph/0502194
- [28] R. Aaij et al. (LHCb), JHEP 12, 026 (2017), 1710.04921
- [29] G. Aad et al. (ATLAS), JHEP **01**, 033 (2021), **2006.09274**
- [30] A.M. Sirunyan et al. (CMS), JHEP **02**, 149 (2019), 1811.**0**6625
- [31] M. Czakon, D. Heymes, A. Mitov (2017), 1704.08551
- [32] M. Czakon, D. Heymes, A. Mitov, Phys. Rev. Lett. 116, 082003 (2016), 1511.00549
- [33] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G.P. Salam, F. Siegert, P. Starovoitov, M. Sutton, Eur. Phys. J. C 66, 503 (2010), 0911.2985
- [34] J.M. Campbell, R.K. Ellis, J. Phys. G 42, 015005 (2015), 1204.1513
- [35] J.M. Campbell, R.K. Ellis, W.T. Giele, Eur. Phys. J. C 75, 246 (2015), 1503.06182
- [36] M. Grazzini, S. Kallweit, M. Wiesemann, Eur. Phys. J. C 78, 537 (2018), 1711.06631
- [37] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, J. Mazzitelli, JHEP 07, 100 (2019), 1906.06535
- [38] M. Czakon, D. Heymes, A. Mitov, JHEP 04, 071 (2017), 1606.03350
- [39] M. Czakon, A. Ferroglia, D. Heymes, A. Mitov, B.D. Pecjak, D.J. Scott, X. Wang, L.L. Yang, JHEP 05, 149 (2018), 1803.07623
- [40] A. Ablat et al., In preparation (2022)
- [41] H. Abramowicz et al. (H1, ZEUS), Eur. Phys. J. C 78, 473 (2018), 1804.01019
- [42] A. Aktas et al. (H1), Eur. Phys. J. C 40, 349 (2005), hep-ex/0411046
- [43] H. Abramowicz et al. (H1, ZEUS), Eur. Phys. J. C 73, 2311 (2013), 1211.1182
- [44] R. Aaij et al. (LHCb), JHEP 08, 039 (2015), 1505.07024
- [45] R. Aaij et al. (LHCb), JHEP 05, 109 (2015), 1503.00963
- [46] T.A. Aaltonen et al. (CDF), Phys. Lett. B 692, 232 (2010), 0908.3914
- [47] V. Khachatryan et al. (CMS), JHEP 03, 156 (2017), 1609.05331
- [48] A.M. Sirunyan et al. (CMS), Eur. Phys. J. C 77, 459 (2017), 1703.01630
- [49] K.J. Golec-Biernat, M. Wusthoff, Phys. Rev. D 59, 014017 (1998), hep-ph/9807513
- [50] F. Caola, S. Forte, J. Rojo, Phys. Lett. B 686, 127 (2010), 0910.3143
- [51] R.D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, Eur. Phys. J. C 78, 321 (2018), 1710.05935
- [52] M. Guzzi, P. Nadolsky, K. Xie, SciPost Phys. Proc. 8, 164 (2022), 2108.01791