THE EUROPEAN
PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study

Richard D. Ball[1], **Alessandro Candido**[2], **Stefano Forte**[2], **Felix Hekhorn**[2], **Emanuele R. Nocera**[3], **Juan Rojo**[4,5,a],
Christopher Schwan[6]

[1] The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland, UK
[2] Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy
[3] Dipartimento di Fisica, Università degli Studi di Torino and INFN, Sezione di Torino, Via Pietro Giuria 1, 10125 Torino, Italy
[4] Department of Physics and Astronomy, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands
[5] Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands
[6] Universität Würzburg, Institut für Theoretische Physik und Astrophysik, 97074 Würzburg, Germany

Abstract We discuss the sensitivity of theoretical predictions of observables used in searches for new physics to parton distributions (PDFs) at large momentum fraction x. Specifically, we consider the neutral-current Drell–Yan production of gauge bosons with invariant masses in the TeV range, for which the forward–backward asymmetry of charged leptons from the decay of the gauge boson in its rest frame is a traditional probe of new physics. We show that the qualitative behaviour of the asymmetry depends strongly on the assumptions made in determining the underlying PDFs. We discuss and compare the large-x behaviour of various different PDF sets, and find that they differ significantly. Consequently, the shape of the asymmetry observed at lower dilepton invariant masses, where all PDF sets are in reasonable agreement because of the presence of experimental constraints, is not necessarily reproduced at large masses where the PDFs are mostly unconstrained by data. It follows that the shape of the asymmetry at high masses may depend on assumptions made in the PDF parametrization, and thus deviations from the traditionally expected behaviour cannot be taken as a reliable indication of new physics. We demonstrate that forward–backward asymmetry measurements could help in constraining PDFs at large x and discuss the accuracy that would be required to disentangle the effects of new physics from uncertainties in the PDFs in this region.

Contents

1 Introduction

An important direction for ongoing and future studies of new physics beyond the Standard Model (BSM) at the Large Hadron Collider (LHC) is the search for novel heavy resonances. The LHC is uniquely suited to direct searches for these resonances, thanks to its unparalleled center of mass energy, $\sqrt{s} = 13.6$ TeV in the recently started Run III, and the high statistics to be accumulated in the coming years, especially in the high-luminosity (HL) phase. For instance, considering representative benchmark BSM scenarios, the HL-LHC is sensitive [1] to searches for sequential Standard Model (SM) W' gauge bosons up to $m_{W'} = 7.8$ TeV, E_6 model Z' gauge bosons up to $m_{Z'} = 5.7$ TeV, and Kaluza-Klein resonances decaying into a $t\bar{t}$ pair up to $m_{KK} = 6.6$ TeV.

The production of such high-mass states proceeds via partonic scattering that involves large values of the momentum fractions x_1 and x_2 of the colliding partons, because the center of mass energy of the partonic collision is $\hat{s} = x_1 x_2 s$. For instance, the on-shell production of a state with invariant mass $m_X = 8$ TeV requires $x_1 x_2 \gtrsim 0.3$, hence for central production at leading order $x_1 = x_2 \approx 0.6$. This is problematic because parton distribution functions (PDFs) [2,3] are poorly known for $x \gtrsim 0.4$, as there is limited data included in current PDF determinations to constrain this kinematic region. Indeed, in the past, claims of possible BSM signals [4] were subsequently traced to poor modeling of the PDFs in the large-x region [5]. The impact of lack of knowledge of the PDFs on BSM searches is thus a delicate issue [6].

Here we wish to further investigate this by specifically considering neutral-current (NC) Drell–Yan (DY) dilepton production and associated observables, frequently used for BSM searches at the LHC. NC Drell–Yan production is one of the cleanest processes in the search for both narrow and broad heavy resonances decaying into dileptons, $pp \rightarrow X \rightarrow \ell^+ \ell^-$, since the two charged leptons can be detected with excellent energy and angular resolution. This also enables the search for smooth, non-resonant distortions with respect to the SM backgrounds, such as those arising in the context of contact interactions or, more generally, induced by Effective Field Theory (EFT) higher-dimensional operators that lead to direct couplings between quarks and leptons [7–10]. Indeed, both ATLAS and CMS have extensively explored this channel in their BSM search program [11–16]. To this purpose, it is mandatory to have a detailed understanding of the dominant SM background, namely dilepton production from quark-antiquark annihilation mediated by a virtual electroweak (EW) boson, $q\bar{q} \rightarrow \gamma^*/Z \rightarrow \ell^+ \ell^-$, with subleading processes involving the quark-gluon and photon-photon initial states.

Drell–Yan production is one of the SM processes which is known to highest perturbative accuracy: indeed, both N³LO QCD results [17] and the full mixed QCD-EW corrections at NNLO [18–22] have become available recently. Therefore, the main uncertainty on theoretical predictions for this process is mostly due to the PDFs, which, as mentioned, are poorly known at large x. Experimentally, uncertainties are minimized when considering observables in which several systematics cancel in part or entirely. An example relevant for the DY process is the forward–backward asymmetry A_{fb} of the angular distribution of the dilepton pair in the center-of-mass frame of the partonic collision, i.e. the asymmetry in the so-called Collins–Soper angle θ^*, recently measured from the Run II dataset by ATLAS [23] and CMS [24]. The sensitivity of this observable to both PDFs and BSM signals has been emphasized recently [25–29], as well as its relevance to extractions of the weak mixing angle $\sin^2 \theta_W$ at the LHC [30]. These studies are mostly restricted to the vicinity

of the Z-boson peak, $m_{\ell\bar{\ell}} \sim m_Z$ with $m_{\ell\bar{\ell}}$ being the dilepton mass, though in a recent study by CMS [24] the forward–backward asymmetry has been used to obtain a lower mass limit (of 4.4 TeV) on a hypothetical Z' heavy gauge boson.

In this work, we assess to which extent different assumptions on the large-x behavior of PDFs, as well as different estimates of the PDF uncertainty in this region, may affect BSM searches, by specifically studying neutral-current Drell–Yan production, and the forward–backward asymmetry in particular. To this purpose, we explain the dependence of the general qualitative features of the asymmetry on the behavior of PDFs, based on an understanding of the analytic dependence of the asymmetry on the partonic luminosities. We then present detailed computations of the forward–backward asymmetry at the LHC, with realistic experimental cuts, using a variety of PDF sets.

We find that first, the large-x PDF shape and uncertainty can differ considerably between different PDF sets, with NNPDF4.0 [31] generally displaying a more flexible shape and a wider uncertainty. And second, that all PDF sets except NNPDF4.0 lead to a qualitative behavior of the asymmetry which in the large-mass multi-TeV region reproduces the shape found around the Z-peak region, even though there is no fundamental reason why this should be the case We will then trace the observed behavior of the asymmetry to that of the underlying PDFs.

The structure of the paper is the following. First in Sect. 2 we review the leading-order (LO) expressions for the Drell–Yan differential distributions and forward–backward asymmetry, in order to explain how the leading qualitative behavior of the asymmetry – specifically the reason for an asymmetry, and its sign – is related to the underlying parton luminosities. We will also show that this LO picture is not qualitatively modified by higher-order perturbative corrections. Then in Sect. 3 we investigate the way the shape of the asymmetry (and specifically its sign) is determined by the large-x behavior of the PDFs. After discussing this in a toy model, we examine current PDF sets: ABMP16 [32], CT18 [33], MSHT20 [34], and NNPDF4.0. Specifically, we compare the behavior of the PDFs and the asymmetry as the final-state dilepton invariant mass is varied. Finally, in Sect. 4 we present predictions for high-mass DY production, specifically the forward–backward asymmetry, at the LHC with realistic experimental cuts, and accounting for NLO QCD and electroweak corrections. For completeness, we present in App. A a comparison to results obtained using the previous, widely used NNPDF3.1 PDF set.

2 Anatomy of Drell–Yan production

The aim of this section is to scrutinize the PDF dependence of the neutral-current Drell–Yan differential cross-section and
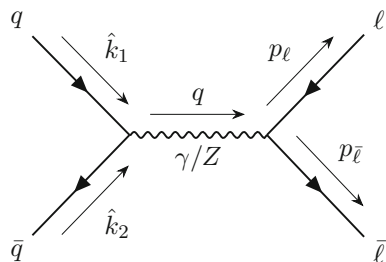
Fig. 1 Neutral-current Drell–Yan production at LO in the quark-antiquark channel

of the associated forward–backward asymmetry by reviewing the LO kinematics, determining LO analytic expressions, and finally comparing these analytical calculations to the results of LO and NLO numerical simulations obtained using MADGRAPH5_AMC@NLO [35] interfaced to PINEAPPL [36,37]. Specifically, we will relate the behavior of the differential distribution and asymmetry to the relevant parton luminosities.

2.1 Drell–Yan kinematics and cross-sections at LO

We consider dilepton production via the exchange of an electroweak neutral gauge boson Z/γ^* in proton–proton collisions:

$$\mathrm{p}(k_1) + \mathrm{p}(k_2) \to Z/\gamma^*(q) \to \ell(p_\ell) + \bar{\ell}(p_{\bar{\ell}}) + X. \quad (2.1)$$

The hadronic differential cross-section $\mathrm{d}\sigma^{\mathrm{pp}\to\ell\bar{\ell}}$ is factorized in terms of PDFs f_i and the partonic cross sections $\mathrm{d}\hat{\sigma}_{ij}$ for incoming partons of species $i,\ j$ as

$$\mathrm{d}\sigma^{\mathrm{pp}\to\ell\bar{\ell}} = \sum_{ij} \int_0^1 \mathrm{d}x_1 \mathrm{d}x_2\, f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2)$$
$$\times \mathrm{d}\hat{\sigma}_{ij}(\hat{k}_1 = x_1 k_1, \hat{k}_2 = x_2 k_2). \quad (2.2)$$

In the sequel we will set the factorization scale μ_F to the invariant mass of the gauge boson, i.e. the dilepton invariant mass, so $\mu_F^2 = m_{\ell\bar{\ell}}^2 = (p_\ell + p_{\bar{\ell}})^2$. The kinematics and Feynman diagram of the LO partonic process in the quark-antiquark channel are shown in Fig. 1. We do not consider photon-initiated processes, as they do not affect the qualitative features of our discussion.

At LO, the momentum fractions of the two incoming partons are fully fixed by knowledge of the invariant mass and rapidity of the gauge boson, i.e. of the dilepton pair $y_{\ell\bar{\ell}} = (y_\ell + y_{\bar{\ell}})/2$:

$$x_1 = \frac{m_{\ell\bar{\ell}}}{\sqrt{s}} \exp(y_{\ell\bar{\ell}}), \quad x_2 = \frac{m_{\ell\bar{\ell}}}{\sqrt{s}} \exp(-y_{\ell\bar{\ell}}), \quad (2.3)$$

where the center of mass energy of the hadronic collision is $s = (k_1 + k_2)^2$ and at LO $m_{\ell\bar{\ell}}^2 = \hat{s} = x_1 x_2 s$. The absolute

dilepton rapidity thus lies in the range $|y_{\ell\bar{\ell}}| \leq \ln(\sqrt{s}/m_{\ell\bar{\ell}})$. Beyond LO there might be extra radiation in the final state, so the LO kinematics provides a lower bound on the momentum fractions of the incoming partons, and all values of the momentum fractions such that $x_{1,2} \geq m_{\ell\bar{\ell}}/\sqrt{s}$ are allowed.

It is useful to define the so-called Collins–Soper angle θ^* [38], which in the hadronic center-of-mass (CoM) frame is defined as

$$\cos\theta^* = \mathrm{sign}(y_{\ell\bar{\ell}}) \cos\theta ,$$
$$\cos\theta \equiv \frac{p_\ell^+ p_{\bar{\ell}}^- - p_\ell^- p_{\bar{\ell}}^+}{m_{\ell\bar{\ell}} \sqrt{m_{\ell\bar{\ell}}^2 + p_{\mathrm{T},\ell\bar{\ell}}^2}}, \quad p^\pm = p^0 \pm p^3. \quad (2.4)$$

It is easy to show that the Collins–Soper angle θ^* coincides with the scattering angle of the lepton in the partonic CoM frame, $\bar{\theta}$. The latter is defined in terms of the lepton momentum as

$$\cos\bar{\theta} \equiv \frac{p_\ell^z}{m_{\ell\bar{\ell}}} , \quad (2.5)$$

where the z axis is along the direction of the incoming quark-antiquark pair. In the partonic CoM frame, of course, $p_\ell^z = -p_{\bar{\ell}}^z$ and $y_{\ell\bar{\ell}} = 0$, so

$$p_\ell^\pm = p_{\bar{\ell}}^\mp = m_{\ell\bar{\ell}} \left(1 \pm \cos\bar{\theta}\right), \quad (2.6)$$

and substituting in Eq. (2.4) it immediately follows that, taking the convention $\mathrm{sign}(y_{\ell\bar{\ell}}) = \mathrm{sign}(0) = +1$, $\cos\theta^* = \cos\theta = \cos\bar{\theta}$. The expression of $\cos\theta$ in Eq. (2.4) is manifestly invariant upon boosts along the z axis, so the identification of θ with the CoM scattering angle $\bar{\theta}$ remains true in any reference frame.

Note that the definition Eq. (2.5) requires a choice for the positive direction of the z axis, which is usually taken along the direction of the incoming fermion (quark). This direction is not experimentally accessible in proton–proton collisions, so the Collins–Soper angle is defined by always taking the positive z axis in the direction of the boosted dilepton pair, i.e., at LO, along the direction of the incoming quark with largest momentum fraction, i.e. by supplementing in the definition a factor $\mathrm{sign}(y_{\ell\bar{\ell}})$. Hence $\cos\theta^* = \cos\bar{\theta}$ ($\cos\theta^* = -\cos\bar{\theta}$) if the momentum fraction of the incoming quark (antiquark) is the largest.

The hard scattering matrix elements that enter the partonic cross-section in Eq. (2.2) are the sum of a pure photon-exchange contribution, a photon-Z interference term, and a pure Z-exchange contribution. Of course, in the region $m_{\ell\bar{\ell}} \gtrsim m_Z$ these contributions are all of the same order. Standard arguments [39] then imply that, because in the Standard Model the photon coupling to leptons is vector while the Z coupling is chiral, the pure photon and pure Z contributions to the cross-section are necessarily even in $\cos\theta^*$ while the interference term is odd.

Specifically, at LO the fully differential hadronic cross-section can be obtained from the well-known result [39] for $e^+e^- \to \mu^+\mu^-$ by replacing the incoming lepton charges with those of the quarks, and accounting for the PDFs, with the result

$$
\frac{\mathrm{d}^3\sigma}{\mathrm{d}m_{\ell\bar\ell}\,\mathrm{d}y_{\ell\bar\ell}\,\mathrm{d}\cos\theta^*}
$$
$$
= \frac{\pi\alpha^2}{3m_{\ell\bar\ell}s}
$$
$$
\times \Bigg((1+\cos^2(\theta^*)) \sum_q S_q
$$
$$
\times \Big[f_q(x_1, m_{\ell\bar\ell}^2) f_{\bar q}(x_2, m_{\ell\bar\ell}^2) + f_q(x_2, m_{\ell\bar\ell}^2) f_{\bar q}(x_1, m_{\ell\bar\ell}^2) \Big]
$$
$$
+ \cos\theta^* \sum_q A_q \,\mathrm{sign}(y_{\ell\bar\ell})
$$
$$
\times \Big[f_q(x_1, m_{\ell\bar\ell}^2) f_{\bar q}(x_2, m_{\ell\bar\ell}^2) - f_q(x_2, m_{\ell\bar\ell}^2) f_{\bar q}(x_1, m_{\ell\bar\ell}^2) \Big] \Bigg),
$$
$$
(2.7)
$$

where α is the QED coupling and the even (symmetric) and odd (antisymmetric) couplings are given by

$$
S_q = e_l^2 e_q^2 + P_{\gamma Z} \cdot e_l v_l e_q v_q + P_{ZZ} \cdot (v_l^2 + a_l^2)(v_q^2 + a_q^2)
$$
$$
A_q = P_{\gamma Z} \cdot 2 e_l a_l e_q a_q + P_{ZZ} \cdot 8 v_l a_l v_q a_q , \qquad (2.8)
$$

in terms of the electric charges e_l, e_q and the vector and axial couplings v_l, v_q and a_l, a_q of the leptons and quarks, and the propagator factors

$$
P_{\gamma Z}(m_{\ell\bar\ell}) = \frac{2 m_{\ell\bar\ell}^2 (m_{\ell\bar\ell}^2 - m_Z^2)}{\sin^2(\theta_W)\cos^2(\theta_W)\left[(m_{\ell\bar\ell}^2 - m_Z^2)^2 + \Gamma_Z^2 m_Z^2 \right]}
$$
$$
(2.9)
$$
$$
P_{ZZ}(m_{\ell\bar\ell}) = \frac{m_{\ell\bar\ell}^4}{\sin^4(\theta_W)\cos^4(\theta_W)\left[(m_{\ell\bar\ell}^2 - m_Z^2)^2 + \Gamma_Z^2 m_Z^2 \right]},
$$
$$
(2.10)
$$

with m_Z and Γ_Z respectively the Z mass and width and θ_W the weak mixing angle. In Fig. 2 we display the symmetric S_q (left) and antisymmetric A_q (right) couplings, Eq. (2.8), for up-like and down-like quarks, as a function of the dilepton invariant mass $m_{\ell\bar\ell}$. Both couplings are around a factor 2 larger for up-like quarks than for down-like quarks, and become $m_{\ell\bar\ell}$-independent for $m_{\ell\bar\ell} \gtrsim 1$ TeV, where they take the asymptotic values $\bar S_q$, $\bar A_q$ obtained by substituting in Eq. (2.8) the large-mass expressions of the propagator factors

$$
\bar P_{\gamma Z} = \frac{2}{\sin^2(\theta_W)\cos^2(\theta_W)}, \qquad \bar P_{ZZ} = \frac{1}{\sin^4(\theta_W)\cos^4(\theta_W)},
$$
$$
(2.11)
$$

to which $P_{\gamma Z}$ and P_{ZZ} respectively reduce up to $O(m_Z^2/m_{\ell\bar\ell}^2)$ corrections.

The interference term proportional to A_q is odd in the Collins–Soper angle $\cos\theta^*$, leading to a forward–backward scattering asymmetry. In a proton–proton collision the initial state is completely symmetric, so the quark and antiquark contributions to the cross-section Eq. (2.7) are necessarily symmetric upon the interchange of the incoming quark and antiquark, with the corresponding momentum fractions fixed at LO by Eq. (2.3). However, as mentioned, there is a sign change in the relation between $\cos\theta^*$ and $\cos\theta$ according to whether the incoming parton with largest momentum fraction is a quark or an antiquark, i.e., when interchanging x_1 with x_2 in the argument of the quark and antiquark PDFs, thereby leading to the result of Eq. (2.7). This leads to a forward–backward asymmetry whenever the quark and antiquark PDFs have different x dependence.

In order to understand the relation of this forward–backward asymmetry in terms of the behavior of the PDFs, it is convenient to rewrite the PDF combinations that contribute to the differential cross-section Eq. (2.7) in terms of symmetric and antisymmetric parton luminosities, defined as

$$
\mathcal{L}_{S,q}(m_{\ell\bar\ell}, y_{\ell\bar\ell}) \equiv f_q(x_1, m_{\ell\bar\ell}^2) f_{\bar q}(x_2, m_{\ell\bar\ell}^2)
$$
$$
+ f_q(x_2, m_{\ell\bar\ell}^2) f_{\bar q}(x_1, m_{\ell\bar\ell}^2) ,
$$
$$
\mathcal{L}_{A,q}(m_{\ell\bar\ell}, y_{\ell\bar\ell}) \equiv \mathrm{sign}(y_{\ell\bar\ell})
$$
$$
\times \Big[f_q(x_1, m_{\ell\bar\ell}^2) f_{\bar q}(x_2, m_{\ell\bar\ell}^2) - f_q(x_2, m_{\ell\bar\ell}^2) f_{\bar q}(x_1, m_{\ell\bar\ell}^2) \Big] ,
$$
$$
(2.12)
$$

where the momentum fractions x_1 and x_2 are given in terms of $m_{\ell\bar\ell}$, $y_{\ell\bar\ell}$, and $\sqrt{s}$ in Eq. (2.3). Note that both parton luminosities are invariant under the interchange $x_1 \leftrightarrow x_2$, upon which $y_{\ell\bar\ell} \to -y_{\ell\bar\ell}$. In terms of these luminosities, the triple differential cross-section Eq. (2.7) takes the compact form

$$
\frac{\mathrm{d}^3\sigma}{\mathrm{d}m_{\ell\bar\ell}\,\mathrm{d}y_{\ell\bar\ell}\,\mathrm{d}\cos\theta^*} = \frac{\pi\alpha^2}{3m_{\ell\bar\ell}s}
$$
$$
\times \Bigg((1+\cos^2(\theta^*)) \sum_q S_q \mathcal{L}_{S,q}(m_{\ell\bar\ell}, y_{\ell\bar\ell})
$$
$$
+ \cos\theta^* \sum_q A_q \mathcal{L}_{A,q}(m_{\ell\bar\ell}, y_{\ell\bar\ell}) \Bigg),
$$
$$
(2.13)
$$

which explicitly displays its symmetry properties upon the transformation $\cos\theta^* \to -\cos\theta^*$, equivalent to a charge conjugation transformation $q \leftrightarrow \bar q$ and $\ell \leftrightarrow \bar\ell$.

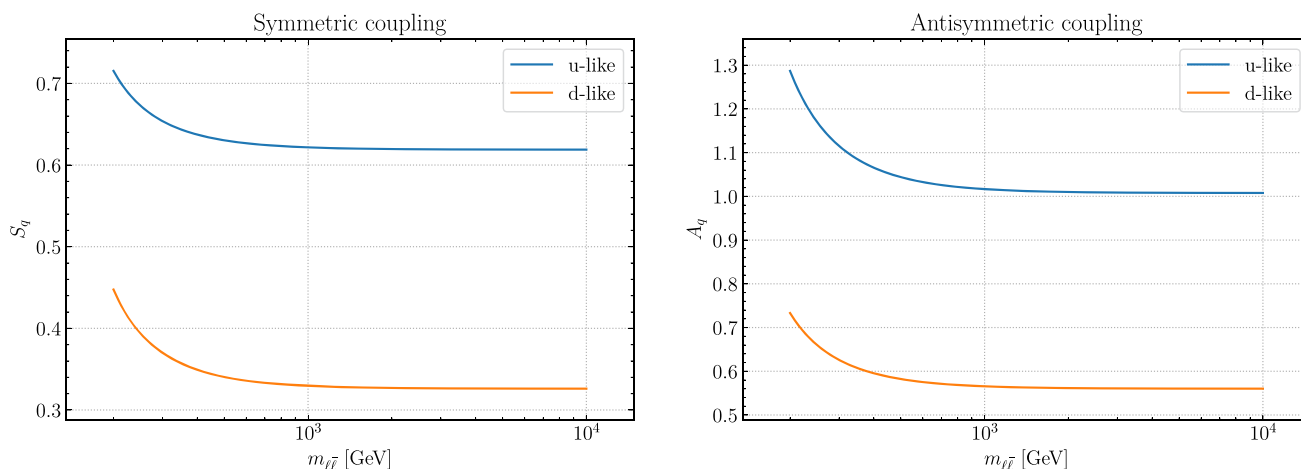The symmetric and antisymmetric parton luminosities Eq. (2.12) can also be expressed in terms of the sum and differ-

Fig. 2 The symmetric S_q (left) and antisymmetric A_q (right) couplings, Eq. (2.8), for up-like and down-like quarks, as a function of the dilepton invariant mass $m_{\ell\bar{\ell}}$.

ence of quark and antiquark PDFs,

$$f_q^{\pm}(x, Q) = f_q(x, Q) \pm f_{\bar{q}}(x, Q), \tag{2.14}$$

where f_q^- is usually called the valence PDF combination, and f_q^+ the total quark PDF. Note that at LO, and more generally in factorization schemes in which PDFs are positive, such as $\overline{\text{MS}}$ [40], f_q^+ is positive while f_q^- in general is not, and $f_q^+ > |f_q^-|$. We can write the symmetric and antisymmetric parton luminosities in Eq. (2.12) as

$$\mathcal{L}_{S,q}(m_{\ell\bar{\ell}}, y_{\ell\bar{\ell}}) = \frac{1}{2}$$
$$\times \left(f_q^+(x_1, m_{\ell\bar{\ell}}^2) f_q^+(x_2, m_{\ell\bar{\ell}}^2) - f_q^-(x_2, m_{\ell\bar{\ell}}^2) f_q^-(x_1, m_{\ell\bar{\ell}}^2) \right) \tag{2.15}$$

$$\mathcal{L}_{A,q}(m_{\ell\bar{\ell}}, y_{\ell\bar{\ell}}) = \frac{\text{sign}(y_{\ell\bar{\ell}})}{2}$$
$$\times \left(f_q^-(x_1, m_{\ell\bar{\ell}}^2) f_q^+(x_2, m_{\ell\bar{\ell}}^2) - f_q^-(x_2, m_{\ell\bar{\ell}}^2) f_q^+(x_1, m_{\ell\bar{\ell}}^2) \right). \tag{2.16}$$

The symmetric luminosity $\mathcal{L}_{S,q}$ is of course positive, and it is dominated by the $f_q^+(x_1, m_{\ell\bar{\ell}}^2) f_q^+(x_2, m_{\ell\bar{\ell}}^2)$ term, which is always larger than the valence contribution $f_q^-(x_2, m_{\ell\bar{\ell}}^2)$ $f_q^-(x_1, m_{\ell\bar{\ell}}^2)$. The sign of the antisymmetric combination, that in turn drives the sign of the forward–backward asymmetry, is in general not determined uniquely. If x_1 is in the region of the valence peak, and x_2 in the small x region, then $f^-(x_1, m_{\ell\bar{\ell}}^2) \gg f^-(x_2, m_{\ell\bar{\ell}}^2)$, and the antisymmetric luminosity is positive provided only that the valence PDF is positive. As we will discuss in Sect. 3, while this is indeed the case in the Z-peak region, it is actually not necessarily the case in the high dilepton mass region relevant for BSM searches.

2.2 Single-differential distributions and the forward–backward asymmetry

Starting from the triple differential cross section, Eq. (2.13), one can define single differential distributions by integrating the other two kinematic variables over the available phase space. In particular, the single-differential distribution in the Collins–Soper angle θ^* is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} = \int_{m_{\ell\bar{\ell}}^{\min}}^{\sqrt{s}} \mathrm{d}m_{\ell\bar{\ell}} \int_{\ln(m_{\ell\bar{\ell}}/\sqrt{s})}^{\ln(\sqrt{s}/m_{\ell\bar{\ell}})} \mathrm{d}y_{\ell\bar{\ell}} \frac{\mathrm{d}^3\sigma}{\mathrm{d}m_{\ell\bar{\ell}} \, \mathrm{d}y_{\ell\bar{\ell}} \, \mathrm{d}\cos\theta^*}, \tag{2.17}$$

where $m_{\ell\bar{\ell}}^{\min}$ is a lower kinematic cut in the dilepton invariant mass. Since Eq. (2.13) falls off steeply with $m_{\ell\bar{\ell}}$, the region with $m_{\ell\bar{\ell}} \gtrsim m_{\ell\bar{\ell}}^{\min}$ will dominate the integral. Given that the dependence of the fully differential cross-section Eq. (2.13) on the Collins–Soper angle factorizes with respect to the PDF dependence, the integration over rapidity and invariant mass does not affect the $\cos\theta^*$ dependence, and the single-differential cross section Eq. (2.17) takes the simple form

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} = (1 + \cos^2\theta^*) \sum_q g_{S,q} + \cos\theta^* \sum_q g_{A,q}, \tag{2.18}$$

where the symmetric and antisymmetric coefficients $g_{S,q}$ and $g_{A,q}$ depend on the quark flavor and on the invariant mass cut $m_{\ell\bar{\ell}}^{\min}$, but not on the Collins–Soper angle itself. The contributions relevant for the forward–backward asymmetry, $g_{A,q}$, are given at LO by

$$g_{A,q} = \frac{\pi\alpha^2}{3s} \int_{m_{\ell\bar{\ell}}^{\min}}^{\sqrt{s}} \frac{\mathrm{d}m_{\ell\bar{\ell}}}{m_{\ell\bar{\ell}}} A_q(m_{\ell\bar{\ell}}) \int_{\ln(m_{\ell\bar{\ell}}/\sqrt{s})}^{\ln(\sqrt{s}/m_{\ell\bar{\ell}})} \mathrm{d}y_{\ell\bar{\ell}} \, \mathcal{L}_{A,q}(m_{\ell\bar{\ell}}, y_{\ell\bar{\ell}}), \tag{2.19}$$
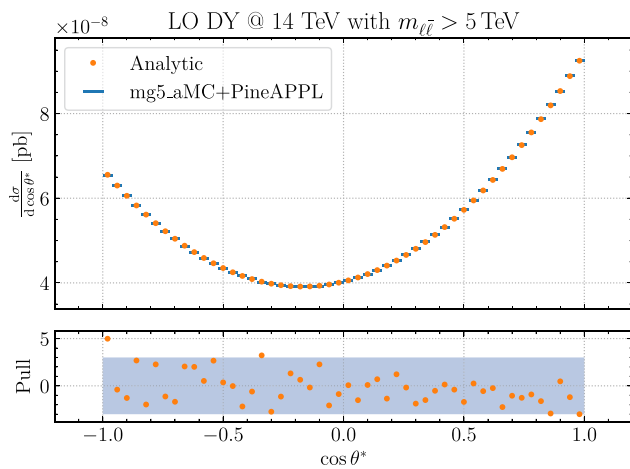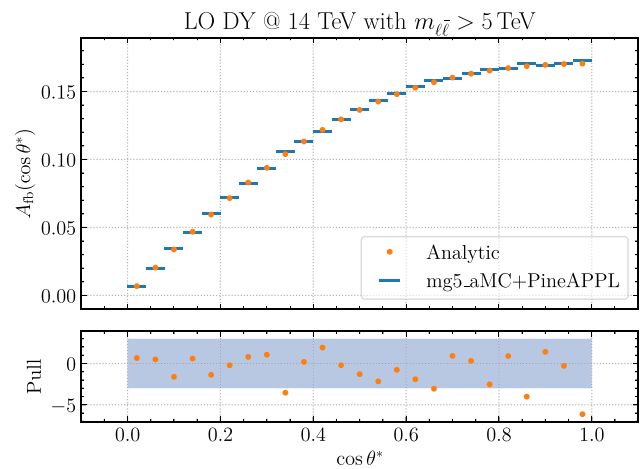
Fig. 3 The single-inclusive differential distribution in the Collins–Soper angle $\cos\theta^*$, Eq. (2.17), and the corresponding forward–backward asymmetry computed at LO, where the analytic calculation Eq. (2.22) is compared with the numerical simulation based on MAD-GRAPH5_AMC@NLO interfaced to PINEAPPL. The bottom panels dis-play the difference between the analytic and numerical calculations relative to the Monte Carlo integration uncertainty. The blue band indicates the 3σ uncertainty interval. One of the replicas of the NNPDF4.0 NNLO PDF set is used as input to the calculation
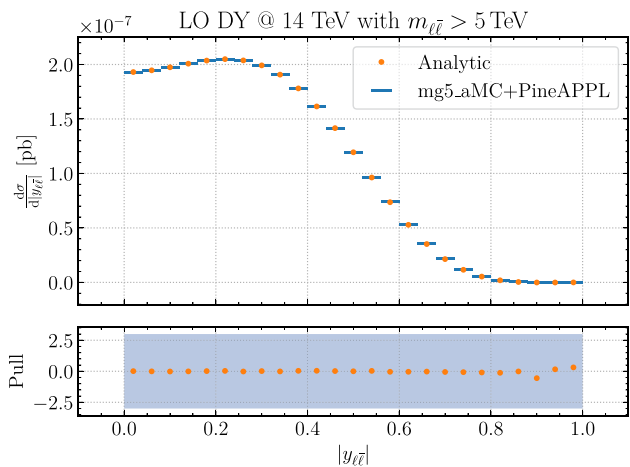

Fig. 4 Same as Fig. 3 but now for the absolute dilepton rapidity distribution $|y_{\ell\bar{\ell}}|$

which in the large-$m_{\ell\bar{\ell}}$ region, expressing the longitudinal momentum integration in terms of x_1 (assuming $x_1 \geq x_2$), becomes

$$g_{A,q} = \frac{\pi\alpha^2\bar{A}_q}{3s} \int\limits_{m_{\ell\bar{\ell}}^{\min}}^{\sqrt{s}} \frac{\mathrm{d}m_{\ell\bar{\ell}}}{m_{\ell\bar{\ell}}} \int\limits_{m_{\ell\bar{\ell}}/\sqrt{s}}^{1} \frac{\mathrm{d}x_1}{x_1} \mathcal{L}_{A,q}(m_{\ell\bar{\ell}},x_1)$$
$$+ \mathcal{O}\left(\frac{m_Z^2}{m_{\ell\bar{\ell}}^2}\right),\tag{2.20}$$

where the $m_{\ell\bar{\ell}}$-independent effective couplings $\bar{A}_q$ are given substituting in Eq. (2.8) the expressions for the asymptotic propagator factors Eq. (2.11).

Upon integration over the Collins–Soper angle, the anti-symmetric contribution vanishes: so for instance the rapidity distribution

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y_{\ell\bar{\ell}}} = \int\limits_{m_{\ell\bar{\ell}}^{\min}}^{\sqrt{s}} \mathrm{d}m_{\ell\bar{\ell}} \int\limits_{-1}^{1} \mathrm{d}\cos\theta^* \frac{\mathrm{d}^3\sigma}{\mathrm{d}m_{\ell\bar{\ell}}\,\mathrm{d}y_{\ell\bar{\ell}}\,\mathrm{d}\cos\theta^*},\tag{2.21}$$

does not depend on terms proportional to A_q. Hence, for BSM searches in which one is interested in the interference terms, as well as for PDF studies in which one is interested in the valence-sea separation, the forward–backward asymmetry is especially relevant. This observable is defined at the differential level as

$$A_{\mathrm{fb}}(\cos\theta^*) \equiv \frac{\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*}(\cos\theta^*) - \frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*}(-\cos\theta^*)}{\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*}(\cos\theta^*) + \frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*}(-\cos\theta^*)},\quad \cos\theta^* > 0,$$
$$\tag{2.22}$$

which in terms of the coefficients introduced in Eq. (2.18) is given at LO by

$$A_{\mathrm{fb}}(\cos\theta^*) = \frac{\cos\theta^*}{(1+\cos^2(\theta^*))}\frac{\sum_q g_{A,q}}{\sum_{q'} g_{S,q'}},\quad \cos\theta^* > 0.\tag{2.23}$$

This shows that the dependence on $\cos\theta^*$ factorizes and the PDF dependence only appears as an overall normalization factor depending on the ratio of $\sum_q g_{A,q}$ and $\sum_q g_{S,q}$, which in turn depend on the antisymmetric and symmetric partonic luminosities $\mathcal{L}_{A,q}$ and $\mathcal{L}_{S,q}$ respectively. Note that the overall sign of A_{fb} remains in general undetermined.

In order to illustrate concretely these results, in Fig. 3 we display the single-inclusive differential distribution in the

$\cos \theta^*$, Eq. (2.17), and the corresponding forward–backward asymmetry, Eq. (2.22) evaluated at LO for $m_{\ell\bar{\ell}}^{\min} = 5\,TeV$. The single-differential rapidity distribution Eq. (2.21) is also shown for reference in Fig. 4. We display both a numerical evaluation based on MADGRAPH5_AMC@NLO interfaced to PINEAPPL, as well as analytic results found using the form Eq. (2.13) of the triple differential luminosity, with all the values of the parameters entering Eqs. (2.8) to (2.10) set to the values used in the MADGRAPH5_AMC@NLO runcard, and performing numerically the integrals in Eqs. (2.17) and (2.21). For validation purposes, no kinematic cuts are applied to the rapidities and transverse momenta of final-state leptons. The PDF input is taken to be given, for illustrative purposes, by one of the replicas of the NNPDF4.0 NNLO set. The relative difference between the analytic and numerical calculation is shown in the bottom panels of Fig. 3 and demonstrates perfect agreement.

While the discussion so far has been presented at LO, its qualitative features are unaffected by higher-order corrections. To illustrate this, in Fig. 5 we compare the LO result from Fig. 3 to the corresponding NLO QCD result. The bottom panels display the NLO K-factor for the $\cos \theta^*$ distribution and the forward–backward asymmetry. Whereas the NLO K-factor in the $\cos \theta^*$ distribution is quite large (around 40%) it exhibits only a mild dependence on the Collins–Soper angle. For A_{fb}, the K-factor is at the 10% level and essentially independent of the value of $\cos \theta^*$.

3 The forward–backward asymmetry and the large-x PDFs

After our general discussion of the Drell–Yan process, we now investigate proton structure at large-x, focusing on its impact on the forward–backward asymmetry $A_{\mathrm{fb}} (\cos \theta^*)$ at large invariant masses. First, we discuss the dependence of the qualitative features of the asymmetry, and specifically its sign, on the behavior of the underlying PDFs: we illustrate this in a toy model, and compare results to a simple and commonly used approximation. Subsequently, we study the large-x behavior of the PDFs from several recent PDF sets: we compare PDFs, luminosities and the LO asymmetry A_{fb} as a function of the dilepton invariant mass $m_{\ell\bar{\ell}}$.

3.1 Qualitative features of A_{fb}

In order to understand the main qualitative features of the $\cos \theta^*$ distribution and of the asymmetry A_{fb} and their dependence on the properties of the underlying PDFs, it is instructive to evaluate predictions based on the same computational setup adopted in Sect. 2, namely LO matrix elements without kinematic cuts, using toy PDFs as input. We consider toy

quark and antiquark PDF with the form

$$xf_q(x) = A_q x^{-a_q} (1-x)^{b_q}, \quad xf_{\bar{q}}(x) = A_{\bar{q}} x^{-a_{\bar{q}}} (1-x)^{b_{\bar{q}}},$$
(3.1)

where A_q and $A_{\bar{q}}$ are normalization constants, irrelevant for this discussion. For simplicity we neglect the scale dependence of the PDFs. We then compute the single-differential distribution Eq. (2.17) and the asymmetry Eq. (2.22) with different assumptions on the large x-behavior of these toy PDFs, i.e. different values of the large-x exponents b_q, $b_{\bar{q}}$.

Since the overall normalization does not affect the shape of the distribution, we set $A_q = A_{\bar{q}} = 1$. Furthermore, since we are not interested in the small-x behavior, we set $a_q = a_{\bar{q}} = 1$. Hence, we consider simple scenarios in which

$$
\begin{aligned}
xf_q^+(x; b_q, b_{\bar{q}}) &= xf_q(x) + xf_{\bar{q}}(x) \\
&= x^{-1} \left[(1-x)^{b_q} + (1-x)^{b_{\bar{q}}} \right],
\end{aligned}
$$
(3.2)

$$
\begin{aligned}
xf_q^-(x; b_q, b_{\bar{q}}) &= xf_q(x) - xf_{\bar{q}}(x) \\
&= x^{-1} \left[(1-x)^{b_q} - (1-x)^{b_{\bar{q}}} \right],
\end{aligned}
$$
(3.3)

with different choices of the parameters b_q and $b_{\bar{q}}$. Specifically, we consider a scenario with $b_q < b_{\bar{q}}$, in particular $(b_q, b_{\bar{q}}) = (3, 5)$; a scenario with $(b_q, b_{\bar{q}}) = (3, 3)$; and a third scenario in which the quark PDFs at large-x fall off more rapidly than the antiquarks, $(b_q, b_{\bar{q}}) = (5, 3)$.

In Fig. 6 we display both the $\cos \theta^*$ single-inclusive distribution Eq. (2.17) and the asymmetry Eq. (2.22). It is apparent that if the antiquark PDFs fall off at large-x faster than the quarks, i.e. when $b_q < b_{\bar{q}}$ the forward–backward asymmetry is positive, while if the converse is true it is negative. Of course if the quark and antiquark PDFs behave in the same way there is no asymmetry. Indeed, the condition for a negative asymmetry is (assuming $x_1 > x_2$)

$$
\begin{aligned}
\mathrm{sign} \left[\mathcal{L}_{A,q} \right] &= \mathrm{sign} \left[\frac{f_q^+(x_2)}{f_q^+(x_1)} - \frac{f_q^-(x_2)}{f_q^-(x_1)} \right] \\
&= \mathrm{sign} \left[\frac{f_q(x_2)}{f_q(x_1)} - \frac{f_{\bar{q}}(x_2)}{f_{\bar{q}}(x_1)} \right], \quad x_1 > x_2.
\end{aligned}
$$
(3.4)

Namely, what determines the sign of the antisymmetric luminosity, and thus of the forward–backward asymmetry, is the relative rate of decrease of the quark and antiquark, or valence and total quark PDFs.

In the simple model that we discussed, this rate of decrease is controlled by the values of the exponents b_q and $b_{\bar{q}}$. The simple model has unphysical features, in that a negative asymmetry corresponds to a negative valence distribution, which conflicts with sum rules. It is easy to construct a more contrived model, in which the valence drops faster than the total quark PDF, yet it remains positive. Also one could argue that Brodsky-Farrar counting rules [41,42] imply that $b_q < b_{\bar{q}}$ as $x \to 1$, so a faster dropping antiquark is favored.
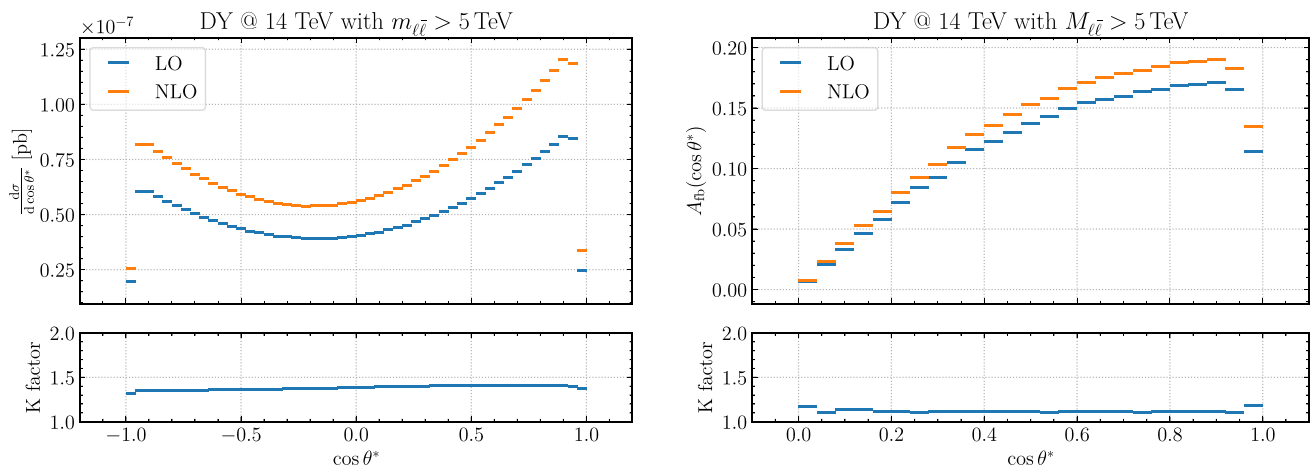
Fig. 5 Same as Fig. 3 now comparing the LO result to the NLO QCD result obtained using MADGRAPH5_AMC@NLO. The K-factor is shown in the lower panel.
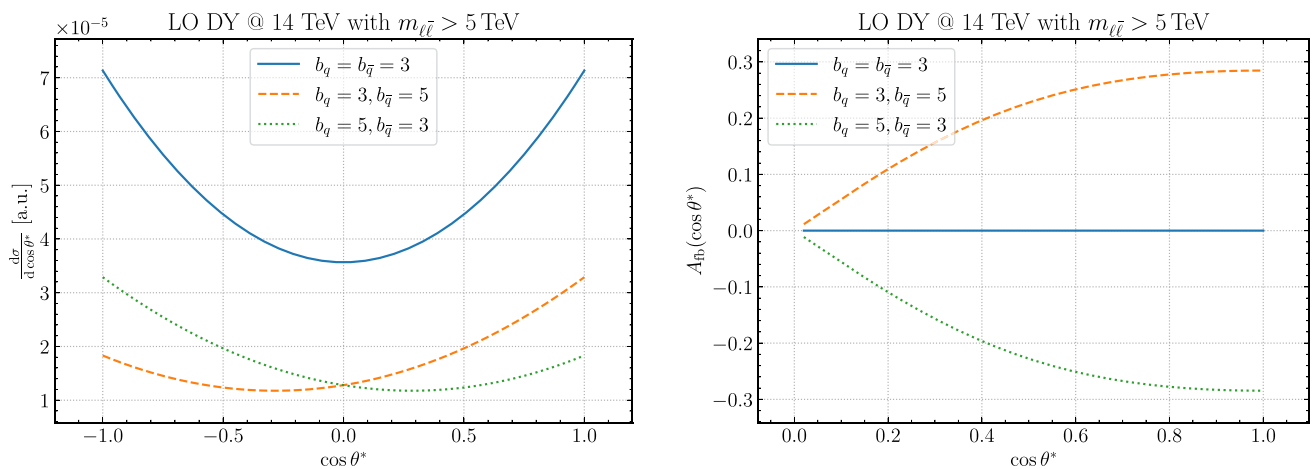

Fig. 6 The single-inclusive $\cos\theta^*$ distribution Eq. (2.17) (left) and the corresponding forward–backward asymmetry (right panel) Eq. (2.22) evaluated using the toy PDFs of Eq. (3.1). No kinematic cuts are applied except for $m_{\ell\bar{\ell}}^{\min} = 5$ TeV

However, counting rules are supposed to only hold asymptotically, so whether they apply in any given region of x is a priori unclear. Again, it is easy to construct more contrived models in which the value of the exponent or effective exponent is x-dependent. However, our purpose is to highlight which features determine the sign of the asymmetry, and not to construct an explicit PDF model. In fact, in Sect. 3.2 we will explicitly exhibit PDFs that do lead to a negative asymmetry, while being consistent with sum rules and not leading to contradiction with asymptotic counting.

It is interesting to note that different conclusions on the asymmetry could be reached by using an approximation to the asymmetry which is quite accurate in the Z peak region. This approximation however turns out to fail at high invariant mass. Indeed, the expression Eq. (2.16) of the antisymmetric luminosity in terms of the valence and total PDF combinations f_q^+ and f_q^- PDF combinations suggests an approximation based on the expectation that the valence is dominant at large x and the sea is dominant at small x. Assuming $x_1 > x_2$,

one then expects that

$$\mathcal{L}_{A,u}(y_{\ell\bar{\ell}}, m_{\ell\bar{\ell}}) \approx \frac{1}{2} f_u^-(x_1, m_{\ell\bar{\ell}}^2) f_u^+(x_2, m_{\ell\bar{\ell}}^2), \quad x_1 > x_2.$$
(3.5)

This is clearly true in the Z-peak region, which motivates the suggestion to use the measurement of A_{fb} as a means to constrain the valence quark combinations [27].

However, while Eq. (3.5) provides a satisfactory approximation in the Z-peak region, it fails at larger $m_{\ell\bar{\ell}}$ values. Indeed, for on-shell Z production, with $\sqrt{s} = 14$ TeV, for a dilepton rapidity with $y_{\ell\bar{\ell}} \sim 2.5$, the limit of the acceptance region of ATLAS and CMS, the colliding partons have $x_1 = 0.09$ and $x_2 = 6 \times 10^{-4}$. So indeed the contribution in which the valence PDF is evaluated at the smallest x value is highly suppressed. But for $m_{\ell\bar{\ell}} = 5$ TeV, the smallest value of x_2, attained when $x_1 = 1$, is $x_2 = 0.35$: so both momentum fractions are large and in fact to the right of the valence peak. In such case, there is no obvious hierarchy between the
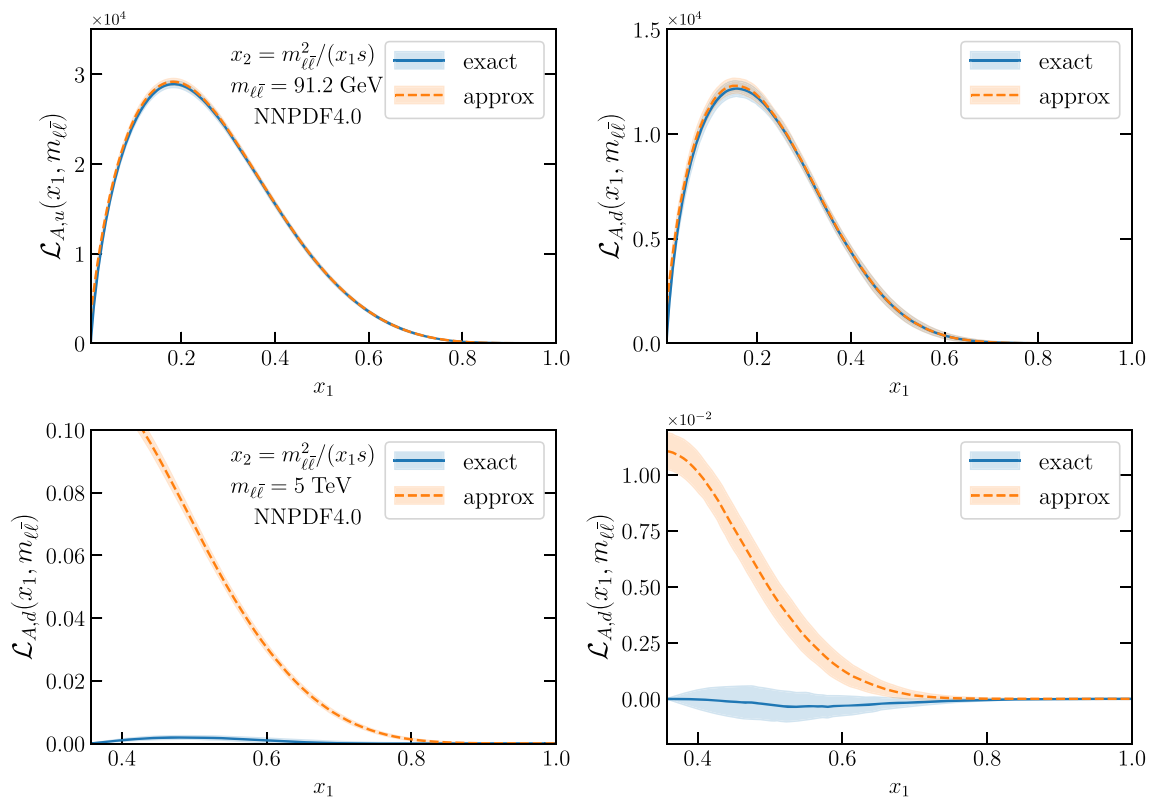
Fig. 7 The antisymmetric partonic luminosity $\mathcal{L}_{A,q}$, Eq. (2.16), for the up and down quarks compared to the approximation Eq. (3.5) in the case of NNPDF4.0 at $m_{\ell\bar{\ell}} = m_Z$ (top) and $m_{\ell\bar{\ell}} = 5$ TeV (bottom panels)

different terms that contribute to to antisymmetric luminosity $\mathcal{L}_{A,q}$.

This is illustrated in Fig. 7, where we compare the antisymmetric luminosity $\mathcal{L}_{A,q}$ for the up and down quarks to the approximation Eq. (3.5), evaluated with NNPDF4.0 NNLO, in the Z-peak region $m_{\ell\bar{\ell}} = m_Z$ and at $m_{\ell\bar{\ell}} = 5$ TeV. While indeed for $m_{\ell\bar{\ell}} = m_Z$ Eq. (3.5) reproduces the exact luminosity, this is not the case for $m_{\ell\bar{\ell}} \gg m_Z$: both the magnitude and the shape of the luminosity are very different. This qualitative behavior is common to all PDF sets: the approximation fails equally badly regardless of the PDF set.

We conclude that there is no simple relation between the sign of the asymmetry and that of the valence PDF, and that the behavior of the asymmetry must be determined by studying the large-x behavior of the quark and antiquark PDFs.

3.2 Parton distributions

We assess now the large-x behavior of the quark and antiquark PDFs in different recent PDF determinations: specifically, we compare ABMP16, CT18, NNPDF4.0, and MSHT20. For completeness, in App. A we also present results obtained with the widely used NNPDF3.1 [43] set.

First, we provide a qualitative assessment of the relative size of the PDFs corresponding to individual quark flavors, both for the total and valence PDFs. In Fig. 8 we compare the total xf_q^+ and valence xf_q^- quark PDF combinations for the up, down, strange, and charm quarks, evaluated at $m_{\ell\bar{\ell}} = 5$ TeV with the NNPDF4.0 NNLO PDF set. The right panels display the corresponding relative 68% CL uncertainties. Note that, because of the way uncertainties are delivered by the various groups, in this and all subsequent plots uncertainties for NNPDF are given are confidence levels (not necessarily Gaussian) determined from the Monte Carlo replica sample, and thus subject to point-to-point fluctuations, while for all other groups these are one-σ Gaussian intervals determined from a Hessian PDF representation.

The leftmost vertical line indicates $x_{\min} = m_{\ell\bar{\ell}}^2/s$, the smallest allowed value of x for dilepton DY production with invariant mass $m_{\ell\bar{\ell}} = 5$ TeV for a collider CoM energy $\sqrt{s} = 14$ TeV. The rightmost vertical line corresponds to the value of x in a symmetric partonic collision where $x_1 = x_2$, namely $x_{\text{sym}} \equiv m_{\ell\bar{\ell}}/\sqrt{s}$.

From Fig. 8 one can observe that for $x \lesssim 0.3$ there is a clear hierarchy $f_u^+ > f_d^+ > f_s^+ > f_c^+$, while for larger x values the strange and charm PDFs become of comparable magnitude. The up and down quarks, both for xf_q^+ and xf_q^-, are significantly larger than the second-generation quark PDFs until $x \simeq 0.7$, and hence dominate the large-$m_{\ell\bar{\ell}}$ differential distributions in Drell–Yan production. PDF uncertainties
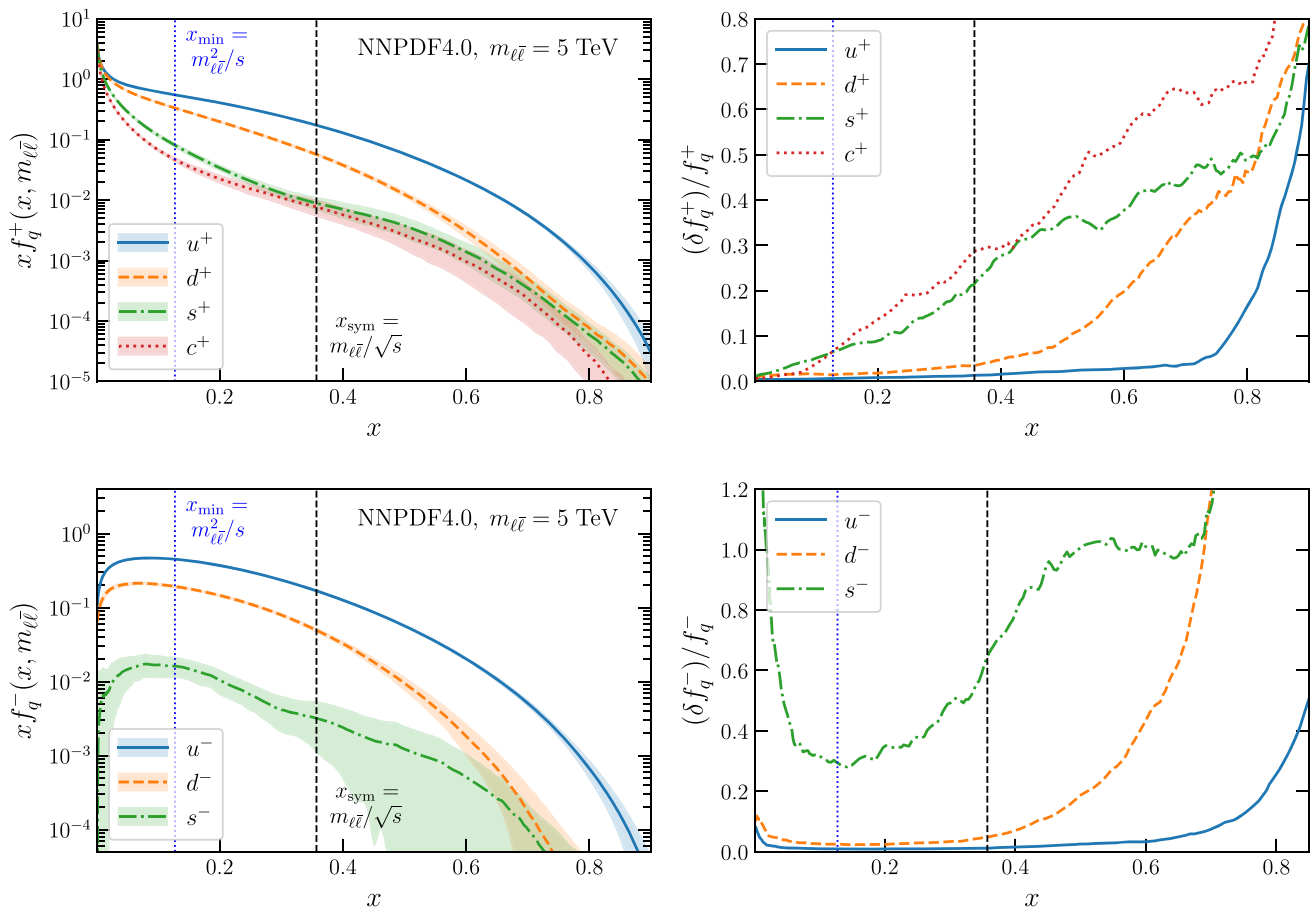
Fig. 8 Comparison of the xf_q^+ (top) and xf_q^- (bottom) quark PDF combinations for the up, down, strange, and charm quarks, evaluated at $m_{\ell\bar{\ell}} = 5$ TeV for NNPDF4.0 NNLO. The right panels display the relative 68% CL uncertainties. The two vertical lines indicate $x_{\min} = m_{\ell\bar{\ell}}^2/s$, the smallest allowed value of x for dilepton DY production for a collider CoM energy $\sqrt{s} = 14$ TeV, and the value of x corresponding to a symmetric partonic collision $x_1 = x_2$, namely $x_{\text{sym}} = m_{\ell\bar{\ell}}/\sqrt{s}$.

grow rapidly with x, reflecting the lack of direct experimental constraints. The same qualitative behavior of the lighter versus heavier flavor PDFs is observed for other PDF sets. Given the hierarchy $f_u^\pm, f_d^\pm \gg f_s^\pm, f_c^\pm$, in the following we will discuss only the behavior of the first-generation quark and antiquark PDFs which are those relevant for the interpretation of neutral-current Drell–Yan production in the kinematic region used for BSM searches.

We next compare the large-x behavior of the four PDF sets ABMP16, CT18, MSHT20, and NNPDF4.0 in Fig. 9 for $m_{\ell\bar{\ell}} = 5$ TeV. We display from top to bottom the absolute PDFs, their ratio to the central NNPDF4.0 value, and their relative 68% CL uncertainties. As in the case of Fig. 8, we indicate with two vertical lines the values of $x_{\min}$ and x_{sym}, both for $m_{\ell\bar{\ell}} = 5$ TeV, and for a smaller and a larger value of $m_{\ell\bar{\ell}}$, namely for $m_{\ell\bar{\ell}} = 3$ TeV and $m_{\ell\bar{\ell}} = 7$ TeV. For clarity, the values of $x_{\min}$ are only shown in the top row of plots, and the values of x_{sym} in the central row. Note that the scale dependence of the PDFs in this range of x and invariant mass is very slight. Indeed, the PDFs shown in Fig. 8 are

essentially unchanged at $m_{\ell\bar{\ell}} = 3$ TeV or $m_{\ell\bar{\ell}} = 7$ TeV; only the corresponding ranges of x_1, x_2 vary significantly.

Good agreement between all PDF sets is found up to around $x \simeq 0.4$. For $m_{\ell\bar{\ell}} = 5$ TeV this corresponds to the value of x_{sym}, i.e. central rapidity. For larger values of $x \gtrsim 0.4$, the up quark PDF xf_u from the NNPDF4.0 set is somewhat suppressed in comparison to the other three sets, which in turn agree among each other. A rather stronger suppression of NNPDF4.0 in comparison to CT18 is observed for the down quark, with MSHT20 and ABMP16 in a somewhat intermediate situation. The opposite behavior is found in the same region $x \gtrsim 0.4$ for antiquark PDFs $xf_{\bar{u}}$ and $xf_{\bar{d}}$: namely, the NNPDF4.0 PDF is significantly larger than that of the other sets. It follows that for a lower invariant mass value $m_{\ell\bar{\ell}} = 3$ TeV, all PDF sets are in agreement in the x range in which they are probed, while for a higher value $m_{\ell\bar{\ell}} = 7$ TeV the disagreement between NNPDF4.0 and the other PDF sets is present for most of the $x \geq x_{\min}$ range.

It is interesting to observe that in the region with $0.4 \lesssim x \lesssim 0.6$ the PDFs are constrained by some fixed-target DIS
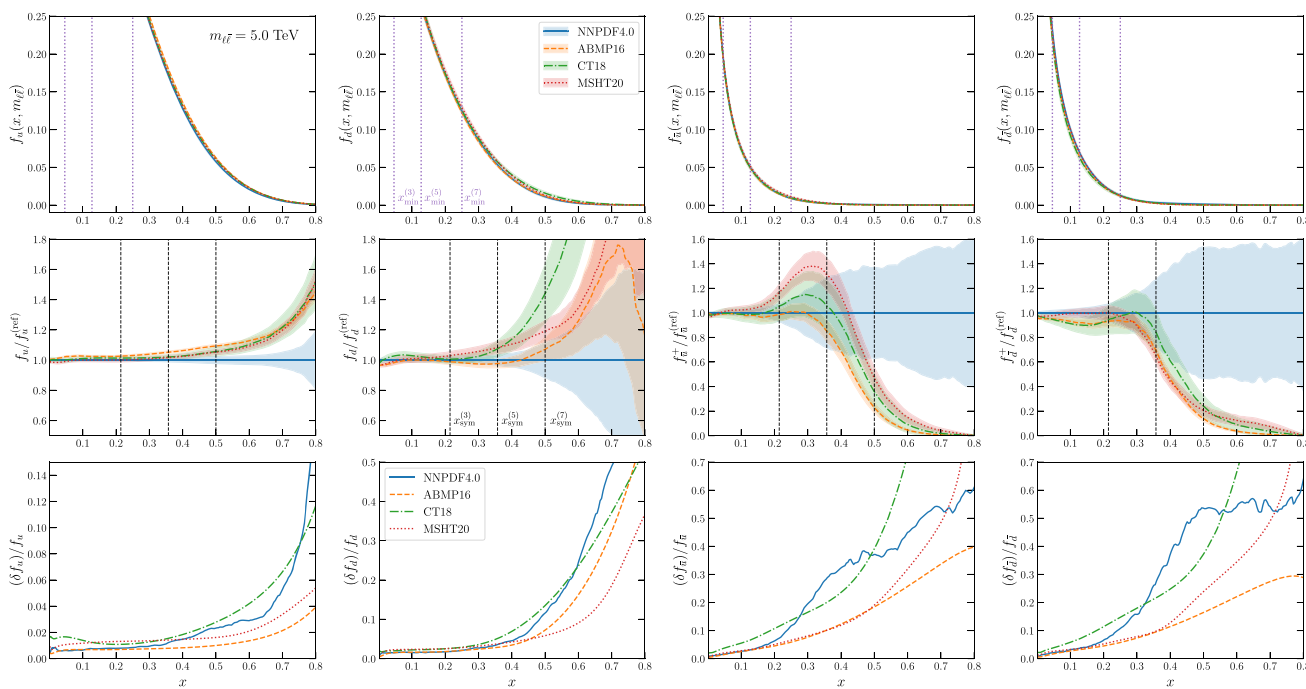
Fig. 9 The up and down quark and antiquark PDFs evaluated at $m_{\ell\bar{\ell}} = 5$ TeV for NNPDF4.0, CT18, MSHT20, and ABMP16 in the x region relevant for high-mass Drell–Yan production. The upper panels display the absolute PDFs, the middle ones their ratio to the central NNPDF4.0 value, and the bottom panels the relative 68% CL uncertainties. The vertical lines in the top row indicate the values of $x_{\min} = m_{\ell\bar{\ell}}^2/s$ and in the central row those of $x_{\text{sym}} = m_{\ell\bar{\ell}}/\sqrt{s}$ for three different values $m_{\ell\bar{\ell}} = 3, 5, 7$ TeV. Note that in the second row the range on the y axis is not the same for quarks and antiquarks, and in the third row also for up and down quarks. Note also that the PDFs, their ratios and their uncertainties are essentially unchanged in the displayed large-x region in the range 1 TeV $< m_{\ell\bar{\ell}} < 7$ TeV

structure functions and by forward W and Z production data from LHCb. Hence, at the edge of the data region NNPDF4.0 starts disagreeing with the other global PDF sets considered here, with the disagreement getting more marked as x grows outside the region covered by the data. Qualitatively, NNPDF4.0 is characterized by the fact that the quark PDFs drop faster as a function of x, and the antiquark PDFs drop less fast as x grows towards $x = 1$. As we will show next, this feature will lead to significant differences in the antisymmetric PDF luminosities $\mathcal{L}_{A,q}$ as the value of the dilepton invariant mass $m_{\ell\bar{\ell}}$ is increased.

The relative PDFs uncertainties, shown in the lower panels in Fig. 9 in all cases grow with x (see also Fig. 8). The largest PDF uncertainties correspond to either CT18 or NNPDF4.0, depending on the x range and the PDF flavor. Specifically, the NNPDF4.0 uncertainties are largest for f_d in the region $x \gtrsim 0.6$ and for $f_{\bar{u}}$ and $f_{\bar{d}}$ when $0.3 \lesssim x \lesssim 0.5$. The smallest PDF uncertainties are displayed by ABMP16 and MSHT20.

The different behavior of the rate of decrease with x of PDFs in the large x region, specifically comparing NNPDF4.0 to other PDF sets, can be seen most clearly from a comparison off effective asymptotic exponents [44]

$$\beta_{a,q}(x, Q) \equiv \frac{\partial \ln |x f_q(x, Q)|}{\partial \ln(1-x)}, \qquad (3.6)$$

which of course for PDFs of the form of Eq. (3.1) just coincide with the exponent b up to $O(1 - x)$ corrections. In Fig. 10 we compare the values of $\beta_{a,q}(x, m_{\ell\bar{\ell}})$ for ABMP16, CT18, MSHT20, and NNPDF4.0 evaluated at $m_{\ell\bar{\ell}} = 5$ TeV for the up and down quark and antiquark PDFs in the x range of Fig. 8.

It is clear that while all PDF sets have a similar effective asymptotic exponent for $x \lesssim 0.35$, a different behavior of NNPDF4.0 in comparison to other determinations sets in for $x \gtrsim 0.4$. Specifically, for quarks the NNPDF4.0 exponents are always larger, and for antiquarks smaller than those found with other PDF sets. Interestingly, whereas for the up quark the effective exponent $\beta_{a,u}$ is approximately constant for all PDF sets when $x \gtrsim 0.4$, with the NNPDF4.0 value being just slightly higher and slowly increasing, for the down quark and all antiquarks this approximately constant behavior is seen for other PDF sets but not for NNPDF4.0. Specifically, for the NNPDF4.0 down quark the exponent slowly but markedly increases for $x \gtrsim 0.3$, together with its uncertainty. In the case of NNPDF4.0 for both antiquarks the exponent rapidly drops in the region $0.3 \lesssim x \lesssim 0.4$. This is consistent with the observation at the PDF level (Fig. 9) that for NNPDF4.0 at large-x, as compared to the other groups, the up and especially the down quark fall off more rapidly, while the antiquark PDFs drop more slowly. Note in particular that for the
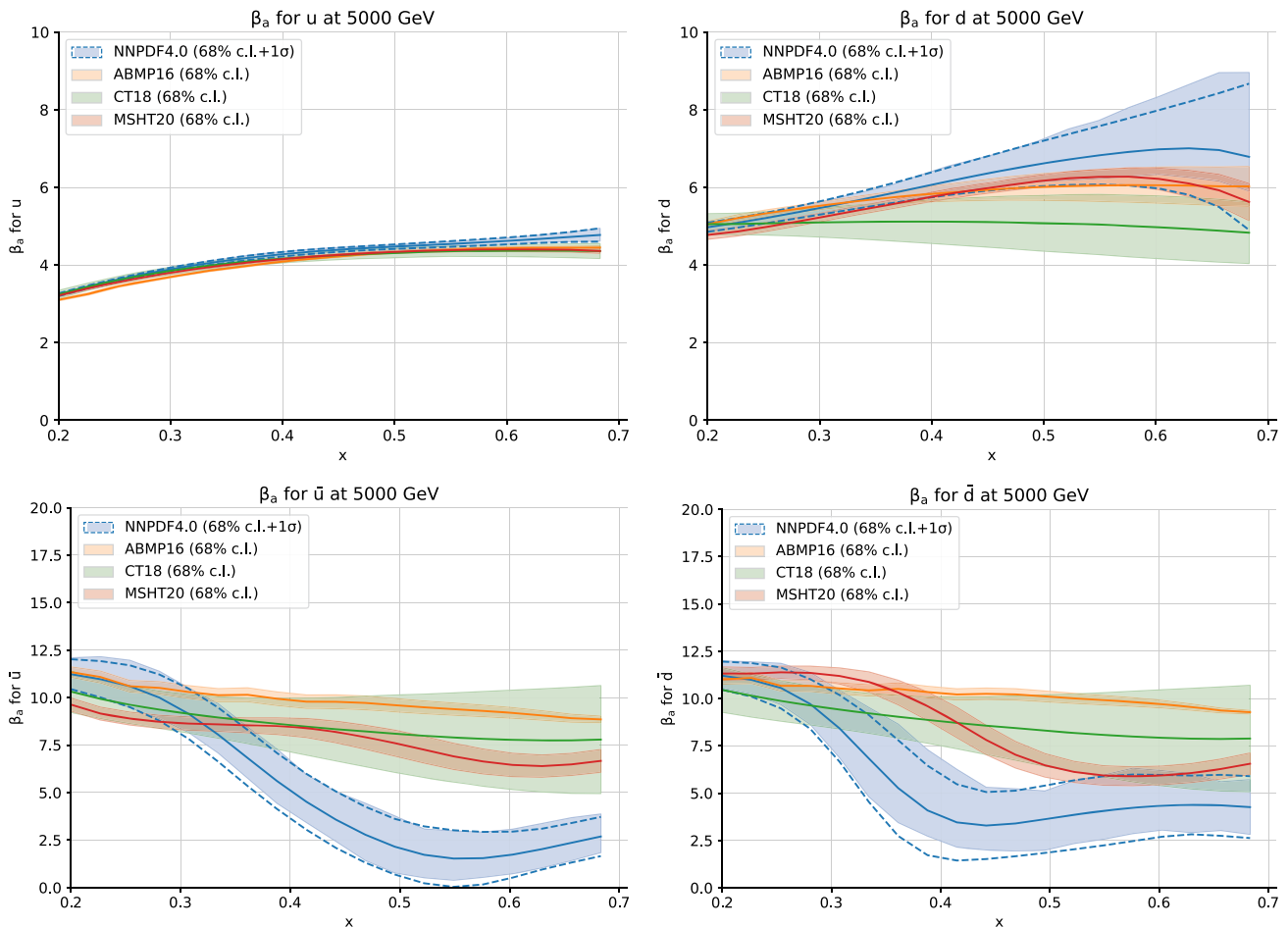
Fig. 10 The large-x asymptotic exponents $\beta_{a,q}(x, m_{\ell\bar{\ell}})$, defined in Eq. (3.6), for ABMP16, CT18, MSHT20, and NNPDF4.0 evaluated at $m_{\ell\bar{\ell}} =$ 5 TeV for the up and down quark and antiquark PDFs

down PDF the antiquark effective exponent is significantly smaller than the quark effective exponent for all $x \gtrsim 0.4$.

The fact that a modification in behavior of the effective down quark and especially antiquark PDFs is observed at the edge of the data region for NNPDF4.0, but not for other PDF sets, suggests that this might be related to the fact that NNPDF4.0 generally adopts a more flexible PDF parametrization in comparison to other groups. Conversely, the fact that other groups display similar behaviors suggests that this is related to their common choice of parametrizing the large x behavior of PDFs as $(1-x)^{\beta_i}$, with the exponents β_i fixed for each PDF flavor or combination of flavors. Also, the uncertainties on the effective exponents $\beta_{a,q}(x, m_{\ell\bar{\ell}})$ tend to be larger for NNPDF4.0 (and also to a lesser extent for CT18) in comparison to those of other groups. Note however that the full PDF uncertainty contains also a contribution from the overall magnitude, which is not captured by the effective exponents displayed here.

3.3 Parton luminosities

We finally turn to the behavior of parton luminosities, with particular regard for the antisymmetric combination which is relevant for the forward–backward asymmetry. As for PDFs, we first assess the qualitative features of the luminosities corresponding to different quark flavors. Specifically, the symmetric $\mathcal{L}_{S,q}$ and antisymmetric $\mathcal{L}_{A,q}$ luminosities Eq. (2.12) for individual flavors are displayed in Fig. 11, evaluated with NNPDF4.0 NNLO for $m_{\ell\bar{\ell}} = 5$ TeV and $\sqrt{s} = 14$ TeV. The left panels display the absolute luminosities (in logarithmic and linear scale respectively for the y and x axes) while the right panels show the corresponding PDF uncertainties (relative and absolute for $\mathcal{L}_{S,q}$ and $\mathcal{L}_{A,q}$, respectively). The bottom and top x-axes in each plot show respectively the values of x_1 and x_2 at which the luminosities are being evaluated, within the allowed range $x \geq x_{\text{sym}} = m_{\ell\bar{\ell}}/\sqrt{s}$, with the convention $x_1 > x_2$.

The symmetric parton luminosities exhibit of course the same hierarchy between flavors as the corresponding PDF
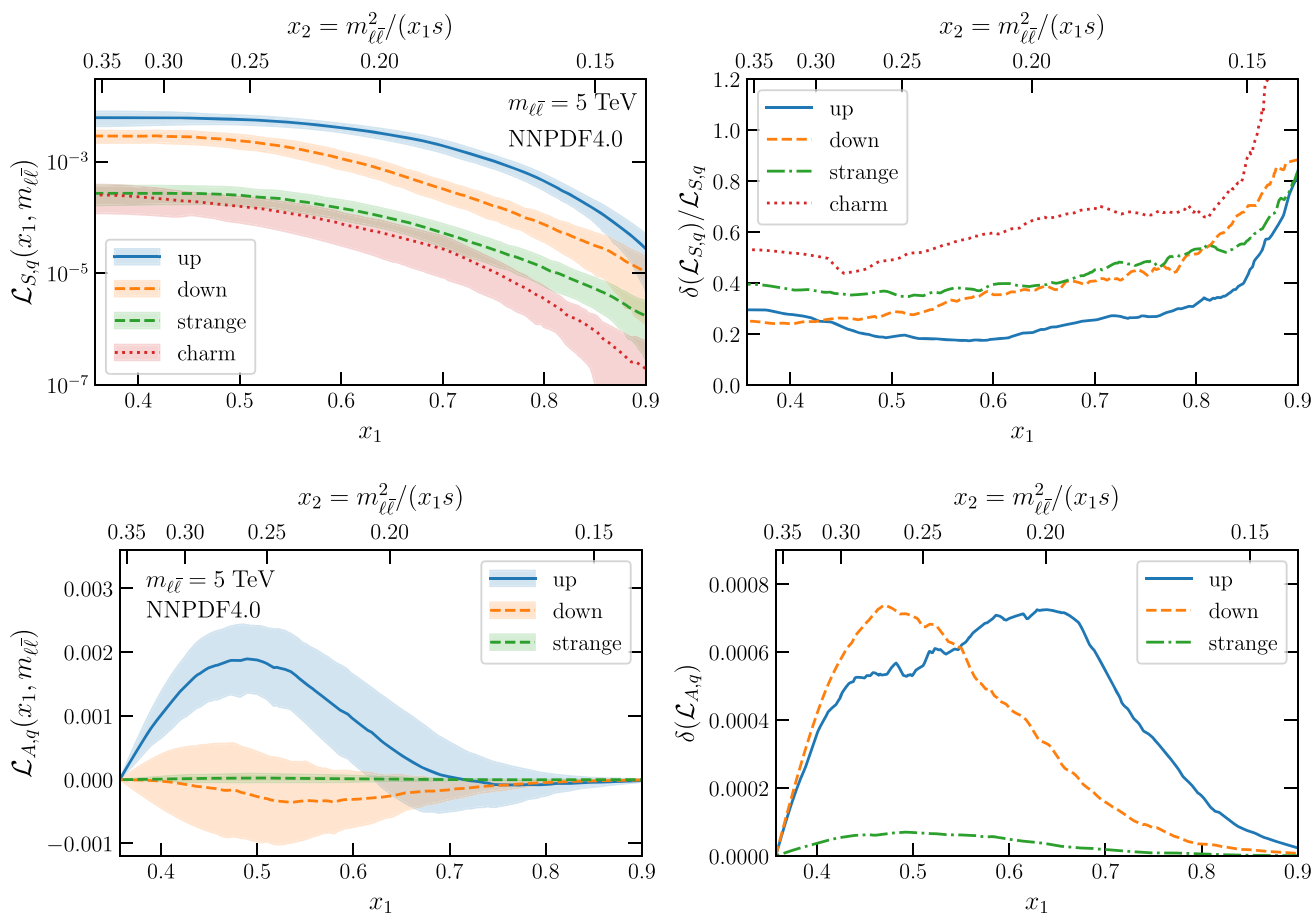
Fig. 11 The symmetric $\mathcal{L}_{S,q}$ (top) and antisymmetric $\mathcal{L}_{A,q}$ (bottom) parton luminosities (left) and relative uncertainties (right) evaluated with NNPDF4.0 NNLO at $m_{\ell\bar{\ell}} = 5$ TeV and $\sqrt{s} = 14$ TeV. The bottom and top x-axes in each plot show respectively the values of x_1 and x_2 at which the luminosities are being evaluated, within the allowed range $x \geq x_{\text{sym}} = m_{\ell\bar{\ell}}/\sqrt{s}$, with the convention $x_1 > x_2$

plots of Fig. 8. The luminosity $\mathcal{L}_{S,q}$ drops rapidly for $x_1 \gtrsim 0.6$. PDF uncertainties depend weakly on x up to $x_1 \gtrsim 0.8$, after which they blow up, and range between $\sim 20\%$ for the up quark luminosity to $\sim 60\%$ for the charm quark one, with down and strange intermediate and of similar magnitude.

As displayed in Fig. 12, the light quark symmetric luminosities of other global PDF sets are qualitatively similar. We show $\mathcal{L}_{S,u}$, $\mathcal{L}_{S,d}$, and their weighted sum that enters the enters the symmetric coefficient $g_{S,q}$ in Eq. (2.18) for the NNPDF4.0, ABMP16, CT18, and MSHT20 at $m_{\ell\bar{\ell}} = 5$ TeV. The luminosities are multiplied by the effective charges S_q defined in Eq. (2.8), and the bottom panels display the corresponding 68% CL PDF uncertainties. Good agreement between the four sets, with a similar shape of $\mathcal{L}_{S,q}$, is observed. The PDF luminosities for the dominant $\mathcal{L}_{S,u}$ contribution are the largest for NNPDF4.0.

Turning to the antisymmetric PDF luminosities $\mathcal{L}_{A,q}$, we note that, for NNPDF4.0, while the up luminosity is positive, the central value of the down luminosity is negative, though the luminosity is compatible with zero at the one sigma level.

Recalling from Fig. 8 that $x f_d^-$ itself is positive for all values of x, this provides an explicit example in which the condition Eq. (3.4) is satisfied without the valence combination being negative. We conclude that for NNPDF4.0, the faster drop of the quark distribution and slower drop of the antiquark distribution that was displayed by the effective exponents of Fig. 10 leads to a negative antisymmetric luminosity, in agreement with Eq. (3.4). The absolute PDF uncertainties are of a similar size for $\mathcal{L}_{A,u}$ and $\mathcal{L}_{A,d}$, with a different shape reflecting the underlying central values.

We compare in Fig. 13 the behavior of the antisymmetric luminosities for all PDF sets for $m_{\ell\bar{\ell}} = 3$ TeV (top) and $m_{\ell\bar{\ell}} = 5$ TeV (bottom). In order to facilitate the understanding of the way the PDF behavior determines that of the asymmetry, we show both the contribution of individual flavors and the total contribution to the antisymmetric coefficient $g_{A,q}$ of Eq. (2.19). Namely, in Fig. 13 the luminosities corresponding to individual flavors are multiplied by the corresponding flavor-dependent effective charges A_q defined in Eq. (2.8): from left to right we display $\mathcal{L}_{A,u}$, $\mathcal{L}_{A,d}$, and their weighted
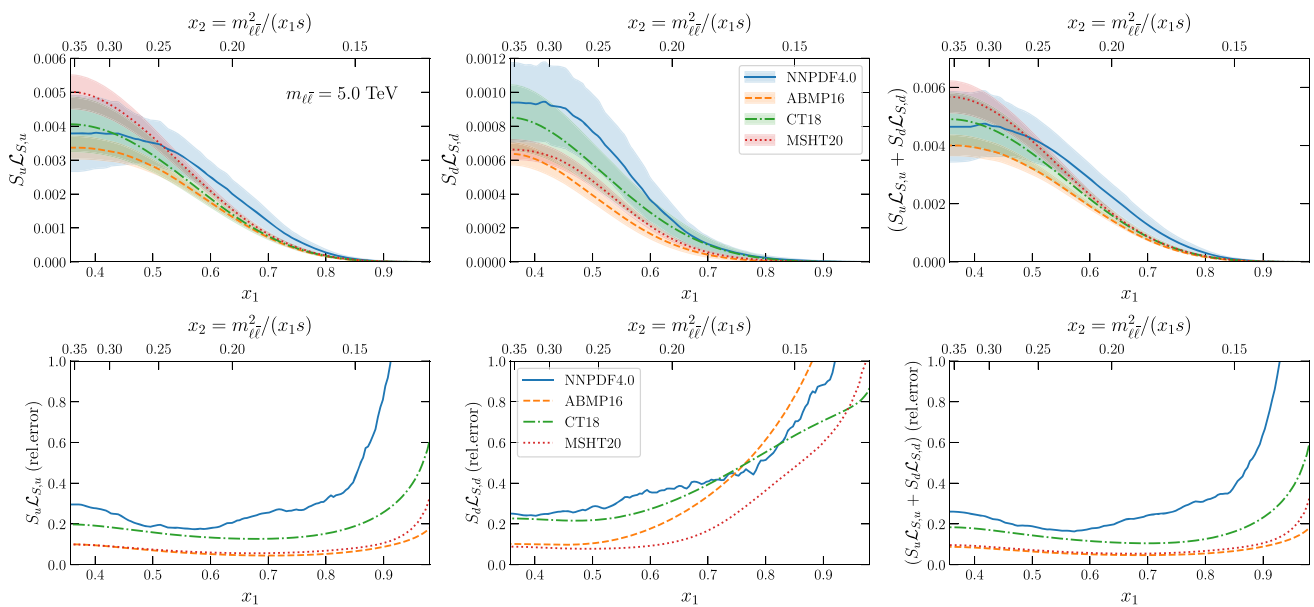
Fig. 12 The symmetric parton luminosities $\mathcal{L}_{S,q}(x_1, m_{\ell\bar{\ell}})$ for the NNPDF4.0, ABMP16, CT18, and MSHT20 NNLO PDF sets for dilepton invariant masses of $m_{\ell\bar{\ell}} = 5$ TeV. The luminosities are multiplied by the effective charges S_q defined in Eq. (2.8). From left to right, we display $\mathcal{L}_{S,u}$, $\mathcal{L}_{S,d}$, and their weighted sum that enters the coefficient $g_{S,q}$ in Eq. (2.18). The bottom panels display the relative 68% CL PDF uncertainties.

sum which determines the sign and magnitude of the total forward–backward asymmetry. The corresponding absolute PDF uncertainties for each of the four PDF sets are displayed in Fig. 14.

Figure 13 shows that for ABMP16, CT18, and MSHT20 the antisymmetric parton luminosities depend only mildly on $m_{\ell\bar{\ell}}$, whereas for NNPDF4.0 they exhibit a strong $m_{\ell\bar{\ell}}$ dependence. Indeed, for dilepton invariant masses of $m_{\ell\bar{\ell}} = 3$ TeV there is good agreement between the three groups, but for $m_{\ell\bar{\ell}} = 5$ TeV the NNPDF4.0 up quark luminosity, while preserving a similar valence-like shape, is suppressed by a factor 2 in comparison to other groups, and the down quark luminosity becomes compatible with zero with a negative central value, as already noted. For all PDF sets and $m_{\ell\bar{\ell}}$ values the weighted sum is dominated by the up quark contribution. The strong scale dependence of $\mathcal{L}_{A,q}$ in NNPDF4.0 reflects the underlying PDF behavior seen in Fig. 9 and highlighted by the effective exponents Fig. 10. As the scale $m_{\ell\bar{\ell}}$ increases, a range of increasingly large x values is probed, for which, in the case of NNPDF4.0, the quark effective exponent slowly increases and the antiquark exponent rapidly drops. This leads to a negative asymmetry, following Eq. (3.4).

A comparison of the corresponding PDF uncertainties, displayed in Fig. 14, clearly shows the transition from the data region to the extrapolation region. For $m_{\ell\bar{\ell}} = 3$ TeV the uncertainty $\delta\mathcal{L}_{A,u}$ is generally small for all sets, with CT18 showing a somewhat larger uncertainty for the up quark, and comparable uncertainties for the down quark for all PDF

sets. As the scale increases to $m_{\ell\bar{\ell}} = 5$ TeV, where the large-x region is probed, the uncertainty increases, though more markedly for NNPDF4.0. For all PDF sets but NNPDF4.0, the uncertainty is approximately unchanged when the scale is further increased, while for NNPDF4.0 it grows markedly.

Finally, in Fig. 15 we display for all PDF sets the ratio of antisymmetric to symmetric couplings

$$R_{\mathrm{fb}} \equiv \frac{\sum_q g_{A,q}}{\sum_{q'} g_{S,q'}} , \qquad (3.7)$$

that, according to Eq. (2.23), determines at leading order the sign and magnitude of the forward–backward asymmetry distribution $A_{\mathrm{fb}}(\cos\theta^*)$. The symmetric and antisymmetric coefficients are obtained by integrating the corresponding symmetric $\mathcal{L}_{S,q}$ and antisymmetric $\mathcal{L}_{A,q}$ partonic luminosities according to Eq. (2.19), and the result is shown as a function of the lower integration cut $m_{\ell\bar{\ell}}^{\min}$. In all cases the correlation between PDF uncertainties in the numerator and the denominator are kept into account.

Figure 15 shows that, consistently with the behavior of the luminosity of Fig. 13, for $m_{\ell\bar{\ell}}^{\min} \lesssim 3$ TeV results agree within uncertainties for all PDF sets. The situation is different for higher dilepton invariant masses $m_{\ell\bar{\ell}}^{\min} \gtrsim 3$ TeV: the ratio R_{fb} starts to decrease for NNPDF4.0, while it remains approximately constant for the other PDF sets. In particular, for NNPDF4.0 the coupling ratio vanishes around $m_{\ell\bar{\ell}}^{\min} \sim 5$ TeV, and it becomes negative for yet larger $m_{\ell\bar{\ell}}^{\min}$ values. It follows that the forward–backward asymmetry in
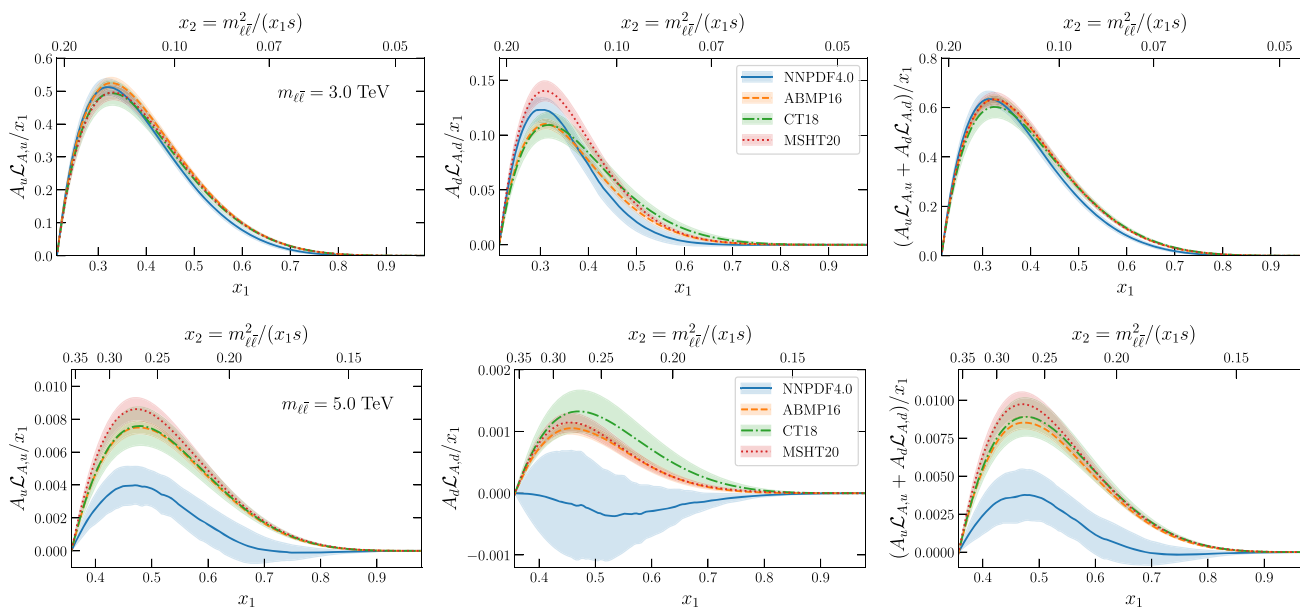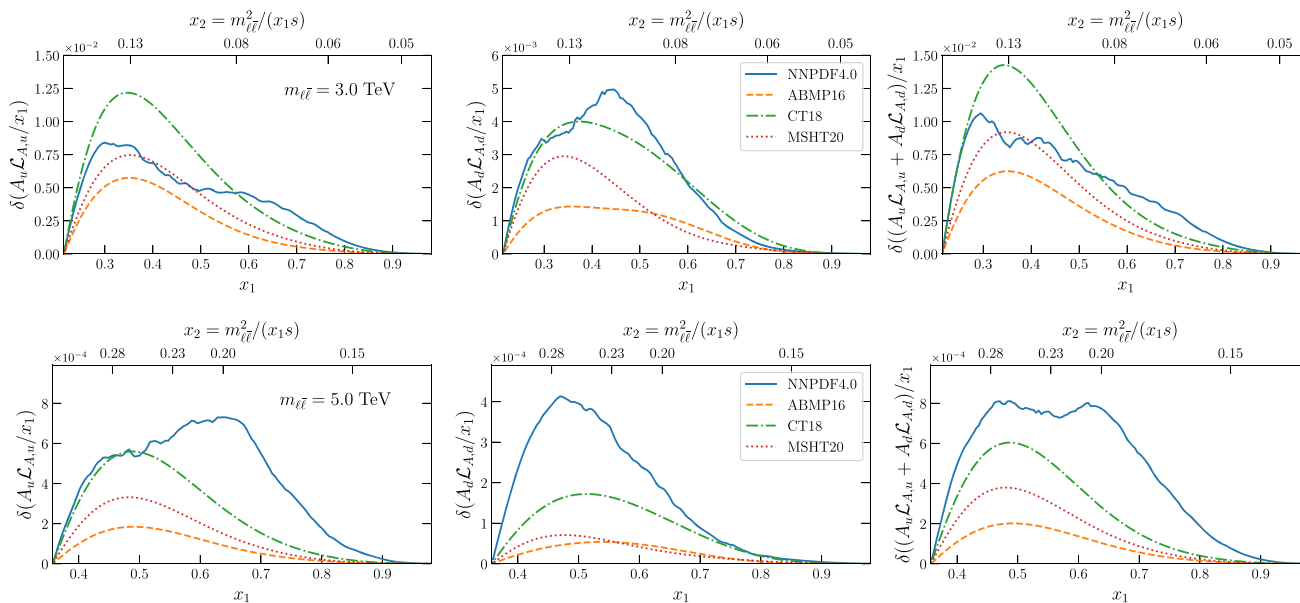
Fig. 13 The antisymmetric parton luminosities $\mathcal{L}_{A,q}(x_1, m_{\ell\bar{\ell}})$ for the NNPDF4.0, ABMP16, CT18, and MSHT20 NNLO PDF sets for dilepton invariant masses of $m_{\ell\bar{\ell}} = 3$ TeV (top) and $m_{\ell\bar{\ell}} = 5$ TeV (bottom). The luminosities are multiplied by the effective charges A_q defined in Eq. (2.8). From left to right, we display $\mathcal{L}_{A,u}$, $\mathcal{L}_{A,d}$, and their weighted sum that enters the coefficient $g_{A,q}$ Eq. (2.19)

Fig. 14 Same as Fig. 13 now for the absolute PDF uncertainties

high-mass Drell–Yan production should decrease and eventually vanish (and possibly even turn negative) in NNPDF4.0 as the $m_{\ell\bar{\ell}}^{\min}$ cut is increased, while for CT18, MSHT20, and ABMP16 it should remain positive with a similar magnitude irrespective of the cut $m_{\ell\bar{\ell}}^{\min}$ adopted.

Figure 16 displays the absolute and relative uncertainties associated to the coupling ratio $R_{\rm fb}$. We observe that NNPDF4.0 shows the most marked increase of the uncertainties in $R_{\rm fb}$ as $m_{\ell\bar{\ell}}^{\min}$ grows. For instance, for $m_{\ell\bar{\ell}}^{\min} \gtrsim 4$ TeV

the absolute PDF uncertainty in NNPDF4.0 is about twice as large as that found using CT18 four times as large as MSHT20, and about one order of magnitude larger than ABMP16. This trend is magnified for the relative uncertainties due to the decrease in the central value of $R_{\rm fb}$ as $m_{\ell\bar{\ell}}^{\min}$ increases.
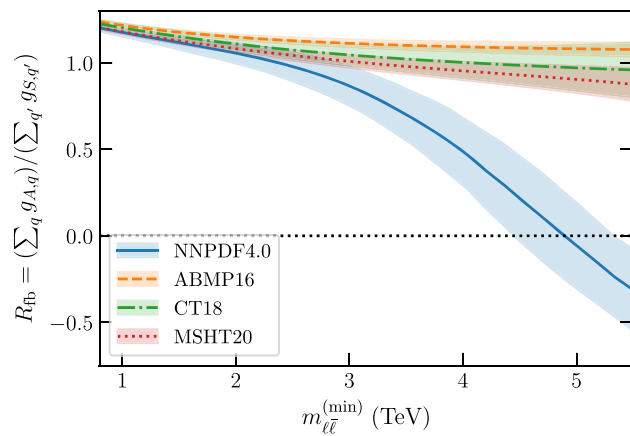
Fig. 15 The coupling ratio $R_{\rm fb}$, Eq. (3.7), that enters the forward–backward asymmetry $A_{\rm fb}(\cos\theta^*)$ at LO, Eq. (2.23), for different PDF sets, as a function of the lower cut in the dilepton invariant mass $m_{\ell\bar\ell}^{\min}$.

4 The Drell–Yan forward–backward asymmetry at the LHC

After the qualitative discussion of the previous sections, here we present results for the $\cos\theta^*$ distributions Eq. (2.17) and the forward–backward asymmetry Eq. (2.22), with NLO QCD and electroweak corrections included and with realistic selection and acceptance cuts for the LHC at $\sqrt{s} = 14$ TeV and different values of the invariant mass $m_{\ell\bar\ell}$ relevant for SM studies and BSM searches.

Computations are performed using MADGRAPH5_AMC@NLO [35], interfaced to PINEAPPL [36,37] to generate fast interpolation grids. In order to account for realistic detector acceptances, we impose phase-space cuts on the transverse momentum and the pseudo-rapidity of the two leading leptons,

$$p_T^\ell > 10\,{\rm GeV}\,, \qquad |\eta_\ell| < 2.4\,. \qquad (4.1)$$

We then consider various regions of dilepton invariant mass $m_{\ell\bar\ell}$: either close to the Z-boson peak (60 GeV $< m_{\ell\bar\ell} <$ 120 GeV), relevant for precision SM studies, or the high-mass region relevant for BSM searches, with various choices of a lower mass invariant cutoff ($m_{\ell\bar\ell} > 3, 4, 5, 6$ TeV). In all cases, in order to facilitate the interpretation of hadron-level results and the connection to the discussion of the PDF features from Sect. 3, we also provide results for the two partonic channels that give the largest contribution to the cross-section. As in Sect. 3, we compare results obtained using the ABMP16, CT18, MSHT20, and NNPDF4.0 PDF sets. In all cases, we use the NNLO sets corresponding to the value $\alpha_s(m_Z) = 0.118$ of the strong coupling. Results obtained using the NNPDF3.1 PDF set are reported in App. A.

Before considering the angular distributions, in Fig. 17 we display the differential distribution in absolute dilepton rapidity $|y_{\ell\bar\ell}|$, defined in Eq. (2.21), for a dilepton invariant

mass of $m_{\ell\bar\ell} > 5$ TeV. This is the kinematic region relevant for searches of high-mass resonances in the dilepton channel at the LHC, e.g. [13,45]. We display the absolute differential distributions with the 68% CL PDF uncertainties (top), the relative PDF uncertainty (center) normalized for each PDF set to the corresponding central prediction, and the pull between the NNPDF4.0 result, taken as a reference, and other sets (bottom). This pull is defined as

$$\text{Pull}_i = \frac{\sigma_{2,i}^{(0)} - \sigma_{1,i}^{(0)}}{\sqrt{\left(\delta\sigma_{2,i}\right)^2 + \left(\delta\sigma_{1,i}\right)^2}}\,, \qquad i = 1,\ldots,n_{\rm bin}\,, \quad (4.2)$$

where $\sigma_{1,i}^{(0)}$ and $\sigma_{2,i}^{(0)}$ are the central values of the theory prediction in the i-th bin of the distribution and $\delta\sigma_{1,i}$, $\delta\sigma_{2,i}$ are the corresponding PDF uncertainties. For the central NNPDF4.0 prediction in the upper panel we also display the contributions from the dominant parton subchannels, namely $u\bar u + c\bar c$ and $d\bar d + s\bar s + b\bar b$.

As discussed in Sect. 2, the $|y_{\ell\bar\ell}|$ distribution depends on the symmetric partonic luminosities $\mathcal{L}_{S,q}$, Eq. (2.15), which in turn are driven by the total PDFs xf_q^+. The $|y_{\ell\bar\ell}|$ distribution is dominated by the $u\bar u$ contribution and its qualitative behavior is found to be similar for the four PDF sets considered. PDF uncertainties are the largest in NNPDF4.0, ranging between 25% and 50%, and the pull between NNPDF4.0 and CT18 and MSHT20 is at most at the 1.5σ level, and slightly larger with ABMP16. The dependence of the $|y_{\ell\bar\ell}|$ distribution on the dilepton mass $m_{\ell\bar\ell}$ is moderate, and the same qualitative features are obtained if $m_{\ell\bar\ell}$ is lowered down to the Z-peak region, or raised to yet higher values. Hence, for the absolute rapidity distribution there is a reasonable agreement between all PDF sets for all scales considered.

We now turn to the differential distribution in $\cos\theta^*$ and the corresponding forward–backward asymmetry $A_{\rm fb}(\cos\theta^*)$. We first consider the Z-peak region, 60 GeV $< m_{\ell\bar\ell} <$ 120 GeV, in Fig. 18. The $\cos\theta^*$ distribution exhibits a small but non-negligible asymmetry, and uncertainties are smallest for NNPDF4.0. The four PDF sets predict a similar behavior and magnitude of the asymmetry $A_{\rm fb}$. PDF uncertainties in the asymmetry are comparable for all PDF sets when $\cos\theta^* \approx 0$, and actually largest for NNPDF4.0 when $\cos\theta^* \approx 1$. In all cases the predictions are compatible within 2σ, with ABMP16 showing larger differences of up to 2.8σ for the $\cos\theta^*$ distribution. Note that the sharp drop-off at the edges $|\cos\theta^*| \approx 1$, appearing in all plots in this section, is a consequence of the phase-space cuts which limit the phase-space volume. Indeed, using LO kinematics

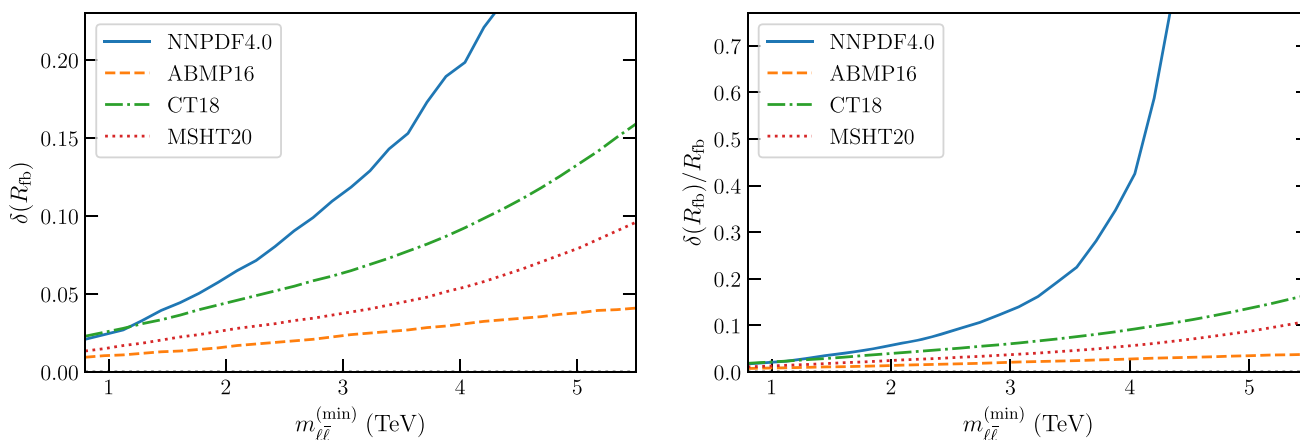$$|\cos\theta^*| = \tanh\left|\frac{\eta_\ell - \eta_{\bar\ell}}{2}\right| = \sqrt{1 - \frac{4(p_T^\ell)^2}{m_{\ell\bar\ell}^2}}, \qquad (4.3)$$

Fig. 16 The absolute (left) and relative (right panel) uncertainties in the coupling ratio R_{fb} shown in Fig. 15.
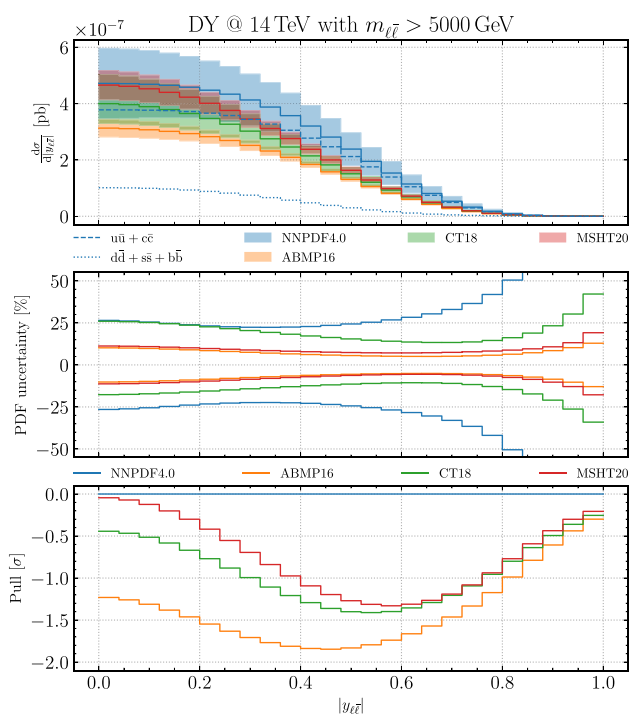

Fig. 17 The differential distribution in absolute dilepton rapidity $|y_{\ell\bar{\ell}}|$, given in Eq. (2.21), for dilepton invariant masses of $m_{\ell\bar{\ell}} > 5$ TeV for neutral current Drell–Yan production at the LHC 14 TeV, obtained using ABMP16, CT18, MSHT20, and NNPDF4.0 NNLO PDFs with $\alpha_s(m_Z) = 0.118$. All uncertainties shown are 68% CL PDF uncertainties, computed at NLO QCD with realistic cuts (see text). We show the absolute distributions (top), relative uncertainties (normalized to the central curve of each set, middle) and the pull with respect to the NNPDF4.0 result, Eq. (4.2) (bottom). For the central NNPDF4.0 prediction the contributions of the $u\bar{u}+c\bar{c}$ and $d\bar{d}+s\bar{s}+b\bar{b}$ parton subchannels are also shown.

so $|\cos\theta^*| \approx 1$ requires a lepton pair with either a large rapidity separation, or a very large invariant mass and small transverse momenta.

As expected from the antisymmetric partonic luminosities studied in Sect. 3.3, the situation is quite different when considering distributions with a higher dilepton invariant mass range. The angular distribution and forward–backward asymmetry in the high-mass region, for different values of the lower cut in the dilepton invariant mass, namely $m_{\ell\bar{\ell}}^{\mathrm{min}} = 3, 4, 5$ and 6 TeV, are respectively shown in Figs. 19 and 20.

Consistent with the underlying parton luminosities, the $\cos\theta^*$ distribution is dominated by $u\bar{u}$ scattering, while $d\bar{d}$ provides a subdominant contribution. When the lower cut is $m_{\ell\bar{\ell}}^{\mathrm{(min)}} = 3$ TeV is used, the four PDF sets are in agreement at the 1σ level: they all display a positive forward–backward asymmetry, and exhibit PDF uncertainties ranging between 10% and 15%. As the invariant mass cut is raised, the qualitative behavior of the angular distribution and asymmetry change substantially for NNPDF4.0, while they remain approximately the same for all other PDF sets, consistent with the behavior of the PDFs and luminosities discussed in Sects. 3.2 and 3.3. Specifically, raising the cut to $m_{\ell\bar{\ell}} \geq 4$ TeV, for NNPDF4.0 the backwards cross-section starts increasing, though the asymmetry remains positive.

For $m_{\ell\bar{\ell}} \geq 5$ TeV the central value of the NNPDF4.0 $\cos\theta^*$ distribution becomes symmetric, though the PDF uncertainty band is rather asymmetric. Also, PDF uncertainties are now the largest for NNPDF4.0, reaching up to 30%. Finally, for $m_{\ell\bar{\ell}} \geq 6$ TeV the central value of forward–backward asymmetry for NNPDF4.0 becomes negative, with the PDF uncertainties increasing further so the asymmetry remains compatible with zero at about the 1.1 σ level. For all other PDF sets there is little change in the shape of the distribution as the dilepton invariant mass cut is increased. Because of the large uncertainty on the NNPDF4.0 result for the $\cos\theta^*$ distribution, even with the highest value of the $m_{\ell\bar{\ell}}^{\mathrm{(min)}}$ cut, where NNPDF4.0 finds a symmetric distributions while all other PDF sets find an asymmetry, the pull is always below 2σ.
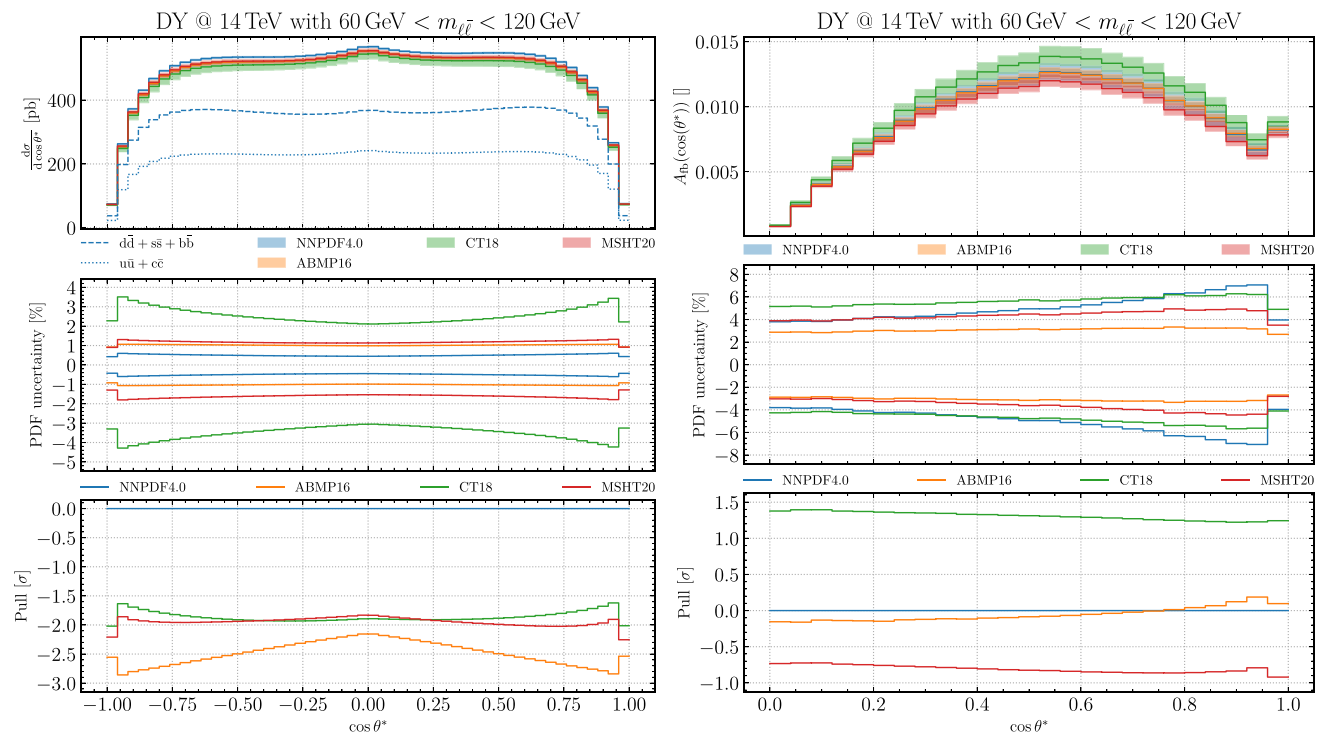
Fig. 18 Same as Fig. 17, now for the differential distribution in $\cos\theta^*$ (left) and the corresponding forward–backward asymmetry $A_{\rm fb}(\cos\theta^*)$ (right), in the Z-peak region defined by 60 GeV $< m_{\ell\bar{\ell}} < 120$ GeV

In Fig. 21 the forward–backward asymmetry with $m_{\ell\bar{\ell}}^{(\rm min)} = 5$ TeV shown in Fig. 20 (bottom left) is shown again, now also including a prediction obtained using the PDF4LHC21 combination of parton distributions [46], specify its compressed Monte Carlo representation [47], based on the CT18, MSHT and NNPDF3.1 PDF sets. Because the PDF uncertainties for NNPDF3.1 are generally, and specifically at large x, rather larger than those on NNPDF4.0 (see also App. A) the uncertainty on the $A_{\rm fb}$ distribution found using the PDF4LHC21 combination is extremely large, and no signal for the asymmetry can be seen. PDF4LHC recommends [46] usage of the combination for BSM searches, and that of individual PDF sets for comparison between data and theory for SM measurements. The results presented here suggest that the uncertainty estimate of NNPDF4.0 in the extrapolation region, which is rather more conservative than that of the other PDF sets shown here, might be desirable and lead to more robust predictions for the forward–backward asymmetry in the high-mass region which is relevant for new physics searches.

5 Summary and outlook

In this work we have scrutinised the PDF dependence of neutral current Drell–Yan production at large dilepton invariant masses $m_{\ell\bar{\ell}}$, focusing on the behavior of the forward–backward asymmetry $A_{\rm fb}$ in the Collins–Soper

angle $\cos\theta^*$, an observable frequently considered in the context of searches for new physics beyond the SM. We have demonstrated that while theoretical predictions for the sign and magnitude of $A_{\rm fb}$ are very similar for all PDF sets in the Z peak region, they depend markedly on the choice of PDF set for large values of $m_{\ell\bar{\ell}}$. We have traced this behavior to that of the PDFs, which agree in the data region, but differ in the large-x region, where PDFs are mostly unconstrained by data.

We have specifically shown that the uncertainty on the asymmetry differs substantially between PDF sets, with NNPDF4.0 displaying a more marked increase as $m_{\ell\bar{\ell}}$ grows, leading to an absolute uncertainty that e.g. for $m_{\ell\bar{\ell}}^{\rm min} \gtrsim 4$ TeV is about twice as large as that found using CT18, four times as large as MSHT20, and about one order of magnitude larger than ABMP16. Also, whereas other PDF sets predict a shape of the asymmetry which is unchanged when $m_{\ell\bar{\ell}}$ increases from the Z-peak region to the TeV range, namely a positive asymmetry implying a larger cross-section for $\cos\theta^* \geq 0$, NNPDF4.0 finds that as $m_{\ell\bar{\ell}}$ increases, the asymmetry is reduced, and the $\cos\theta^*$ distribution becomes symmetric when $m_{\ell\bar{\ell}}^{\rm min} \sim 5$ TeV.

We have traced this behavior to that of the underlying PDFs in the large-x region, where PDFs are mostly unconstrained by data. Specifically we have seen that in this region NNPDF4.0 has generally wider uncertainties. Also, while for all PDF sets the quark and antiquark distributions vanish as
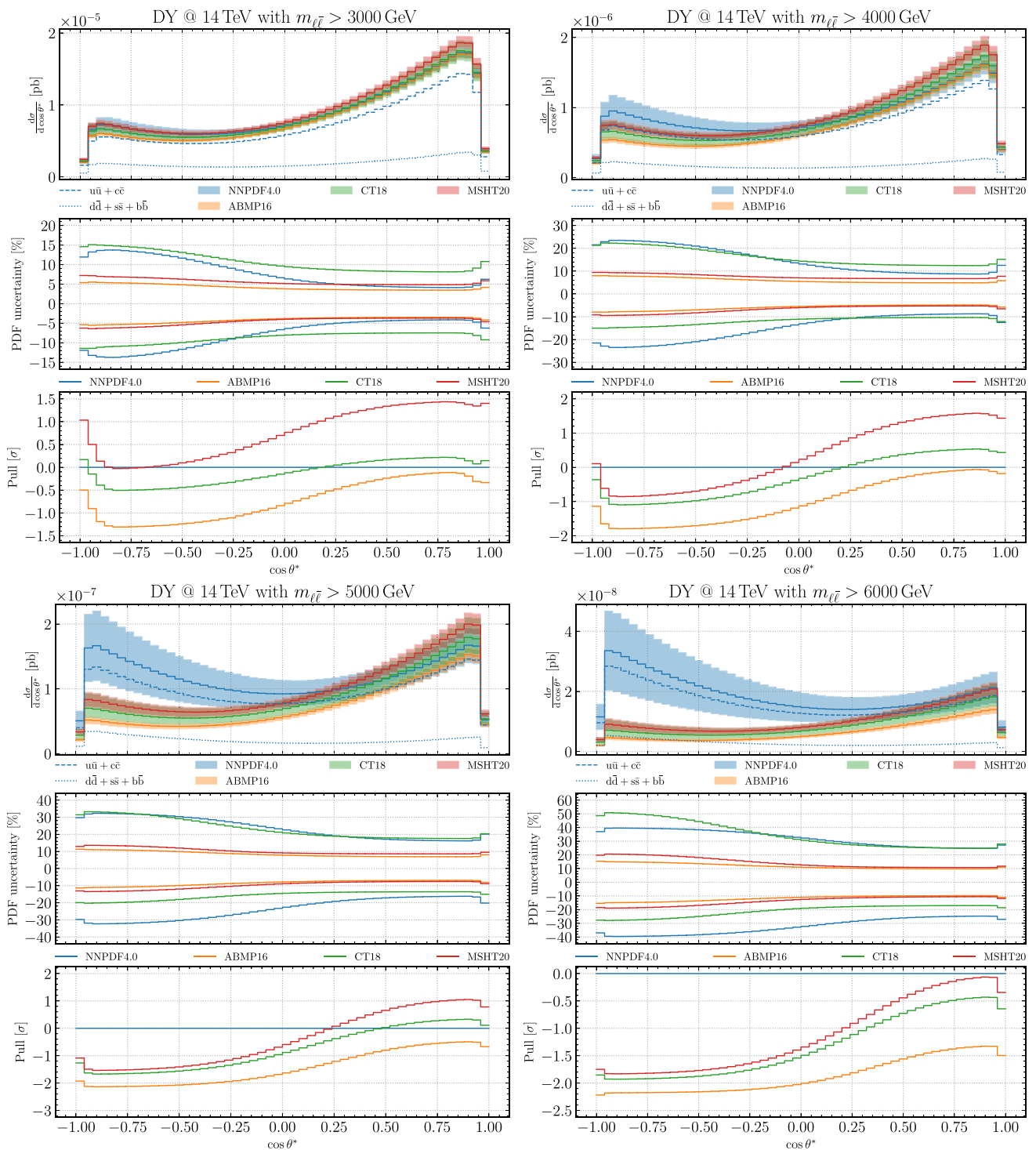
Fig. 19 Same as Fig. 18 (left) for different values of the lower cut in the dilepton invariant mass: $m_{\ell\bar{\ell}} \geq 3, 4, 5,$ and 6 TeV respectively
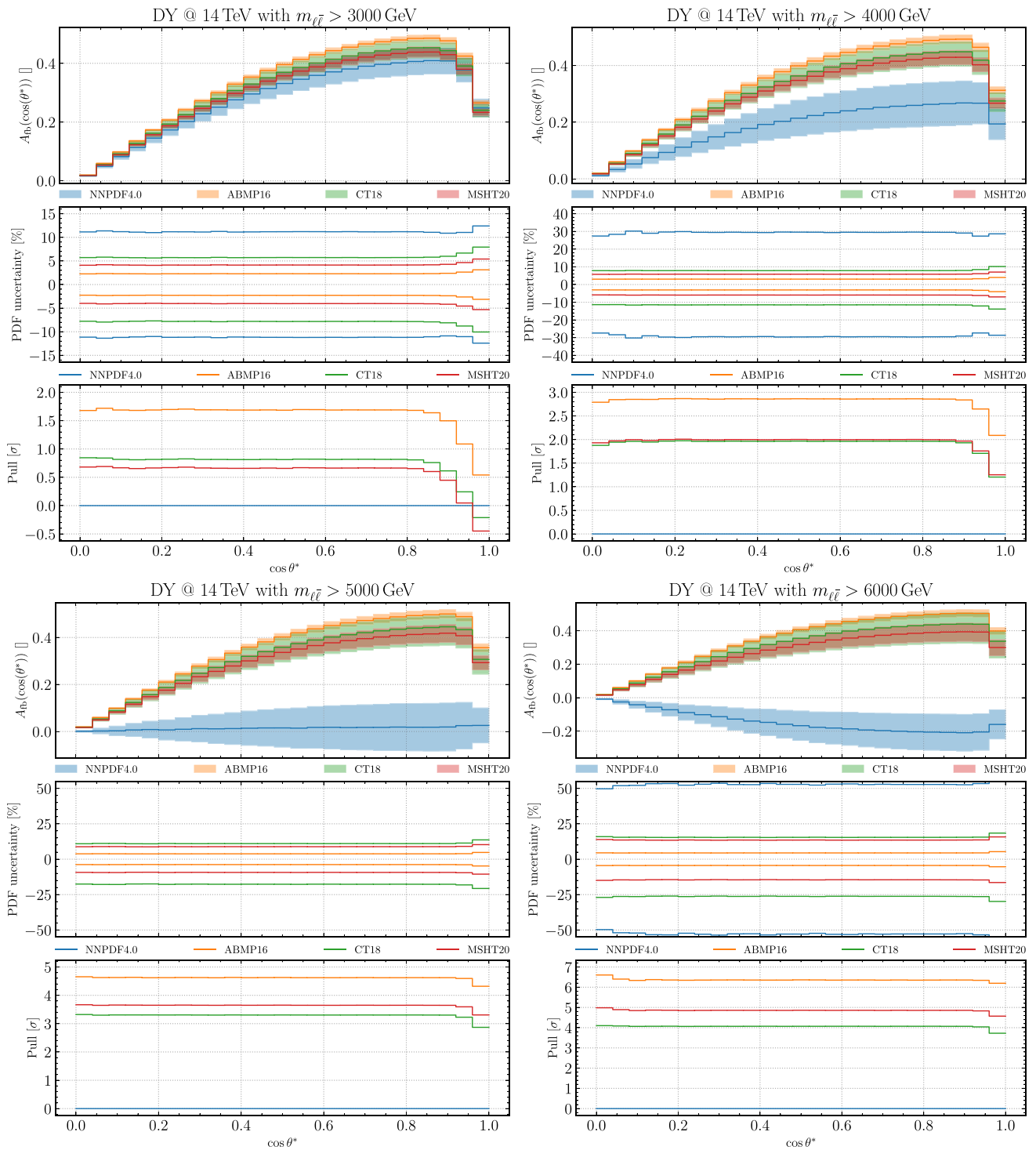
Fig. 20 Same as Fig. 18 (right) for different values of the lower cut in the dilepton invariant mass: $m_{\ell\bar{\ell}}^{\min} = 3, 4, 5,$ and 6 TeV
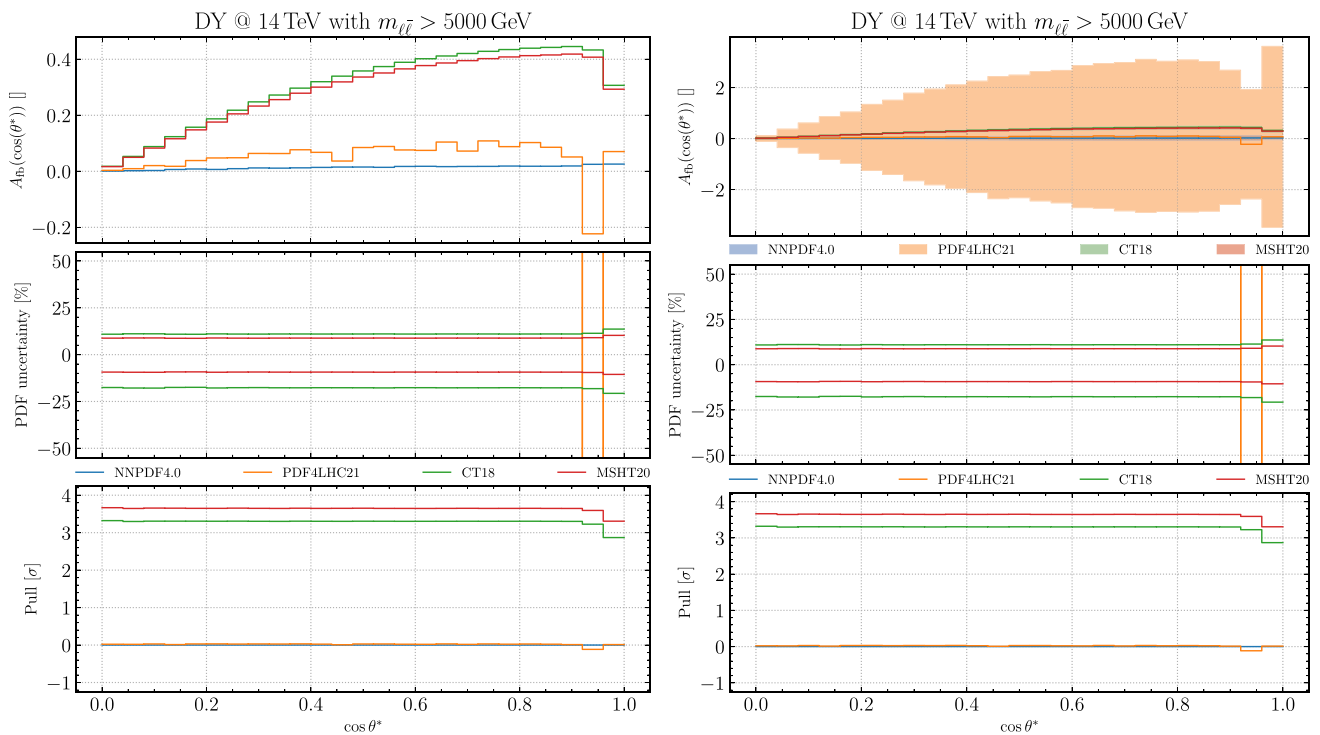
Fig. 21 Same as Fig. 20 (bottom left), now also including the PDF4LHC21 prediction. In the left plot the PDF uncertainty is not shown, in the right plot the scale on the y axis is suitably expanded. Note that the PDF4LHC percentage PDF uncertainty does not show as it falls outside the plot

a power of $(1 - x)$ as $x \to 1$, for all groups but NNPDF4.0 this power is constant for light quarks to the right of the valence peak, while for NNPDF4.0 it changes as x increases, slowly for up quarks, more rapidly for down quarks and even more rapidly for antiquarks. All this suggests that the different behavior of NNPDF4.0 is due to its more flexible PDF parametrization.

Our general conclusion is that the behavior of the forward–backward asymmetry observed at lower invariant masses is not necessarily reproduced at large masses if flexible enough PDFs are used: the characteristic positive asymmetry observed for low $m_{\ell\bar{\ell}}$ values can be washed out in the high-mass region. Hence, deviations from the traditional expectation of a positive forward–backward asymmetry in high-mass Drell–Yan cannot be taken as an indication of BSM physics, at least based on our current understanding of proton structure in the large-x region.

Turning the argument around, future measurements of the $\cos\theta^*$ distribution and the associated forward–backward asymmetry $A_{\rm fb}$ when included in PDF determinations could help in constraining PDFs at large x. For instance, Fig. 19 indicates that for $m_{\ell\bar{\ell}}^{\min} = 5$ TeV and $\sqrt{s} = 14$ TeV the asymmetry $A_{\rm fb}$ can be as large as 50% for ABMP16 while it vanishes (within large uncertainties) in the case of NNPDF4.0. By rebinning the $\cos\theta^*$ distribution, for an integrated luminosity of $\mathcal{L} = 6$ ab^{-1}, corresponding to the combination

at ATLAS and CMS at the end of the HL-LHC data-taking period, $\mathcal{O}(10)$ events are expected in the backward region, with an statistical uncertainty of $\delta_{\rm stat} \sim 30\%$ which could be sufficient to discriminate between these two limiting scenarios at the 2σ level.

Higher event counts are expected if the $m_{\ell\bar{\ell}}$ cut is loosened, though one is then less sensitive to the large-x region where differences between PDF sets and their uncertainties are the largest. Ultimately, the constraining power of high-mass Drell–Yan in general and of the forward–backward asymmetry in particular can only be addressed by means of a dedicated projections based on binned pseudo-data such as those carried out for the HL-LHC and the Electron Ion Collider in e.g. [48,49]. While we leave this exercise for a future study, the investigations presented in this work indicate that $A_{\rm fb}$ at high-invariant masses represents a promising and mostly unexplored channel to pin down large-x light quark and antiquark PDFs at the HL-LHC.

While in this work we have focused on the forward–backward asymmetry in neutral-current Drell–Yan production, similar considerations apply for other processes relevant for BSM searches at high mass at the LHC. Indeed, the HL-LHC will be sensitive to a broad range of hypothetical new massive particles, from resonances in the m_{jj} dijet invariant mass distribution up to 11 TeV, heavy vector triplet resonances decaying into a diboson VV' pair up to 5 TeV, and

gluinos with masses up to $m_{\tilde{g}} = 3$ TeV in the minimal supersymmetric standard model (MSSM) with a massless lightest SUSY particle [1].

For all these channels, a robust understanding of PDFs and their uncertainties at large x, including the role of methodological and model assumptions, will be necessary to fully exploit the HL-LHC discovery potential for BSM signatures. Conversely, once BSM phenomena have been excluded in some high-energy channel, the corresponding search can be unfolded into a measurement to provide direct constraints on the PDFs in this key large-x region, which in turn will enhance the reach of other searches. It would be very interesting to perform a detailed study, in the same vein as Ref. [48], of the impact of future HL-LHC data on large-x PDFs and the prospect of asymmetry measurements in searches for new physics, but this will be left for future work.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The PDF sets used in this work are available via the LHAPDF interface, and the calculation tools used for the LHC simulations are also publicly available.]

A A_{fb} in NNPDF3.1

In this appendix we compare partonic luminosities and LHC differential distributions obtained with NNPDF4.0 in Sects. 3 and 4 with those based on its predecessor NNPDF3.1, as well as with a variant of NNPDF4.0 where positivity is imposed at the level of observable cross-sections but not at the PDF level, as was the case in NNPDF3.1, which we will denote NNPDF4.0(3.1pos).

Figure 22 compares the symmetric partonic luminosities $\mathcal{L}_{S,q}$ evaluated for $m_{\ell\bar{\ell}} = 5$ TeV. The three sets are found to agree within uncertainties, with NNPDF4.0 having the smallest uncertainties. This increase in precision arises only marginally due to the more restrictive positivity constraints, since predictions with the NNPDF4.0(3.1pos) variant are close to the baseline NNPDF4.0, especially for the $u\bar{u}$ contribution, for both central values and uncertainties. The comparison in Fig. 22 indicates that phenomenological predictions for high-mass Drell–Yan production based on NNPDF3.1 are expected to be consistent within errors with those of NNPDF4.0 for the contributions symmetric in $\cos\theta^*$, such as the $|y_{\ell\bar{\ell}}|$ distribution.

The antisymmetric luminosities $\mathcal{L}_{A,q}$, relevant for the forward–backward asymmetry, are displayed in Fig. 23 for $m_{\ell\bar{\ell}} = 3$ and 5 TeV respectively. Their qualitative behavior is similar for all PDF sets, with a marked decrease of PDF uncertainties first from NNPDF3.1 to NNPDF4.0(3.1pos) then to NNPDF4.0. Specifically, the qualitative $m_{\ell\bar{\ell}}$ dependence of $\mathcal{L}_{A,q}$ remains unchanged. Namely, the positive A_{fb} found for $m_{\ell\bar{\ell}} = 3$ TeV decreases as the dilepton invariant mass is increased. Hence also for the component of the Drell–Yan cross-section which is odd in $\cos\theta^*$ we expect LHC predictions based on NNPDF3.1 to be consistent with those obtained from NNPDF4.0.

These expectations are confirmed by Fig. 24, which shows the dilepton rapidity $|y_{\ell\bar{\ell}}|$ and the Collins–Soper angle $\cos\theta^*$ distributions for neutral-current DY production at the LHC
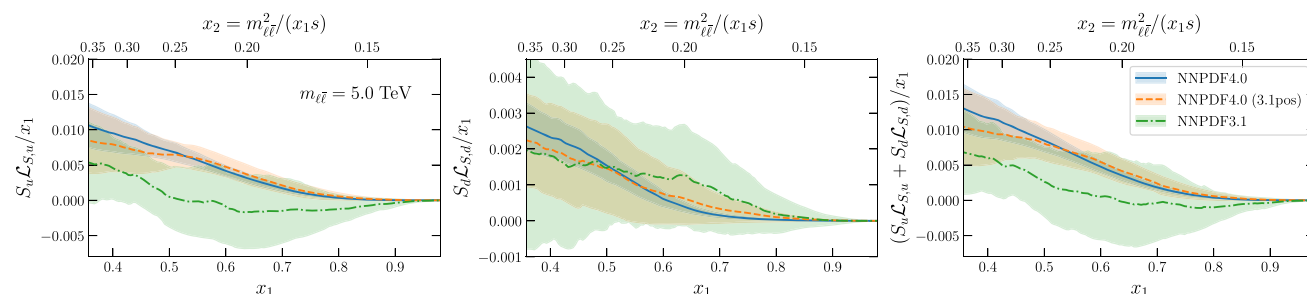

Fig. 22 Same as Fig. 12 (upper panels) comparing NNPDF4.0, NNPDF4.0(3.1pos), and NNPDF3.1
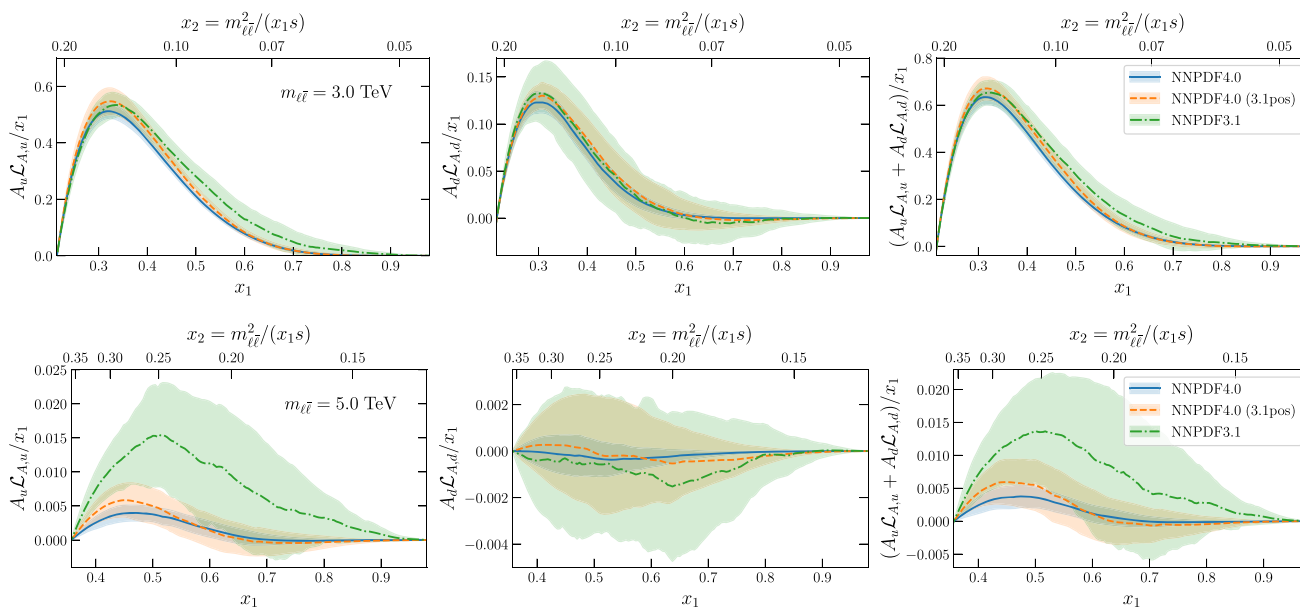
Fig. 23 Same as Fig. 13 for the antisymmetric partonic luminosities $\mathcal{L}_{A,q}$, comparing NNPDF4.0, NNPDF4.0(3.1pos), and NNPDF3.1.
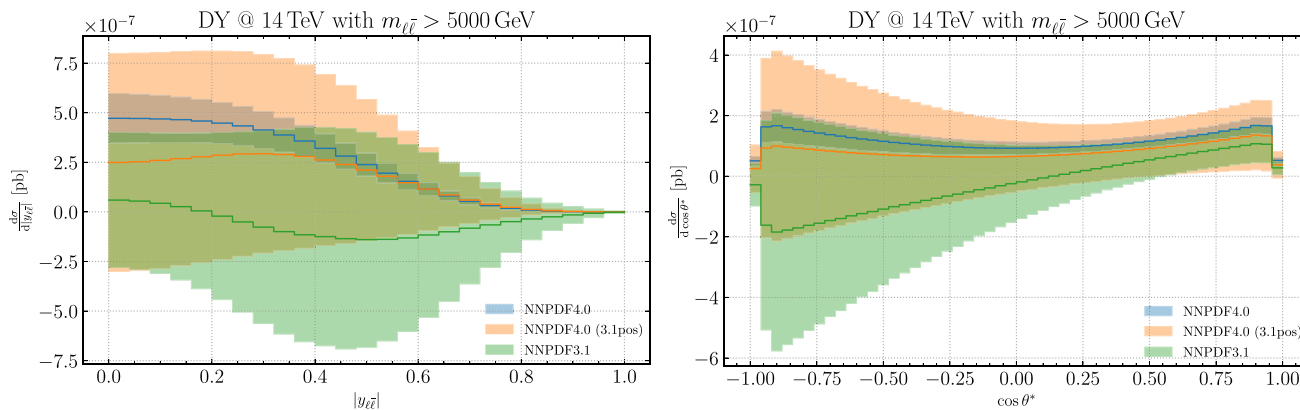

Fig. 24 Same as Figs. 17 and 19 for the absolute dilepton rapidity $|y_{\ell\bar{\ell}}|$ (left) and the $\cos\theta^*$ (right) distributions for dilepton invariant masses of $m_{\ell\bar{\ell}} \geq 5$ TeV comparing NNPDF4.0, NNPDF4.0(3.1pos), and NNPDF3.1.

14 TeV for dilepton invariant masses of $m_{\ell\bar{\ell}} \geq 5$ TeV, comparing the baseline NNPDF4.0 predictions with those from NNPDF3.1 and NNPDF4.0(3.1pos). Indeed, good agreement within the three PDF sets is observed with a significant reduction of PDF uncertainties between NNPDF3.1 and NNPDF4.0, consistent with the behaviour exhibited by the corresponding partonic luminosities.

References

1. X. Cid Vidal et al., Report from Working Group 3: Beyond the Standard Model physics at the HL-LHC and HE-LHC. CERN Yellow Rep. Monogr. **7**, 585–865 (2019). arXiv:1812.07831
2. J. Gao, L. Harland-Lang, J. Rojo, The structure of the proton in the LHC precision era. Phys. Rept. **742**, 1–121 (2018). arXiv:1709.04922
3. K. Kovařík, P. M. Nadolsky, D. E. Soper, Hadronic structure in high-energy collisions. Rev. Mod. Phys. **92**(4) (045003) (2020). arXiv:1905.06957
4. CDF Collaboration, F. Abe et al., Inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev. Lett. **77**, 438–443 (1996). arXiv:hep-ex/9601008
5. H.L. Lai, J. Huston, S. Kuhlmann, F.I. Olness, J.F. Owens, D.E. Soper, W.K. Tung, H. Weerts, Improved parton distributions from global analysis of recent deep inelastic scattering and inclusive jet data. Phys. Rev. D **55**, 1280–1296 (1997). arXiv:hep-ph/9606399
6. W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, E. Laenen, S. Marzani, J. Rojo, NLO+NLL squark and gluino production cross-sections with threshold-improved parton distributions. Eur. Phys. J. C **76**(2), 53 (2016). arXiv:1510.00375
7. SMEFiT Collaboration, J. J. Ethier, G. Magni, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang, Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC. JHEP **11**, 089 (2021). arXiv:2105.00006

8. S. Dawson, P. P. Giardino, A. Ismail, Standard model EFT and the Drell-Yan process at high energy. Phys. Rev. D **99**(3), 035044 (2019). arXiv:1811.12260

9. J. Ellis, M. Madigan, K. Mimasu, V. Sanz, T. You, Top, Higgs, Diboson and electroweak fit to the standard model effective field theory. JHEP **04**, 279 (2021). arXiv:2012.02779

10. A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali, C. Voisey, Parton distributions in the SMEFT from high-energy Drell-Yan tails. JHEP **07**, 122 (2021). arXiv:2104.02723

11. ATLAS Collaboration, G. Aad et al., Search for contact interactions and large extra dimensions in the dilepton channel using proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector. Eur. Phys. J. C **74**(12), 3134 (2014). arXiv:1407.2410

12. ATLAS Collaboration, G. Aad et al., Search for new non-resonant phenomena in high-mass dilepton final states with the ATLAS detector. JHEP **11**, 005 (2020). arXiv:2006.12946. [Erratum: JHEP 04, 142 (2021)]

13. ATLAS Collaboration, G. Aad et al., Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector. Phys. Lett. B **796**, 68–87 2019). arXiv:1903.06248

14. C.M.S. Collaboration, A.M. Sirunyan et al., Search for resonant and nonresonant new phenomena in high-mass dilepton final states at $\sqrt{s}$ = 13 TeV. JHEP **07**, 208 (2021). arXiv:2103.02708

15. ATLAS Collaboration, G. Aad et al., Search for new phenomena in final states with two leptons and one or no b-tagged jets at $\sqrt{s}$ = 13 TeV using the ATLAS detector. Phys. Rev. Lett. **127**(14), 141801 (2021). arXiv:2105.13847

16. C.M.S. Collaboration, A.M. Sirunyan et al., Search for contact interactions and large extra dimensions in the dilepton mass spectra from proton-proton collisions at $\sqrt{s}$ = 13 TeV. JHEP **04**, 114 (2019). arXiv:1812.10443

17. C. Duhr, B. Mistlberger, Lepton-pair production at hadron colliders at N^{3}LO in QCD. JHEP **03**, 116 (2022). arXiv:2111.10379

18. F. Buccioni, F. Caola, M. Delto, M. Jaquier, K. Melnikov, R. Röntsch, Mixed QCD-electroweak corrections to on-shell Z production at the LHC. Phys. Lett. B **811**, 135969 (2020). arXiv:2005.10221

19. F. Buccioni, F. Caola, H.A. Chawdhry, F. Devoto, M. Heller, A. von Manteuffel, K. Melnikov, R. Röntsch, C. Signorile-Signorile, Mixed QCD-electroweak corrections to dilepton production at the LHC in the high invariant mass region. JHEP **06**, 022 (2022). arXiv:2203.11237

20. R. Bonciani, F. Buccioni, N. Rana, A. Vicini, Next-to-Next-to-Leading Order Mixed QCD-Electroweak Corrections to on-Shell Z Production, Phys. Rev. Lett. **125**(23), 232004 (2020). arXiv:2007.06518

21. R. Bonciani, L. Buonocore, M. Grazzini, S. Kallweit, N. Rana, F. Tramontano, A. Vicini, Mixed Strong-electroweak corrections to the Drell-Yan Process. Phys. Rev. Lett. **128**(1), 012002 (2022). arXiv:2106.11953

22. T. Armadillo, R. Bonciani, S. Devoto, N. Rana, A. Vicini, Two-loop mixed QCD-EW corrections to neutral current Drell-Yan. JHEP **05**, 072 (2022). arXiv:2201.01754

23. ATLAS Collaboration, M. Aaboud et al., Measurement of the Drell-Yan triple-differential cross section in pp collisions at $\sqrt{s}$ = 8 TeV. JHEP **12**, 059 (2017). arXiv:1710.05167

24. CMS Collaboration, A. Tumasyan et al., Measurement of the Drell-Yan forward-backward asymmetry at high dilepton masses in proton-proton collisions at $\sqrt{s}$ = 13 TeV. arXiv:2202.12327

25. J. Fiaschi, F. Giuli, F. Hautmann, S. Moretti, Enhancing the Large Hadron Collider sensitivity to charged and neutral broad resonances of new gauge sectors. JHEP **02**, 179 (2022). arXiv:2111.09698 newpage

26. J. Fiaschi, F. Giuli, F. Hautmann, S. Moretti, Lepton-Charge and Forward-Backward Asymmetries in Drell-Yan Processes for Precision Electroweak Measurements and New Physics Searches. Nucl. Phys. B **968**, 115444 (2021). arXiv:2103.10224

27. E. Accomando et al., PDF Profiling Using the Forward-Backward Asymmetry in Neutral Current Drell-Yan Production. JHEP **10**, 176 (2019). arXiv:1907.07727

28. E. Accomando, J. Fiaschi, F. Hautmann, and S. Moretti, Neutral current forward–backward asymmetry: from θ_W to PDF determinations. Eur. Phys. J. C **78**(8), 663 (2018). arXiv:1805.09239. [Erratum: Eur.Phys.J.C 79, 453 (2019)]

29. J. Fiaschi, F. Giuli, F. Hautmann, S. Moch, and S. Moretti, **Z**′-boson dilepton searches and the high-**x** quark density. arXiv:2211.06188

30. CMS Collaboration, A. M. Sirunyan et al., Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV. Eur. Phys. J. C **78**(9), 701 (2018). arXiv:1806.00863

31. NNPDF Collaboration, R. D. Ball et al., The path to proton structure at 1% accuracy. Eur. Phys. J. C **82**(5), 428 (2022). arXiv:2109.02653

32. S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Parton distribution functions, α_s, and heavy-quark masses for LHC Run II. Phys. Rev. D **96**(1), 014011 (2017). arXiv:1701.05838

33. T.-J. Hou et al., New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, Phys. Rev. D **103**(1), 014013 (2021). arXiv:1912.10053

34. S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, R. S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs, Eur. Phys. J. C **81**(4), 341 (2021). arXiv:2012.04684

35. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. JHEP **1407**, 079 (2014). arXiv:1405.0301

36. S. Carrazza, E.R. Nocera, C. Schwan, M. Zaro, PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes. JHEP **12**, 108 (2020). arXiv:2008.12789

37. C. Schwan, A. Candido, F. Hekhorn, S. Carrazza, N3PDF/pineappl: v0.5.5, Aug (2022)

38. J.C. Collins, D.E. Soper, Angular distribution of dileptons in high-energy hadron collisions. Phys. Rev. D **16**, 2219 (1977)

39. M.E. Peskin, D.V. Schroeder, *An Introduction to quantum field theory* (Addison-Wesley, Reading, USA, 1995)

40. A. Candido, S. Forte, F. Hekhorn, Can $\overline{\text{MS}}$ parton distributions be negative? JHEP **11**, 129 (2020). [arXiv:2006.07377]

41. S.J. Brodsky, G.R. Farrar, Scaling laws at large transverse momentum. Phys. Rev. Lett. **31**, 1153–1156 (1973)

42. S.J. Brodsky, G.R. Farrar, Scaling Laws for Large Momentum Transfer Processes. Phys. Rev. D **11**, 1309 (1975)

43. NNPDF Collaboration, R. D. Ball et al., Parton distributions from high-precision collider data. Eur. Phys. J. C **77**(10), 663 (2017). arXiv:1706.00428

44. R. D. Ball, E. R. Nocera, J. Rojo, The asymptotic behaviour of parton distributions at small and large x, Eur. Phys. J. C **76**(7), 383 (2016). arXiv:1604.00024

45. CMS Collaboration, V. Khachatryan et al., Search for narrow resonances in dilepton mass spectra in proton-proton collisions at $\sqrt{s}$ = 13 TeV and combination with 8 TeV data. Phys. Lett. B **768**, 57–80 (2017). arXiv:1609.05391

46. PDF4LHC Working Group Collaboration, R. D. Ball et al., The PDF4LHC21 combination of global PDF fits for the LHC Run III. J. Phys. G **49**(8), 080501 (2022). arXiv:2203.05506

47. S. Carrazza, J.I. Latorre, J. Rojo, G. Watt, A compression algorithm for the combination of PDF sets. Eur. Phys. J. C **75**, 474 (2015). arXiv:1504.06469

48. R. Abdul Khalek, S. Bailey, J. Gao, L. Harland-Lang, and J. Rojo, Towards Ultimate Parton Distributions at the High-Luminosity LHC. Eur. Phys. J. C **78**(11), 962 (2018). arXiv:1810.03639

49. R. A. Khalek, J. J. Ethier, E. R. Nocera, and J. Rojo, Self-consistent determination of proton and nuclear PDFs at the Electron Ion Collider. Phys. Rev. D **103**(9), 096005 (2021). arXiv:2102.00018