Charm and bottom hadroproduction with heavy-quark mass renormalized in the $\overline{\text{MS}}$ scheme

........,^{a}........^{b}

^{a}*II. Institute for Theoretical Physics, Hamburg University*
  Luruper Chaussee 149, D–22761 Hamburg, Germany
^{b}*....*

  E-mail: x.y@desy.de

ABSTRACT: We present predictions for charm (bottom) production at hadron colliders making use of the $\overline{\text{MS}}$ renormalization scheme for the charm (bottom) mass, as an alternative to the widely used on-shell renormalization scheme. We compare these two possibilities, by showing single and double differential distributions including NLO QCD corrections complemented by the effects of phenomenological fragmentation functions, and total cross-sections at approximate NNLO accuracy. We briefly sketch possible interesting applications. This study is motivated by the observation that precise determinations of the value of the charm and bottom masses in the $\overline{\text{MS}}$ scheme are available, while the perturbative expansion for the conversion of the charm and bottom mass values from the $\overline{\text{MS}}$ scheme to the on-shell scheme is not converging, due to the smallness of the scales involved ($\sim$ very few GeV), not so far from the value of Λ_{QCD}, and to infrared renormalon singularities. In case of charm, a lack of convergence is already visible up to the 4-loop order, i.e. the accuracy at which the conversion to the pole mass value has been computed so far. We also provide public numerical routines, which can be used in the MCFM framework.

Contents

1 Introduction

Charm-anticharm and bottom-antibottom pair production are among the most frequent inelastic processes occurring in hadronic collisions at modern high-energy colliders, with cross-sections smaller than for the dijet case but well above those for the top-antitop case, as follows from the values of the masses of the various final-state partons, the available phase-space, and the structure of the Standard Model (SM) Lagrangian.

Collider experiments are able to detect the products of the hadronization / fragmentation of intermediate-mass quarks, i.e. heavy mesons and baryons, as well as the jet associated to them, i.e. b-jets and c-jets. B-hadrons and D-hadrons are reconstructed by their decay products, with decay vertices displaced with respect to the primary vertices. This operation requires good tracking, vertexing and particle identification capabilities. The measurements are indeed easier to perform in case of B-hadrons than for D-hadrons, because the proper lifetimes of the first ones are longer than those of the latter. In the reconstruction of the transverse momentum p_T of the decaying hadrons one has to take into account that at very low p_T, decay tracks can be affected by multiple scattering and that trigger efficiencies play a role in determining the minimum p_T of the decay products. On the other hand, the angle between the decay tracks become smaller as the p_T of the decaying object increases, challenging the reconstruction process at large p_T, made even more difficult by the lower statistics. At large p_T, b-jet measurements, which not require a fully-exclusive reconstruction of the properties of the heavy decaying hadrons and do not distinguish

between different B-hadron species, are still performed, made possible by the level of sophistication of b-jet tagging algorithms, that is rapidly increasing also thank to deep learning techniques. On the other hand, c-jet tagging is more difficult, but new promising methods and analyses adopting them are under development.

On the theory side, analytical formulas at the NLO accuracy for the cross-sections for the hadroproduction of massive quark pairs, as well as for one-massive-quark inclusive production, are known since many years. The resummation of logarithms of (p_T^2/m^2), relevant when the transverse momentum of the heavy-quark p_T is much larger than its mass m, has been performed up to the next-to-leading-logarithmic (NLL) accuracy. The resummation of Sudakov logarithms related to radiation with $p_T \to 0$, as well as the one of threshold logarithms and high-energy logarithms, have also been worked out (up to various degrees of accuracy) and presented in a number of papers. The transition from quarks to hadrons is described either by Fragmentation Functions, or by matching NLO matrix-elements to parton shower and hadronization approaches. Differential cross-sections are known at NNLO accuracy only for top-antitop hadroproduction, eventually complemented by the resummation of threshold logarithms at the netx-to-NLL (NNLL) accuracy, but not yet for the charm and bottom cases. In case of the latter, however, approximate total NNLO cross-sections are also available.

One of the inputs of all aforementioned calculations are the values of the heavy-quark masses. The SM is not able, by itself, to predict the values of the quark masses, which are thus fundamental parameters in this framework. In general, different renormalization schemes can be used for relating the bare masses to the renormalized ones, and predictions for physical observables on the basis of a truncated pQCD series, relying on renormalized quantities, are scheme dependent. It follows that even the perturbative convergence can vary from scheme to scheme.

In this paper we describe the use of the $\overline{\text{MS}}$ scheme, as an alternative to the most widely used on-shell scheme, for the renormalization of the charm (bottom) masses in charm (bottom) production at hadron colliders. We investigate the perturbative convergence in both schemes, by providing comparisons between physical quantities calculated at various levels of accuracy, and we sketch interesting applications within and beyond collider physics.

The implementation of the calculation of single and double differential distributions, including NLO QCD corrections and using the $\overline{\text{MS}}$ scheme for renormalizing the charm (bottom) mass is described in Section 2. The details of the corresponding numerical routines released as an addendum to this paper are collected in the Appendix. Predictions obtained using this implementation are presented in Section 3, whereas predictions for total cross-sections, up to approximate NNLO accuracy, are presented in Section 4, together with considerations on the convergence of the perturbative expansion in the strong coupling constant. Examples of possible applications are briefly discussed in Section 5. Finally, our conclusions are summarized in Sec-

2   Implementation

1)...formulas....

In this work light quarks are assumed to be massless. On the other hand, for the heavy-quark masses, different renormalization schemes can be adopted. In the following we focus on the cases of the $\overline{\text{MS}}$ and on-shell schemes and we explicitly recall the conversion relation between the values of the renormalized masses in these two schemes. Other mass choices would also be possible, e.g. the potential subtractive mass was suggested as a possibility to produce an improved perturbative convergence for heavy quark production near threshold (i.e. at energies slightly above $2m_Q$) in Ref. [1] and the 1S mass have been introduced in Ref. [2] as a way of stabilizing the position of the peak of the vector-current-induced total cross section of $t\bar{t}$ production in e^+e^- collisions, as a function of $\sqrt{s}$, for $\sqrt{s} \sim 2\ m_t$. Limited to the case of top quarks, the top quark jet mass [3], was introduced in the framework of Soft Collinear Effective Theory.

In the $\overline{\text{MS}}$ scheme, the renormalized mass of the heavy-quark vary with the renormalization scale μ_R. The evolution follows the renormalization group equation (RGE)

$$\mu_R^2 \frac{dm(\mu_R)}{d\mu_R^2} = -\gamma(\alpha_S(\mu_R))m(\mu_R) \tag{2.1}$$

where $\gamma(\alpha_S(\mu_R)) \equiv \sum_{r=1}^{\infty} \gamma_r(\alpha_S(\mu_R)/\pi)^r$ is the anomalous dimension, known up to four-loop order in perturbation theory, depending on the coupling constant α_S, also renormalized in the same scheme. Explicit formulas for the γ_1, γ_2, γ_3 and γ_4 coefficients, depending on the number of active light quark flavours at the scale μ_R, are given in the literature.

Precise determinations of the $\overline{\text{MS}}$ masses of heavy-quarks are available from a number of studies.

In the on-shell scheme, the renormalized mass coincides with the pole in the Fourier transform of the Feynman propagator of the renormalized quark field and is thus the same at all scales. Considering its infrared finiteness to all orders in perturbation theory, this is the natural mass parameter for the description of heavy-quark systems at scales governed by the mass of the heavy quark. However, this definition is based on the concept of quarks as asymptotic states, which can rise a number of issues. One of them is that the pole mass is not an observable, because of confinement.

The on-shell (also called pole) mass M can be obtained from the $\overline{\text{MS}}$ running mass m at a renormalization scale μ_R, in terms of an perturbative expansion in the

strong coupling constant α_S:

$$M = m(\mu_R)\left(1 + c_1\left(\frac{\alpha_S}{\pi}\right) + c_2\left(\frac{\alpha_S}{\pi}\right)^2 + c_3\left(\frac{\alpha_S}{\pi}\right)^3 + \mathcal{O}(\alpha_S^3)\right) \qquad (2.2)$$

with the coefficients c_1, c_2, c_3 given by

$$c_1 = C_F\left(1 + \frac{3}{4}L\right) \qquad (2.3)$$

$$c_2 = ... \qquad (2.4)$$

$$c_3 = ... \qquad (2.5)$$

where $C_F = 4/3$ is one of the SU(3) color factors, $L \equiv \ln(\mu_R^2/m(\mu_R)^2)$, and α_S is evaluated at the $m(\mu_R)$ scale and renormalized in the $\overline{\text{MS}}$ scheme in the theory with a number of active flavours equal to $n_f = n_{lf} + 1$ at and above the heavy-quark threshold scale, assumed to be equal to its mass. The number of light flavours is $n_{lf} = 3$ (4) for charm (bottom) production, respectively.

Considering the particular case of $m(m)$, i.e. the $\overline{\text{MS}}$ mass renormalized at the scale $\mu_R = m(\mu_R)$, $m(m) \equiv m(\mu_R = m(\mu_R))$, the logarithmic terms L cancel and, evaluating eq. 2.2 numerically, one gets [4]:

$$M = m(m)\left(1 + 1.333\left(\frac{\alpha_S}{\pi}\right) + (13.44 - 1.041 n_{lf})\left(\frac{\alpha_S}{\pi}\right)^2 + \right.$$
$$\left. + \left(190.595 - 27.0 n_{lf} + 0.653 n_{lf}^2\right)\left(\frac{\alpha_S}{\pi}\right)^3 + \mathcal{O}(\alpha_S^3)\right) \qquad (2.6)$$

On the other hand, inverting this relation and making a Taylor expansion, one gets $m(m)$ as a function of the pole mass

$$m(m) = M\left(1 - 1.333\left(\frac{\alpha_S}{\pi}\right) + (-11.67 + 1.041 n_{lf})\left(\frac{\alpha_S}{\pi}\right)^2 + \right.$$
$$\left. + \left(-170.0 + 24.8 n_{lf} - 0.653 n_{lf}^2\right)\left(\frac{\alpha_S}{\pi}\right)^3 + \mathcal{O}(\alpha_S^3)\right) \qquad (2.7)$$

When using eq. 2.7, we evaluate α_S in the $\overline{\text{MS}}$ scheme in the theory with n_f active flavours at the scale $\mu_R = M$.

While α_S in eq. 2.2, 2.6, 2.7 is renormalized in the $\overline{\text{MS}}$ scheme, the matrix-elements, as well as the parton distribution functions (PDF) and the associated α_S evolution used in the fixed-order massive calculations presented in this paper are all defined in the decoupling scheme, with a fixed number of light flavours $n_{lf} = 3$ for charm production and $n_{lf} = 4$ for bottom production at all scales. Using the decoupling relations, it is possible to express the PDFs, α_S and the partonic cross-sections in the $\overline{\text{MS}}$ scheme in terms of those in the decoupling scheme, once fixed a

scale μ. Alternatively, it is also possible to re-express eq. 2.6 in terms of α_S in the decoupling scheme, obtaining

$$M = ... \tag{2.8}$$

In practice, although terms for the expansion in eq. 2.2, 2.6, 2.7 are known up to 4-loops [5], we truncate it at two-loops (order α_S^2) for computing the NNLO cross-sections and at one-loops (order α_S) for computing NLO cross-sections, respectively.

In this work, the α_S evolution as a function of μ_R and the corresponding α_S values entering in eq. 2.2, 2.6 and 2.7 and other parts of the fixed-order computation are evaluated retaining three-loops for producing NNLO cross-sections and two-loops for producing the NLO ones, respectively, although other choices would indeed be possible, as follows from the fact that solving RGE equations, relating running quantities at different μ_R scales, is an all-loop operation.

The presence of infrared renormalon singularities in the relation of conversion from the pole mass to the $\overline{MS}$ mass manifests in practice as factorially growing terms in the expansion in eq. 2.2, that does not converge. Treating the series as an asymptotic expansion, the ambiguity in its resummation, affecting the value of the pole mass, is of the order of $\mathcal{O}(\Lambda_{QCD})$.

While the pole mass is plagued by the renormalon ambiguity uncertainty, the short-distance mass definitions are renormalon-free.

2) may be add a plot with comparison between m(μ_R) as a function of μ_R, and M (constant) as a function of μ_R ?

3)...MNR and MCFM....

4) some example of cross-check between the two codes (MCFM, MNR) and a pair of figures showing they give the same results

3   Predictions for differential cross-sections

1) Comparisons between LO and NLO for a number of differential distributions (i.e. all possible kinds of $d\sigma/dX$ implemented in the MNR/MCFM program, with no cuts at all, and with different sets of cuts, for both charm and bottom. Also add the case where $X = m_{Q\bar{Q}}$ considering that this is interesting for some ongoing experimental analyses (see CMS) for extracting the running mass of the top quark (and a similar methodology may be can be extended even to the extraction of charm and bottom

masses ?). The reason of studying also LO, even if, in case of charm, I agree it has gigantic uncertainty bands, is that we want to have idea of the convergence of the cross-sections when using running mass with respect to the case of using pole masses. To do this, we would ideally better compare NLO and NNLO, however, considering that we do not have any NNLO differential distributions, the only thing we can do for now for studying convergence of differential distribution is comparing NLO with LO. I think one could put in the paper some differential K-factor, making ratios of differential distributions at NLO and LO, to see if K-factors using running masses are a bit smaller than those using the pole masses or not and give some quantitative number to the people. One could compare the perturbative convergence in case of charm and in case of bottom. The LO predictions One could put just in a few plots (e.g. when drawing distributions without cuts or with purely theoretical cuts), not necessarily in the plots with the experimental cuts and explicit comparison with experimental data. In fact, I am thinking to distributions of charm and bottom quarks, not even D-mesons or B-mesons.

2) Check if, using stronger systems of cuts, the implementation in the running mass scheme behaves better than the other or if the behaviour of scale uncertainties remains similar to what we have already observed for the LHCb cuts (i.e. no strong changes).

Among other things, I would look, maybe, at large p_T, again for case of charm and bottom quarks (and not the mesons). I have not a specific suggestion for a system of cuts, but, in general, the more exclusive the system of cuts are, the most I would expect distributions to change between two different implementations.

3) Possible double differential distributions $\frac{d^2\sigma}{dp_T dy}$, $\frac{d^2\sigma}{dm_{Q\bar{Q}} dy}$ and importance for comparison with experimental measurements. One can make reference to recent experimental studies of top quarks by CMS.

4) One can try to use different scales. One could e.g. also try to use Czakon/Mitov suggested scales for top, and show how they work (or do not work) for bottom and charm.

4   Predictions for total cross-sections

1) Comparison of LO, NLO, NNLO cross-sections using Hathor with on-shell vs $\overline{\text{MS}}$ scheme. Even if we have already presented some of these comparisons in our old paper, one could make further updated ones, to see the convergence of the perturbative series, and comparisons using the most modern PDFs.

5 Possible applications

5.1 PDF fits

1) Qualitative discussion on possible outcome from the PROSA fit redone with the modified MCFM and all the rest the same: does it happens something to the fitted charm mass ? And the χ^2 ? And the PDF uncertainty band ? I think this is not a new PROSA paper, which would indeed require to add new data in the fit, etc....but more an exercise, showing if using running mass, can have an impact or not.

One of the possible applications of the calculations for heavy-quark hadroproduction with the $\overline{\text{MS}}$ mass definition could be PDF fits which use heavy-flavour measurements from the LHC to constrain the gluon distribution at low values of x [6–8]. Because of the large scale dependence of the NLO calculations for charm hadroproduction, in such fits it seems feasible to include only ratios of cross sections, which are constructed using measurements at different values of rapidity and/or transverse momentum, or at different centre-of-mass energies. As an example, in the PROSA analysis [6] ratios of charm and bottom hadron production cross sections [9, 10] as a function of rapidity to the cross section in the $3 < y < 3.5$ interval for each p_T bin were used, together with the inclusive DIS data from HERA [11–13]. These ratios afforded a reduced scale dependence, but at the same time they have reduced sensitivity to the heavy-quark mass. We repeat the PROSA analysis using the $\overline{\text{MS}}$ heavy-quark masses both in the calculations of the DIS structure functions and charm and bottom hadroproduction cross sections, while all other settings are as in Ref. [6], and observe only small impact on χ^2 and fitted PDFs, which are well within the uncertainties. Even these small differences are driven mainly by the change in the predictions for the HERA data, because the LHCb data in the format of the normalised cross sections do not provide notable sensitivity neither to the heavy-quark mass scheme, not to the value of the heavy-quark mass. As a result, the fitted $\overline{\text{MS}}$ masses come out compatible with the results obtained in Refs. [12–15] where the HERA data alone were analysed to determine the heavy-quark $\overline{\text{MS}}$ masses.

2) Inclusion of measurements of total cross-sections for charm hadroproduction in (future) NLO and NNLO PDF fits: discussion at qualitative level.

The NLO predictions for differential charm hadroproduction at the LHC have very large ($> 100\%$ in some regions of the phase space) scale uncertainties. Furthermore most of the PDF fitting groups provide fits at NNLO accuracy. This prevents in-

cluding existing data on charm and bottom hadroproduction in their state-of-the-art fits. In this context measurements of the total charm hadroproduction cross section would be beneficial, because they can be confronted in the fits with NNLO predictions [16–19] which have reduced scale uncertainties. However, no such measurements have been performed to date. On the other hand, ALICE [20, 21], ATLAS [22] and LHCb [9, 23, 24] experiments provided measurements of charm production in different kinematic regions which cover more a larger part of the phase space. One can reliably extract the total cross section by extrapolating these measurements to the full phase space. Below we perform such extrapolation, provide values for the total charm cross section at different center-of-mass energies and their ratios together with experimental and theoretical uncertainties arising from the extrapolation procedure, compare to theoretical predictions at NNLO obtained which are obtained using different PDF sets, and demonstrate how these data can help reducing PDF uncertainties.

The existing most precise LHC measurements of charm production are summarised in Table 1. The ALICE measurements at 5 and 7 TeV cover the central region $|y| < 0.5$, the LHCb measurements at 5, 7 and 13 TeV provide coverage of the forward region $2 < y < 4.5$, and the ATLAS measurement at 7 TeV essentially bridges the gap by providing data at $|\eta| < 2.1$. However, while both ALICE and LHCb provide measurements nearly in the full p_T range starting from 0 GeV, AT-LAS reports the cross sections only at $p_T > 3.5$ GeV, thus leaving the bulk of the corresponding p_T kinematic range unmeasured. Furthermore, it turns out that the most precise data for ALICE and LHCb are available for D^0 production, while this final state was not measured by ATLAS. Given these arguments, at 5 and 7 TeV we extrapolate ALICE and LHCb measurements of D^0 production to the full phase space. In order to maintain least dependence on the theoretical predictions, both ALICE and LHCb measurements are extrapolated to nearby regions of y, namely to $0 < |y| < 1.5$ and $|y| > 1.5$, respectively:

$$\sigma_{\text{total}} = \sigma_{\text{ALICE}} \times K_{\text{ALICE}} + \sigma_{\text{LHCb}} \times K_{\text{LHCb}} \times 2$$

$$K_{\text{ALICE}} = \frac{\sigma^{NLO}_{|y|<1.5}}{\sigma^{\text{NLO}}_{|y|<0.5}}, \qquad K_{\text{LHCb}} = \frac{\sigma^{NLO}_{|y|>1.5}}{\sigma^{\text{NLO}}_{2<|y|<4.5}}. \tag{5.1}$$

Here σ_{ALICE} and σ_{LHCb} are the ALICE and LHCb data, respectively, and σ^{NLO} is the theoretical prediction. The factor 2 in the second term takes into account that the LHCb data are provided for only $2 < y < 4.5$ and need to be extrapolated to $2 < |y| < 4.5$ (we take that the cross sections at $2 < y < 4.5$ and $-4.5 < y < -2$ are equal). Also the measurements are extrapolated into the full range of p_T (not shown in Eq. 5.1 for brevity), which implies only 1% correction for the LHCb data at 7 TeV provided for $p_T < 8$ GeV, and even smaller corrections for the other data sets. This procedure is used to obtain the total charm cross section at 5 and 7 TeV, while at 13

Measurement	Final state	Kinematic region
ALICE 5 TeV [21]	D^0, D^+, D^{*+}, D_s^+	$0 < p_T < 36$ GeV, $\lvert y \rvert < 0.5$
ALICE 7 TeV [21]	D^0, D^+, D^{*+}, D_s^+	$0 < p_T < 36$ GeV, $\lvert y \rvert < 0.5$
ATLAS 7 TeV [22]	D^+, D^{*+}, D_s^+	$3.5 < p_T < 100$ GeV, $\lvert \eta \rvert < 2.1$
LHCb 5 TeV [24]	D^0, D^+, D^{*+}, D_s^+	$0 < p_T < 10$ GeV, $2 < y < 4.5$
LHCb 7 TeV [9]	$D^0, D^+, D^{*+}, D_s^+, \Lambda_c$	$0 < p_T < 8$ GeV, $2 < y < 4.5$
LHCb 13 TeV [23]	D^0, D^+, D^{*+}, D_s^+	$0 < p_T < 15$ GeV, $2 < y < 4.5$

Table 1. Summary of most precise measurements of charm production at the LHC.

TeV we extrapolate solely the LHCb measurement since no other data are available at this energy. We also compute ratios of cross sections at different centre-of-mass energies $R_{7/5} = \sigma_{7\ \text{TeV}}/\sigma_{5\ \text{TeV}}$ and $R_{13/7} = \sigma_{13\ \text{TeV}}/\sigma_{7\ \text{TeV}}$ which benefit from partial cancellation of theoretical uncertainties.

The theoretical prediction σ^{NLO} are computed using the NLO predictions with the $\overline{\text{MS}}$ masses described in the previous sections, supplemented with phenomenological non-pertutbative fragmentation functions to describe the $c \to D^0$ transition. The factorisation and renormalisation scales are chosen to be $\mu_r = \mu_f = \sqrt{4m_c(m_c)^2 + p_T^2}$ and varied by a factor of two up and down (both simultaneously and independently for μ_r and μ_f) to estimate the uncertainties. The $\overline{\text{MS}}$ charm quark mass is set to $m_c(m_c) = 1.275 \pm 0.030$ GeV [25]. The proton is described by the PROSA PDF set [6]. We chose this set because among all FFNS PDF sets which are available today the PROSA set is expected to have the most reliable gluon distribution at low x thanks to the heavy-quark data used for its determination. Furthermore, to estimate the PDF uncertainties, the extrapolation is performed using the ABMP16 [26], CT14 [27], MMHT2014 [28], JR14 [29], NNPDF3.1 [30] and HERAPDF FF3A [31] NLO PDF sets, and the envelope covering the PROSA PDF uncertainties and the difference obtained using any of the additional PDF sets is constructed. This conservative procedure is essential, because the theoretical calculations for highest values of y involve the gluon PDF at very low values of x (up to 4×10^{-8}) where it is not directly covered by data in any of the PDF fits (even in the PROSA fit which includes the charm data up to $y = 4.5$ as measured by LHCb). The fragmentation of charm quarks into D^0 mesons is described by the Kartvelishvili function with $\alpha_K = 4.4 \pm 1.7$ [6], while the fragmentation fraction $f(c \to D^0)$ cancels for the extrapolation factors in Eq. 5.1. All theoretical uncertainties are assumed to be fully correlated for cross sections in different kinematic regions and at different centre-of-mass energies. Robustness of the extrapolation procedure is checked by varying the boundary y between the kinematic regions into which the ALICE and LHCb

measurements are extrapolated by ± 0.5 (at the same time, this variation tests consistency of the ALICE and LHCb data). Furthermore, we did another check of the extrapolation procedure by computing predictions for the ALICE and LHCb data using the POWHEGBOX + PYTHIA6 framework which provide NLO predictions matched to parton showers [Maria could write more details here] and found them to be consistent with our NLO calculations within the theoretical uncertainties.

The results of extrapolation are reported in Table 2. The scale, mass, PDF and fragmentation uncertainties are added in quadrature to obtain the total theoretical uncertainty assigned to the extrapolated results. The experimental uncertainties of the input data are propagated to the extrapolated cross sections and reported separately. The experimental uncertainties of the input data sets are assumed to be fully uncorrelated. The experimental and theoretical extrapolation uncertainties are of the same size. The total uncertainties are obtained by adding the experimental and theoretical uncertainties in quadrature, and amount to $\approx 10\%$.

The extrapolated cross sections and their ratios are compared to NNLO predictions obtained using different PDF sets. Namely, we used the ABMP16 [32], CT14 [27], MMHT2014 [28], JR14 [29], NNPDF3.1 [30] and HERAPDF [31] NNLO PDF sets. The cross sections are computed using the HATHOR program [33] interfaced in xFitter [34]. The factorisation and renormalisation scales are chosen to be $\mu_r = \mu_f = 2m_c(m_c)$ and varied by a factor of two up and down (both simultaneously and independently for μ_r and μ_f) to estimate the uncertainties. The $\overline{\text{MS}}$ charm quark mass is set to $m_c(m_c) = 1.275$ GeV [25]. Figure 1 shows the extrapolated cross sections and their ratios compared to NNLO predictions. For the NNLO predictions, separately the theoretical uncertainty which arise from the scale variations and the PDF uncertainty is shown. All theoretical predictions agree with the data within uncertainties, but noticeably the MMHT2014, HERAPDF2.0 and CT14 PDF sets have uncertainties which are larger than both scale and data extrapolation uncertainties for some of the observables. In particular, for the cross sections at 13 TeV the predictions of the MMHT2014 and HERAPDF2.0 sets are consistent with negative values within uncertainties, while the CT14 predictions has a large positive uncertainty which greatly exceeds our extrapolated cross section. Therefore, these PDF sets could benefit from data on charm production cross sections or their ratios included in the fits. Remarkably, also the scale uncertainties appear to be different when using different PDF sets. Among the considered observables, the most conclusive one is $R_{7/5}$ for which both data extrapolation and theoretical scale uncertainties are moderate ($\approx 10\%$). Our extrapolated value for this observable can be used in future NNLO PDF fits to constrain the gluon PDF at low x. The other ratio $R_{13/7}$ has a larger extrapolation uncertainty suffering from the lack of experimental measurements of charm production in the central rapidity region at 13 TeV.

Observable \Unc. [%]	NLO μ var.	$m_c(m_c) \pm 0.03$ GeV	$\alpha_K \pm 1.7$	PDF	$y_{\text{ALICE,LHCb}} \pm 0.5$	Total th.	Exp.	Total
$\sigma(c\bar{c})_{5\text{TeV}}/\text{mb} = 5.254$	$^{+0.8}_{-0.6}$	$^{-0.1}_{+0.1}$	$^{-2.0}_{+1.1}$	$^{+4.8}_{-1.5}$	$^{-2.0}_{+2.2}$	$^{+5.0}_{-2.5}$	± 4.3	$^{+6.6}_{-5.0}$
$\sigma(c\bar{c})_{7\text{TeV}}/\text{mb} = 6.311$	$^{+0.7}_{-0.6}$	$^{-0.1}_{+0.1}$	$^{-2.0}_{+1.1}$	$^{+7.8}_{-1.9}$	$^{-2.2}_{+2.4}$	$^{+7.9}_{-2.8}$	± 6.5	$^{+10.2}_{-7.1}$
$\sigma(c\bar{c})_{13\text{TeV}}/\text{mb} = 11.298$	$^{+0.7}_{-2.9}$	$^{+0.2}_{-0.2}$	$^{+1.5}_{-0.6}$	$^{+0.0}_{-2.9}$	n/a	$^{+1.6}_{-4.1}$	± 6.1	$^{+6.3}_{-7.3}$
$R_{7/5} = 1.201$	$^{+0.1}_{-0.0}$	$^{+0.0}_{-0.0}$	$^{-0.0}_{+0.0}$	$^{+2.9}_{-0.4}$	n/a	$^{+2.9}_{-0.4}$	± 7.8	$^{+8.3}_{-7.8}$
$R_{13/7} = 1.790$	$^{+1.3}_{-3.5}$	$^{+0.2}_{-0.2}$	$^{+3.6}_{-1.7}$	$^{+1.0}_{-8.5}$	n/a	$^{+3.9}_{-9.3}$	± 8.9	$^{+9.7}_{-12.9}$

Table 2. Extrapolated total charm production cross sections and their ratios at different centre-of-mass energies. The correlation factor between $R_{7/5}$ and $R_{13/7}$ is -0.61.

To demonstrate the feasibility of the provided observables to constrain the PDFs, we employ a profiling technique [35] based on minimizing χ^2 function between data and theoretical predictions taking into account both data and theoretical uncertainties arising from PDF variations. The analysis is performed using the xFitter program [34]. We consider the MMHT2014 PDF set and the ratios $R_{7/5}$ and $R_{13/7}$ which exhibit least scale uncertainties. The correlation of $R_{7/5}$ and $R_{13/7}$ because of the common input 7 TeV data sets is taken into account. The PDF uncertainties are included in χ^2 through nuisance parameters, and the values of these nuisance parameters at the minimum are interpreted as optimised (or profiled) PDFs, while their uncertainties determined using the tolerance criterion of $\Delta\chi^2 = 1$ correspond to the new PDF uncertainties. The original and profiled MMHT2014 gluon PDF is shown in Fig. 2 at the scale $Q^2 = 10$ and 100 GeV^2. The profiled distribution exhibits greatly reduced uncertainties at low x, and in this region the distribution is shifted towards larger values of the gluon density.

5.2 Prompt neutrino fluxes

....Discussion at qualitative level about reduction of uncertainties.

6 Conclusions

Appendix: description of the routines added to MCFM

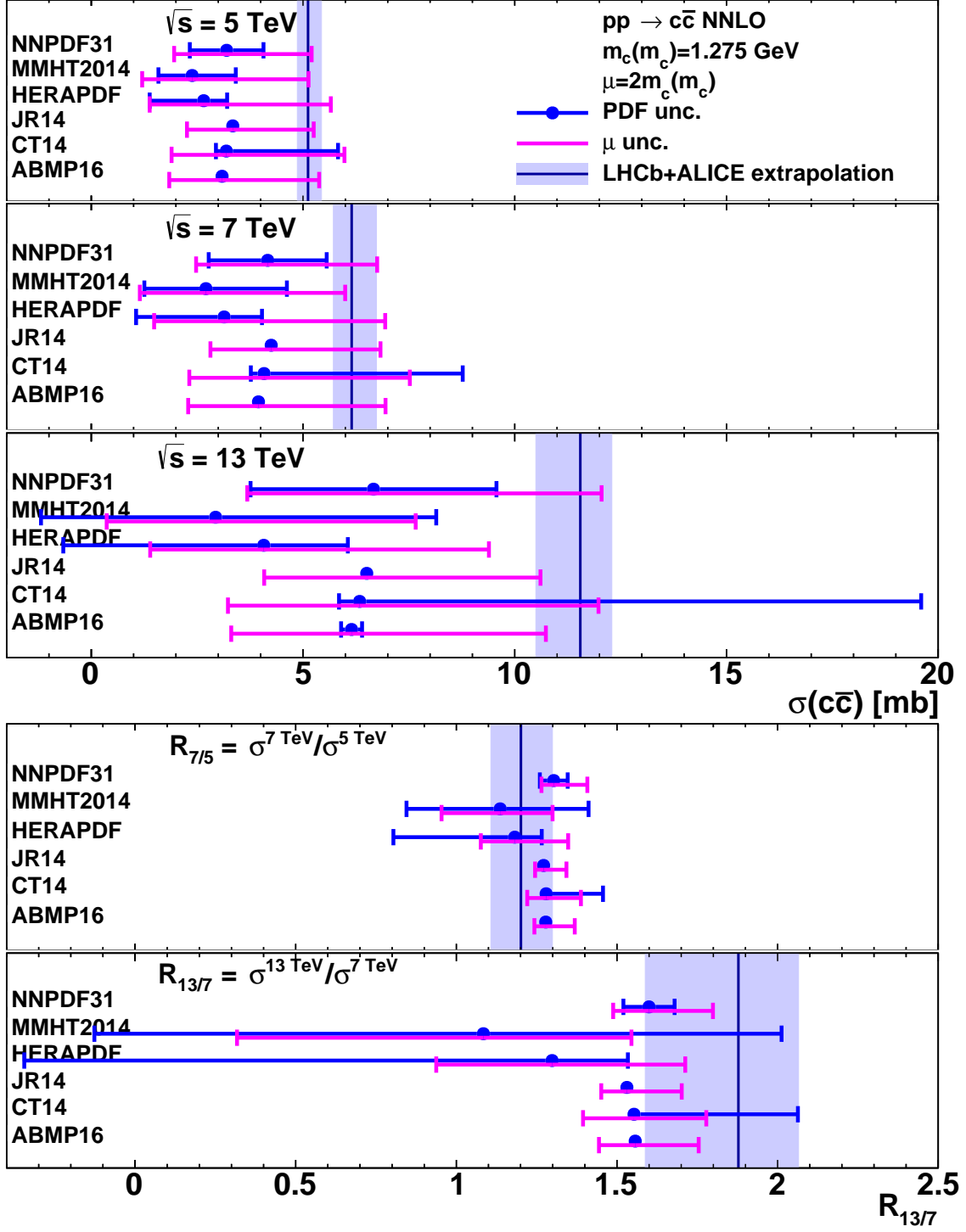...name of routines, what they do, how to compile....

Figure 1. Comparison of the extrapolated total charm production cross sections and their ratios with the NNLO theoretical predictions using different PDF sets.
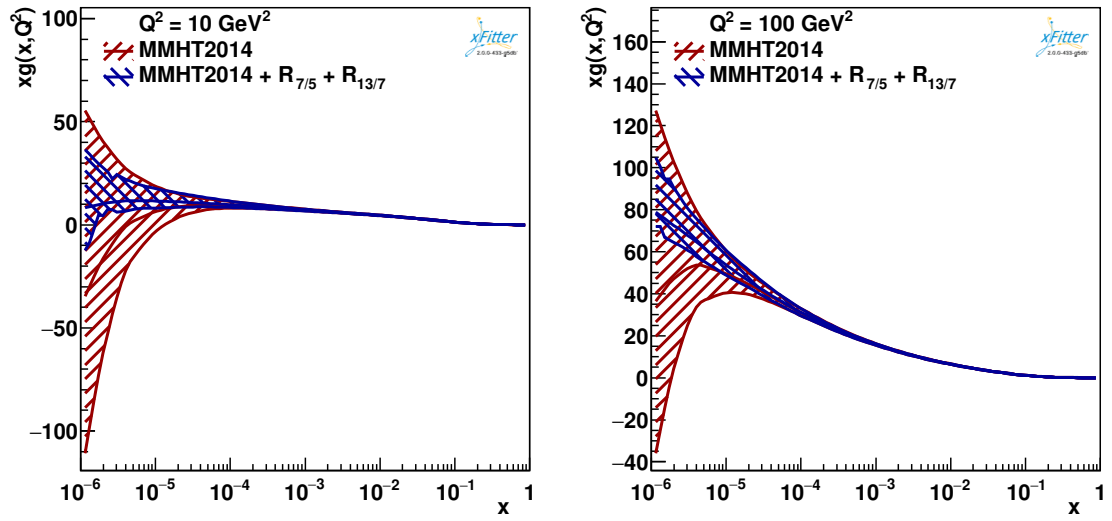
Figure 2. The gluon distribution of the original and profiled MMHT2014 set at $Q^2 = 10$ GeV^2 (left) and $Q^2 = 100$ GeV^2 (right).

References

[1] M. Beneke, *A Quark mass definition adequate for threshold problems*, *Phys. Lett.* **B434** (1998) 115–125, [hep-ph/9804241].

[2] A. H. Hoang and T. Teubner, *Top quark pair production close to threshold: Top mass, width and momentum distribution*, *Phys. Rev.* **D60** (1999) 114027, [hep-ph/9904468].

[3] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, *Jets from massive unstable particles: Top-mass determination*, *Phys. Rev.* **D77** (2008) 074010, [hep-ph/0703207].

[4] K. G. Chetyrkin and M. Steinhauser, *Short distance mass of a heavy quark at order α_s^3*, *Phys. Rev. Lett.* **83** (1999) 4001–4004, [hep-ph/9907509].

[5] P. Marquard, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, *Quark Mass Relations to Four-Loop Order in Perturbative QCD*, *Phys. Rev. Lett.* **114** (2015), no. 14 142002, [arXiv:1502.01030].

[6] **PROSA** Collaboration, O. Zenaiev et al., *Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x*, *Eur. Phys. J.* **C75** (2015), no. 8 396, [arXiv:1503.04581].

[7] R. Gauld, J. Rojo, L. Rottoli, and J. Talbert, *Charm production in the forward region: constraints on the small-x gluon and backgrounds for neutrino astronomy*, *JHEP* **11** (2015) 009, [arXiv:1506.08025].

[8] R. Gauld and J. Rojo, *Precision determination of the small-x gluon from charm production at LHCb*, *Phys. Rev. Lett.* **118** (2017), no. 7 072001, [arXiv:1610.09373].

[9] **LHCb** Collaboration, R. Aaij et al., *Prompt charm production in pp collisions at sqrt(s)=7 TeV*, *Nucl. Phys.* **B871** (2013) 1–20, [arXiv:1302.2864].

[10] **LHCb** Collaboration, R. Aaij et al., *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **08** (2013) 117, [arXiv:1306.3663].

[11] **H1, ZEUS** Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, *JHEP* **01** (2010) 109, [arXiv:0911.0884].

[12] **H1, ZEUS** Collaboration, H. Abramowicz et al., *Combination and QCD Analysis of Charm Production Cross Section Measurements in Deep-Inelastic ep Scattering at HERA*, *Eur. Phys. J.* **C73** (2013), no. 2 2311, [arXiv:1211.1182].

[13] **ZEUS** Collaboration, H. Abramowicz et al., *Measurement of beauty and charm production in deep inelastic scattering at HERA and measurement of the beauty-quark mass*, *JHEP* **09** (2014) 127, [arXiv:1405.6915].

[14] **H1, ZEUS** Collaboration, H. Abramowicz et al., *Combination and QCD analysis of*

charm and beauty production cross-section measurements in deep inelastic ep scattering at HERA, *Eur. Phys. J.* **C78** (2018), no. 6 473, [arXiv:1804.01019].

[15] **xFitter Developers' Team** Collaboration, V. Bertone et al., *A determination of $m_c(m_c)$ from HERA data using a matched heavy-flavor scheme*, *JHEP* **08** (2016) 050, [arXiv:1605.01946].

[16] 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** (2012) 132001, [arXiv:1204.5201].

[17] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **12** (2012) 054, [arXiv:1207.0236].

[18] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **01** (2013) 080, [arXiv:1210.6832].

[19] 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** (2013) 252004, [arXiv:1303.6254].

[20] **ALICE** Collaboration, S. Acharya et al., *Measurement of D-meson production at mid-rapidity in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C77** (2017), no. 8 550, [arXiv:1702.00766].

[21] **ALICE** Collaboration, S. Acharya et al., *Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE*, *Eur. Phys. J.* **C79** (2019), no. 5 388, [arXiv:1901.07979].

[22] **ATLAS** Collaboration, G. Aad et al., *Measurement of $D^{*\pm}$, $D^{\pm}$ and $D_s^{\pm}$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Nucl. Phys.* **B907** (2016) 717–763, [arXiv:1512.02913].

[23] **LHCb** Collaboration, R. Aaij et al., *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **03** (2016) 159, [arXiv:1510.01707]. [Erratum: JHEP05,074(2017)].

[24] **LHCb** Collaboration, R. Aaij et al., *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV*, *JHEP* **06** (2017) 147, [arXiv:1610.02230].

[25] **Particle Data Group** Collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev.* **D98** (2018), no. 3 030001.

[26] S. Alekhin, J. Bl"umlein, and S. Moch, *NLO PDFs from the ABMP16 fit*, *Eur. Phys. J.* **C78** (2018), no. 6 477, [arXiv:1803.07537].

[27] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev.* **D93** (2016), no. 3 033006, [arXiv:1506.07443].

[28] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. **C75** (2015), no. 5 204, [arXiv:1412.3989].

[29] P. Jimenez-Delgado and E. Reya, *Delineating parton distributions and the strong coupling*, Phys. Rev. **D89** (2014), no. 7 074049, [arXiv:1403.1852].

[30] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. **C77** (2017), no. 10 663, [arXiv:1706.00428].

[31] **H1, ZEUS** Collaboration, H. Abramowicz et al., *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. **C75** (2015), no. 12 580, [arXiv:1506.06042].

[32] S. Alekhin, J. Bl"umlein, S. Moch, and R. Placakyte, *Parton distribution functions, α_s, and heavy-quark masses for LHC Run II*, Phys. Rev. **D96** (2017), no. 1 014011, [arXiv:1701.05838].

[33] 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** (2011) 1034–1046, [arXiv:1007.1327].

[34] S. Alekhin et al., *HERAFitter*, Eur. Phys. J. **C75** (2015), no. 7 304, [arXiv:1410.4412].

[35] H. Paukkunen and P. Zurita, *PDF reweighting in the Hessian matrix approach*, JHEP **12** (2014) 100, [arXiv:1402.6623].