ZU-TH 25/24 CERN-TH-2024-062

# QCD predictions for vector boson plus hadron production at the LHC

**S. Caletti,**[a] **A. Gehrmann–De Ridder,**[a,b] **A. Huss,**[c] **A. Rodriguez Garcia,**[a] **G. Stagnitto**[d]

[a] *Institute for Theoretical Physics, ETH, CH-8093 Zürich, Switzerland*
[b] *Department of Physics, University of Zürich, CH-8057 Zürich, Switzerland*
[c] *Theoretical Physics Department, CERN, CH-1211 Geneva 23, Switzerland*
[d] *Università degli Studi di Milano-Bicocca & INFN, Piazza della Scienza 3, I-20126 Milano, Italy*

*E-mail:* scaletti@phys.ethz.ch, gehra@phys.ethz.ch, alexander.huss@cern.ch, adrianro@phys.ethz.ch, giovanni.stagnitto@unimib.it

Abstract: The identification of a hadron in the final state of hadron-collider events that feature a leptonically decaying vector boson can provide essential information on the parton content of the colliding protons. Moreover, the study of hadrons inside jets can provide deeper insights into the fragmentation dynamics. We provide theoretical predictions for specific observables involving either the production of a $Z$ boson in association with light charged hadrons inside a jet or the production of a $W$ boson together with a charmed hadron. We present results for various fragmentation functions and compare our predictions with measurements by LHCb and ATLAS at $\sqrt{s} = 13$ TeV. Our predictions are obtained using the antenna subtraction formalism which has been extended to cope with infrared singularities associated to the fragmentation processes in a hadron-collider environment at NLO accuracy.

arXiv:2405.17540v2 [hep-ph] 4 Oct 2024

## Contents


**1 Introduction** **1**

**2 Identified hadron production at hadron colliders** **3**
- 2.1 Subtraction at next-to-leading order 5
  - 2.1.1 Subtraction at real level 6
  - 2.1.2 Subtraction at virtual level 7

**3 Infrared structure of vector boson and hadron production** **9**
- 3.1 Leading order 10
- 3.2 Next-to-leading order 10
  - 3.2.1 Real matrix element contributions 10
  - 3.2.2 Real subtraction terms 12
  - 3.2.3 Virtual matrix element contributions 13
  - 3.2.4 Virtual subtraction terms 14

**4 $Z$ boson in association with a hadron inside a jet** **17**
- 4.1 Observable definition and setup 17
- 4.2 Results for unidentified charged hadron production 18
- 4.3 Results for charged pion production 19

**5 $W$ boson in association with a charmed hadron** **21**
- 5.1 Fiducial cuts and numerical setup 23
- 5.2 Results for the $|\eta_\ell|$ observable 24
- 5.3 Results for the $p_{T,h}$ observable 27

**6 Conclusions** **29**


## 1 Introduction

With the rapidly growing data set from the LHC and the reduction of both statistical and systematic uncertainties, more and more observables are becoming accessible for precision studies. One such frontier comprises processes with identified hadrons from light or heavy quark flavours.

Identified hadron production is described in perturbative QCD through the production of partons (quarks and gluons) which subsequently fragment into hadrons. This parton-to-hadron transition is a non-perturbative process which can be parametrized in terms of process-independent fragmentation functions (FFs). The latter describe the probability of a parton fragmenting into a hadron carrying some fraction of its momentum. These FFs

fulfil Altarelli-Parisi evolution equations in their resolution scale [1], which are in complete analogy to the evolution of parton distributions functions (PDFs) in the nucleon.

Identified hadron production can arise in two categories of collider observables: *inclusive* observables, which are differential in the hadron momentum but inclusive in the kinematics of all other final state particles; *exclusive* observables, where the hadrons are identified in events that are characterized by requiring the additional presence of jets and/or gauge bosons in the final state.

One-particle inclusive cross sections have been measured for a variety of hadron species in electron-positron, lepton-hadron and hadron-hadron collisions. On the theory side, using the factorization properties of QCD, the identified hadron production cross section can be written as the convolution of a process-dependent coefficient function for the production of partons with a universal fragmentation function for the parton-to-hadron transition. The coefficient functions, corresponding to the parton-level cross section can be calculated in perturbative QCD to the desired orders. At present, these coefficient functions are known to NLO for hadron-hadron collisions [2] and to NNLO for $e^+e^-$ annihilation [3, 4] and lepton-proton scattering [5].

The recent focus on exclusive observables is largely driven by precision studies of vector boson production in association with identified hadrons at the LHC, with the presence of leptons from the vector boson decay providing a clean experimental signature. These exclusive observables can then provide essential information on the quark flavour decomposition of the colliding protons: for instance, $W$-boson plus charmed hadron processes are important handles on the strange content of the proton [6–10].

Other exclusive observables require the presence of a hadron inside a jet [**?** ]. These processes can provide significant constraints on fragmentation functions. Recently, observables involving the production of a $Z$-boson in association with light hadrons inside a jet have been performed by the LHCb collaboration [11].

Theory predictions for any of these exclusive observables require the use of a parton-level event generator implementing all parton-level contributions to a given order and applying the experimental kinematical requirements on the final state event selection. For the computation of observables related to processes with hadrons and involving parton-to-hadron fragmentation functions, any subtraction method originally developed for jet production needs to be extended: the subtraction procedure must keep track of the identified parton momentum fraction in the unresolved emissions, which is usually integrated over for purely jet observables. Such extensions are available at NLO for dipole subtraction [12] and FKS subtraction [13]. At NNLO, recent work towards the description of fragmentation processes yielded results for heavy hadron production in top quark decays [14] in the residue subtraction scheme as well as for photon fragmentation [15, 16] in the antenna subtraction scheme. An extension of the antenna subtraction formalism to incorporate arbitrary hadron fragmentation processes at lepton colliders has been outlined in [17]. Very recently, an interface to MG5_aMC@NLO [18] to compute exclusive processes with fragmentation at NLO through the usage of a hybrid subtraction scheme has been presented [19]. Other theoretical developments concern the interface of fixed-order NLO calculations to parton showers for processes with identified hadrons in the final state: for instance, NLO+PS predictions for

$W$ plus charmed hadron production have been produced in [20, 21].

In this paper, we consider theory predictions for $Z$-boson plus light hadron and $W$-boson plus heavy hadron production. We present the essential ingredients of the antenna formalism extended to cope with infrared singularities associated to fragmentation processes at NLO accuracy. This work is a first step towards obtaining NNLO predictions for vector boson plus identified hadron production at the LHC.

The paper is structured as follows. In section 2, we present the general framework describing hadron fragmentation processes for hadron collider observables up to NLO level. In section 3 we give explicit expressions for the subtraction terms and their integrated counterparts needed for the computation of hadron production in association with a $Z$ or $W$ boson. As a first application of our formalism, in sections 4 and 5 we present theoretical predictions using specific parametrizations for the fragmentation functions and perform in-depth comparisons with LHCb and ATLAS data at 13 TeV. Finally we summarize our findings in Section 6.

## 2 Identified hadron production at hadron colliders

We begin by recollecting the main features of the different ingredients that enter the computation of hadron-collider observables associated to the production of identified hadrons at NLO.

We consider a generic proton–proton to hadron process of the form:

$$\mathrm{p} + \mathrm{p} \to h(K_h) + X\,(\,+\mathrm{jets}\,)\,, \tag{2.1}$$

where the hadron $h$ with momentum $K_h$ may or may not be inside a jet. The fully differential cross section for a process of this kind can be written in a factorized form

$$\mathrm{d}\sigma^h = \sum_{i,j}\sum_{p}\int \frac{\mathrm{d}\xi_1}{\xi_1}\frac{\mathrm{d}\xi_2}{\xi_2}\mathrm{d}\eta\, f_i(\xi_1,\mu_F^2) f_j(\xi_2,\mu_F^2) D_p^h(\eta,\mu_D^2)\,\mathrm{d}\hat{\sigma}_p^{i,j}(\eta,\mu_F^2,\mu_D^2)\,, \tag{2.2}$$

where $f_{i,j}$ are the parton distribution functions (PDFs) with momentum fractions $\xi_{1,2}$ and $D_p^h$ denotes the fragmentation function (FF) describing the transition of a parton $p$ into a hadron $h$ carrying the momentum fraction $\eta = K_h/k_p$. After mass factorization, both the PDFs and FF acquire a dependence on the factorisation ($\mu_F$) and fragmentation scale ($\mu_D$), respectively, which compensates against the corresponding scale dependence of the hard scattering cross section $\mathrm{d}\hat{\sigma}_p^{i,j}(\eta,\mu_F^2,\mu_D^2)$. A sum over the initial-state ($i$, $j$) and final-state ($p$) partons is performed. Finally, the one-parton exclusive cross section $\mathrm{d}\hat{\sigma}_p^{i,j}$ is calculated perturbatively as a series expansion in $\alpha_s$ and reads

$$\mathrm{d}\hat{\sigma}_p(\eta,\xi_{1,2}) = \mathrm{d}\hat{\sigma}_p^{\mathrm{LO}}(\eta,\xi_{1,2}) + \left(\frac{\alpha_s}{2\pi}\right)\mathrm{d}\hat{\sigma}_p^{\mathrm{NLO}}(\eta,\xi_{1,2}) + \left(\frac{\alpha_s}{2\pi}\right)^2\mathrm{d}\hat{\sigma}_p^{\mathrm{NNLO}}(\eta,\xi_{1,2}) + \ldots\,, \tag{2.3}$$

where we specify the fragmenting parton denoted as $p$ but leave the initial state $(i,j)$ implicit. In the following, we describe how the parton-level cross section given in eq. (2.3) is computed up to NLO. We closely follow the notation adopted in [17], where the formalism needed to incorporate arbitrary hadron fragmentation processes for lepton-collider observables has been outlined up to NNLO.

For a specific initial state denoted as $(i, j)$ and a fragmenting parton denoted as $p$ in eq. (2.2), the leading order (LO) cross section is defined as the integration over the $n$-particle phase space of the tree-level Born partonic cross section as

$$\mathrm{d}\hat{\sigma}_p^{\mathrm{LO}}(\eta, \xi_{1,2}) = \int_n \mathrm{d}\hat{\sigma}_p^{\mathrm{B}}(\eta, \xi_{1,2})\,, \tag{2.4}$$

with,

$$\mathrm{d}\hat{\sigma}_p^{\mathrm{B}}(\eta, \xi_{1,2}) = \mathcal{N}_{\mathrm{B}}\,\mathrm{d}\Phi_n(k_1, \ldots, k_n; \xi_1 P_1, \xi_2 P_2)\,\frac{1}{S_n}\,M_n^0(k_1, \ldots, k_n)\,J_n^{(m)}(\{k_1, \ldots, k_n\}_n, \eta k_p)\,. \tag{2.5}$$

Here, $\mathcal{N}_{\mathrm{B}}$ denotes the normalisation factor at Born-level, $S_n$ a symmetry factor related to the final-state particles, $M_n^0$ the tree-level matrix element for $2 \to n$ scattering, and $\mathrm{d}\Phi_n$ the phase space for a $n$-parton final state with total four-momentum $Q = \xi_1 P_1 + \xi_2 P_2$. Throughout the remainder of the paper, we will define the initial-state partons $(i, j)$ as $(1, 2)$, use $p_{1,2} = \xi_{1,2} P_{1,2}$, and leave the PDF momentum fractions and the flavour content of the initial-state partons implicit.

The jet function $J_n^{(m)}$ in eq. (2.5) implicitly describes how the jet cross sections are computed. In particular, it includes the jet clustering algorithm and application of cuts on jet observables in case their reconstruction is part of the process definition. Compared to standard jet functions, the jet function $J_n^{(m)}$ here requires modification for the computation presented in this work: It depends explicitly on the momentum fraction $\eta$ relating the momentum $k_p$ of the identified parton and the momentum of the identified hadron $K_h = \eta k_p$. This enables to define and apply cuts on observables that depend explicitly on the momentum fraction $\eta$ carried by the hadron.

At higher orders, infrared divergences due to the emission of soft and/or collinear radiation appear in the parton-level cross section. For sufficiently inclusive observables these divergences are guaranteed to cancel between real and virtual contributions. A subtraction method is needed however, in order to regulate these divergences at intermediate stages of the calculation. In the antenna subtraction formalism, used in this work, subtraction terms are constructed with products of antenna functions and reduced matrix elements of lower multiplicity. The antenna functions capture all unresolved radiation between a pair of hard radiators, thus reproducing the behaviour of the matrix element in the singular limits. The integrated subtraction terms are then obtained by analytically integrating the antenna functions over their respective factorised phase space. Those integrated subtraction terms are then added back at the virtual level in order to cancel the explicit poles in the virtual matrix elements and additional mass-factorisation counterterms present in a hadron-collider setup. Most of the ingredients needed to deal with the computation of hadron-collider jet observables at NNLO using the antenna formalism [22] can also be applied here in the context of identified hadron production. However, the explicit identification of a final-state parton in $\mathrm{d}\hat{\sigma}_p^{i,j}$ spoils the cancellation mechanism of collinear divergences between real and virtual contributions. For instance, a hard object composed of a single quark or a collinear quark–gluon cluster are distinguished as separate objects in this case. The associated collinear divergence remains uncancelled and is instead absorbed into the

bare fragmentation function through mass factorization counterterms. This renormalization procedure introduces a dependence on an arbitrary fragmentation scale $\mu_D$ at which this subtraction is performed. In the context of the application of the antenna subtraction formalism, this procedure has been presented in [17].

To account for the presence of an identified particle, and to guarantee that the hadron-level cross section can be computed as a convolution of the parton-level cross section and the parton-to-hadron fragmentation function as given in eq. (2.2), one needs to keep track of the momentum fraction of the fragmenting parton at all stages of the computation of the short-distance cross section. As a consequence, the construction of subtraction terms at unintegrated and integrated level within the antenna subtraction formalism must be modified accordingly.

In general, these subtraction terms will involve so-called fragmentation antennae, where an identified parton is tracked and its momentum dependence is kept explicitly. The integrated subtraction terms will involve the integration of these fragmentation antennae over the relevant phase spaces while retaining their dependence on the momentum fraction of the fragmenting parton. After integrating over all kinematical variables except the momentum fraction of the identified parton, these integrated fragmentation antennae will have the proper structure to be combined with virtual contributions and mass-factorisation counterterms associated with final-state collinear divergences of the identified parton. The analytic pole cancellation occurs before the convolution with the fragmentation function has taken place and the finite remainder can be evaluated numerically using a parton-level event generator, like NNLOJET [**?** ].

In the remainder of this section we present the general structure of the unintegrated and integrated subtraction terms needed to account for the presence of identified-hadron production in a hadron-collider environment at NLO.

### 2.1 Subtraction at next-to-leading order

At NLO level, the one-parton exclusive cross section as given in eq. (2.3) comprises two separately divergent contributions: the real-emission corrections containing implicit soft and collinear divergences and the virtual loop-corrections with explicit divergences regulated as $1/\varepsilon$ poles in dimensional regularisation. These two types of corrections feature different particle multiplicities, thus requiring a separate numerical integration. The antenna subtraction scheme at NLO level is used to construct a real subtraction term $\mathrm{d}\hat{\sigma}_p^{\mathrm{S}}$ and a virtual subtraction term $\mathrm{d}\hat{\sigma}_p^{\mathrm{T}}$ in order to deal with these divergences in the intermediate stages of the computation. As a consequence, for a specified identified parton $p$, the $\eta$ dependent parton-level cross section $\mathrm{d}\hat{\sigma}_p^{\mathrm{NLO}}(\eta)$ in eq. (2.3) may be written as:

$$\mathrm{d}\hat{\sigma}_p^{\mathrm{NLO}}(\eta) = \int_{n+1} \left[\mathrm{d}\hat{\sigma}_p^{\mathrm{R}}(\eta) - \mathrm{d}\hat{\sigma}_p^{\mathrm{S}}(\eta)\right] + \int_{n} \left[\mathrm{d}\hat{\sigma}_p^{\mathrm{V}}(\eta) - \mathrm{d}\hat{\sigma}_p^{\mathrm{T}}(\eta)\right] , \tag{2.6}$$

with the two integrals over each particle multiplicity $n$ and $n+1$ being separately finite and thus suitable for a numerical implementation.

### 2.1.1 Subtraction at real level

The real partonic cross section $\mathrm{d}\hat{\sigma}_p^{\mathrm{R}}$ has the same structure as the Born-level cross section in eq. (2.5) with an additional parton and a corresponding real-level normalisation factor. It is composed of several terms that cover all the possibilities of having one of its $n+1$ partons becoming unresolved with a further decomposition according to their colour structure. Similarly, we decompose the real subtraction term $\mathrm{d}\hat{\sigma}_p^{\mathrm{S}}$ with an identified parton $p$ into a sum of contributions in which parton $j$ can become unresolved

$$\mathrm{d}\hat{\sigma}_p^{\mathrm{S}} = \sum_j \mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S}} \,. \tag{2.7}$$

The term $\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S}}$ covers all relevant colour orderings and can be further split into two parts:

$$\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S}} = \mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,non\text{-}id.}p} + \mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,id.}p} \,. \tag{2.8}$$

Here, the subtraction term denoted as $\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,non\text{-}id.}p}$ deals with contributions in which the unresolved parton $j$ is not colour connected to the identified parton $p$. This type of term is constructed using the standard antenna formalism with conventional NLO antenna functions $X_3^0$. In addition, the jet function ensures that the fragmenting particle is hard, i.e. cannot go unresolved. Since their construction is well established, we will not discuss $\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,non\text{-}id.}p}$ further here.

The subtraction term $\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,id.}p}$ is instead composed of contributions in which the unresolved parton $j$ is colour connected to the identified parton $p$. In this case, the dependence on the fragmentation momentum fraction carried by the identified parton must be retained explicitly. For the case of initial-state parton 1 being the colour-connected partner of $p$, a generic real subtraction term of this type takes the following form:

$$\begin{aligned}\mathrm{d}\hat{\sigma}_{p,j}^{\mathrm{S,id.}p} &= \mathcal{N}_{\mathrm{R}} \mathrm{d}\Phi_{n+1}(k_1,\ldots,k_p,\ldots,k_{n+1};p_1,p_2)\,\frac{1}{S_{n+1}}\,X_3^{0,\mathrm{id.}p}(p_1;k_j,k_p^{\mathrm{id.}}) \\ &\quad\times M_n^0(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};xp_1,p_2)\,J(\{\ldots,\tilde{k}_p,\ldots\}_n,\eta\,z\,\tilde{k}_p)\,,\end{aligned} \tag{2.9}$$

where $\mathcal{N}_{\mathrm{R}} = \mathcal{N}_{\mathrm{B}}\,\overline{C}(\epsilon)/C(\epsilon)$, with

$$C(\epsilon) = \frac{(4\pi e^{-\gamma_E})^\epsilon}{8\pi^2}\,, \quad \overline{C}(\epsilon) = (4\pi e^{-\gamma_E})^\epsilon\,. \tag{2.10}$$

In this equation $X_3^{0,\mathrm{id.}p}$ is a standard three-parton initial–final antenna function involving the identified particle with momentum denoted as $k_p$. The momentum of the identified particle after an $n+1 \to n$ mapping, as it enters in the reduced matrix element, is denoted as $\tilde{k}_p$. The information on the momentum fraction $z$ must be retained through the mapping such that the hadron momentum of the subtraction term is given by $k_h = \eta z \tilde{k}_p$. With the common definition of the momentum fraction $x$ related to initial state emissions (with $p_1$ denoting the momentum of an initial state parton),

$$x = \frac{Q^2}{2p_1\cdot q}\,, \quad q^2 = (p_q - k_p - p_1)^2 = -Q^2\,, \tag{2.11}$$

in the case of initial–final kinematics, the momentum fraction $z$ can be chosen as:

$$z = x\frac{(k_p - p_1)^2}{Q^2} = \frac{s_{1p}}{s_{1p} + s_{1j}}\,, \tag{2.12}$$

to satisfy the desired properties. In particular, the momentum entering the jet function at this level is $\eta z \tilde{k}_p$ as given in eq. (2.9).

### 2.1.2 Subtraction at virtual level

The virtual-level subtraction term for the one-particle exclusive cross section in eq. (2.6) is given by the combination of the integrated real-level subtraction terms and mass factorisation counterterms as:

$$\mathrm{d}\hat{\sigma}_p^{\mathrm{T}}(\eta) = -\int_1 \mathrm{d}\hat{\sigma}_p^{\mathrm{S}}(\eta) - \mathrm{d}\hat{\sigma}_p^{\mathrm{MF}}(\eta). \tag{2.13}$$

Focussing on the first term of this equation, the integration of standard antenna functions with initial–final kinematics is expanded using the phase space factorisation presented in [23]. Here we focus on the integration of the fragmentation antennae containing the identified particle as given in eq. (2.9). The first step is to factorise the initial–final phase space and include the explicit integration over the momentum fraction $z$. This results in

$$\begin{aligned}\mathrm{d}\Phi_{n+1}(k_1,\ldots,k_p,k_j,\ldots,k_{n+1};p_1,p_2) &= \mathrm{d}\Phi_n(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};xp_1,p_2)\frac{\mathrm{d}x}{x}\mathrm{d}z \\ &\times \frac{Q^2}{2\pi}\mathrm{d}\Phi_2(k_p^{\mathrm{id.}},k_j;p_1,q)\delta\Big(z - \frac{s_{1p}}{s_{1p}+s_{1j}}\Big).\end{aligned} \tag{2.14}$$

Similarly to the photon fragmentation case [15], we define the integrated fragmentation antenna function denoted as $\mathcal{X}_3^{0,\mathrm{id.}p}(x,z)$ over the two-parton phase space as present in eq. (2.14) as:

$$\begin{aligned}\mathcal{X}_3^{0,\mathrm{id.}p}(x,z) &= \frac{1}{C(\epsilon)}\int \mathrm{d}\Phi_2(k_p^{\mathrm{id.}},k_j;p_1,q)\frac{Q^2}{2\pi}\,X_3^0(p_1;k_j,k_p^{\mathrm{id.}})\delta\Big(z - \frac{s_{1p}}{s_{1p}+s_{1j}}\Big) \\ &= \frac{Q^2}{2}\frac{e^{\gamma_E\epsilon}}{\Gamma(1-\epsilon)}\left(Q^2\right)^{-\epsilon}\mathcal{J}(x,z)\,X_{1,jp}^{0,\mathrm{id.}p}(x,z)\,.\end{aligned} \tag{2.15}$$

Note that no integration is needed at this stage, i.e. just an expansion in distribution is required here. The Jacobian factor present in eq. (2.15) is given by

$$\mathcal{J}(x,z) = (1-x)^{-\epsilon}x^{\epsilon}z^{-\epsilon}(1-z)^{-\epsilon}\,. \tag{2.16}$$

The latter factor arises from writing the two-body phase space integral as a single integral over $z$. The product of the $x$- and $z$-dependent Jacobian and antenna functions in eq. (2.15) leads to terms of the form $(1-x)^{-1-\varepsilon}$ and $(1-z)^{-1-\varepsilon}$ which regulate the end-point soft divergences. These can be expanded in term of distributions according to

$$(1-y)^{-1-k\varepsilon} = -\frac{1}{k\varepsilon}\delta(1-y) + \sum_n \frac{(k\varepsilon)^n}{n!}\mathcal{D}_n(y)\,, \tag{2.17}$$

with plus-distributions of the form

$$\mathcal{D}_n(y) = \Big(\frac{\log^n(1-y)}{1-y}\Big)_+ . \tag{2.18}$$

Using the integrated form of the fragmentation antenna as given in eq. (2.15) the integrated subtraction term for an identified parton denoted as $p$ with momentum $k_p$ before the mapping, is given by

$$\begin{aligned}\int_1 \mathrm{d}\hat{\sigma}^{\mathrm{S,id.}p}_{p,j} = \mathcal{N}_{\mathrm{V}} \int \frac{\mathrm{d}x}{x}\mathrm{d}z\, \mathrm{d}\Phi_n(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};\mathcal{Q})\, \frac{1}{S_n}\mathcal{X}_3^{0,\mathrm{id.}p}(x,z) \\ \times M_n^0(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};xp_1,p_2)\, J(\{\ldots,\tilde{k}_p,\ldots\}_n,\eta\, z\, \tilde{k}_p)\,,\end{aligned} \tag{2.19}$$

with $\mathcal{N}_{\mathrm{V}} = \mathcal{N}_{\mathrm{R}} C(\epsilon) = \mathcal{N}_{\mathrm{B}}\, \overline{C}(\epsilon)$.

To achieve explicit pole cancellation at virtual level with an $n$-particle configuration as formulated in eq. (2.6), we further need to include mass-factorisation counterterms to the integrated subtraction terms (2.19). In the case of hadron-collider observables with an identified parton in the short distance cross section, those mass-factorisation counterterms will be needed for regulating collinear divergences coming from two different origins: initial-state collinear divergences associated with PDFs and final-state collinear singularities involving an identified parton that are associated with a FF.

With that in mind, one needs to consider mass-factorisation counterterms which are built with splitting kernels depending on both the initial-state momentum fraction $x$ as well as the final-state momentum fraction of the fragmenting parton $z$: For the one identified parton denoted as $p$, the general structure of this mass-factorisation counterterm has the following form:

$$\begin{aligned}\mathrm{d}\hat{\sigma}^{\mathrm{MF,id.}p}_{p,(if)}(\eta) = -\mathcal{N}_{\mathrm{V}} \int \frac{\mathrm{d}x}{x}\mathrm{d}z\, \mathrm{d}\Phi_n(k_1,\ldots,\tilde{k}_p,\ldots,k_n;xp_1,p_2) \\ \times \Gamma^{(1)}_{rs;kp}(x,z)\, \frac{1}{S_n} M_n^0(k_1,\ldots,\tilde{k}_p,\ldots,k_n;xp_1,p_2) \\ \times J(\{k_1,\ldots,\tilde{k}_p,\ldots,k_n\}_n,\eta\, z\, \tilde{k}_p)\,.\end{aligned} \tag{2.20}$$

We define the $x$- and $z$-dependent mass-factorisation kernel $\Gamma^{(1)}_{rs;kp}(x,z)$ as the sum of two terms depending either on $x$, and the factorisation scale ($\mu_F$) or on $z$ and the fragmentation scale ($\mu_D$) individually. It reads:

$$\Gamma^{(1)}_{rs;kp}(x,z) = \delta(1-z)\delta_{ps}\mu_F^{-2\epsilon}\Gamma^{(1)}_{k\leftarrow r}(x) + \delta(1-x)\delta_{kr}\mu_D^{-2\epsilon}\Gamma^{(1)}_{p\leftarrow s}(z)\,. \tag{2.21}$$

In this formula, the fragmenting parton is denoted by $p$ and there is in addition an implicit sum over $r$ and $s$, respectively denoting the particle types of the initial- and final-state partons. The general form of the mass-factorisation counterterms needed in space-like and time-like kinematics is well known. In the context of the antenna subtraction formalism, explicit expressions have been presented in [22] and [17] respectively.

Combining the expressions for the integrated subtraction term given in eq. (2.19) and for the mass factorisation counterterm as given in eq. (2.20), one can express the virtual

subtraction term given in eq. (2.13) in terms of fragmentation dipoles denoted as $\boldsymbol{J}_2^{(1),\mathrm{id}.p}$ below:

$$
\begin{aligned}
\mathrm{d}\hat{\sigma}_p^{\mathrm{Tid}.p}(\eta) &= -\mathcal{N}_{\mathrm{V}} \int \frac{\mathrm{d}x}{x} \mathrm{d}z\, \mathrm{d}\Phi_n(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};\mathcal{Q})\, \frac{1}{S_n} \left[ \mathcal{X}_3^{0,\mathrm{id}.p}(x,z) - \Gamma^{(1)}_{rs;kp}(x,z) \right] \\
&\quad \times M_n^0(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};xp_1,p_2)\, J(\{\ldots,\tilde{k}_p,\ldots\}_n,\eta\, z\, \tilde{k}_p)\,, \\
&= -\mathcal{N}_{\mathrm{V}} \int \frac{\mathrm{d}x}{x} \mathrm{d}z\, \mathrm{d}\Phi_n(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};\mathcal{Q})\, \frac{1}{S_n} \times \boldsymbol{J}_2^{(1),\mathrm{id}.p}(p_1,\tilde{k}_p,x,z) \\
&\quad \times M_n^0(k_1,\ldots,\tilde{k}_p,\ldots,k_{n+1};xp_1,p_2)\, J(\{\ldots,\tilde{k}_p,\ldots\}_n,\eta\, z\, \tilde{k}_p)\,.
\end{aligned}
\tag{2.22}
$$

## 3 Infrared structure of vector boson and hadron production

In this section, we provide explicit expressions needed to compute observables for the process $\mathrm{p\,p} \to V + h$ ($V = Z, W$) at NLO accuracy in the antenna subtraction formalism. The formulae are given in a form valid both for $Z$ and $W$ production as indicated by the label $V$ appearing in the matrix element and phase space. Before discussing the individual contributions, we first provide some general comments on our notation of the matrix elements and the overall factors used in the remainder of this section.

- The matrix element denoted as $B_n^\ell$ corresponds to processes involving one correlated $q\bar{q}$ pair, $n$ gluons and a vector boson at $\ell$ loops. At NLO, we will also require the sub-leading colour matrix element $\tilde{B}_n^\ell$, the $N_F$ part $\hat{B}_n^\ell$, as well as $C_n^\ell$ including two $q\bar{q}$ pairs and $D_n^\ell$, which is the interference part of the processes with two $q\bar{q}$ pairs of identical flavour.

- Although we employ a non-diagonal CKM matrix in our computation for the process involving a $W$ boson, we refrain from explicitly including the flavour assignment and CKM dependence for clarity and generality between the neutral- and charged-current processes. This information can easily be reinstated from the given formulae if necessary.

- For $V = W$, the initial state denoted as $q\bar{q}$ corresponds to $q\bar{q}''$, with $q''$ being the isospin partner of $q$. We assume that the flavour of $q$ ($\bar{q}$) is fixed (or becomes $\bar{q}''$ ($q''$) if a $W$ is coupled) while new quark flavours that appear at NLO are denoted as $q'$ ($\bar{q}'$), also admitting $q = q'$ ($\bar{q} = \bar{q}'$).

- The overall colour-independent factors that are denoted as $N_{ij}$ for $i, j = q, \bar{q}, g$ below depend on the initial-state configuration and are defined as

$$
N_{qg} = N_{\bar{q}g} = \frac{g_s^2 \mathcal{C}_V^2 (N^2-1)}{4N(N^2-1)} \quad \text{and} \quad N_{q\bar{q}} = N_{q\bar{q}'} = N_{qq} = N_{qq'} = \frac{g_s^2 \mathcal{C}_V^2 (N^2-1)}{[2N]^2}, \tag{3.1}
$$

where $g_s^2 = 4\pi\alpha_s$ and $\mathcal{C}_V = 2(4\pi\alpha_V(M_V))$ are the QCD and electroweak couplings respectively, where the latter differ for the $Z$ and $W$ cases [23]:

$$\alpha_W = \frac{G_F\, M_W^2\sqrt{2}}{4\pi}\,, \qquad \alpha_Z = \frac{G_F\, M_Z^2\sqrt{2}}{64\pi}\,. \tag{3.2}$$

- For ease of readability, we fix the position of the identified parton to be 3 with an additional superscript (id.), while the initial-state particles are assigned to positions 1 and 2.

### 3.1 Leading order

At leading order, the short-distance cross section for $V+h$ receives contributions from three different sub-processes with initial states, $qg$, $\bar{q}g$ and $q\bar{q}$:

$$\mathrm{d}\hat{\sigma}^{\mathrm{LO}} = \mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{(qg)} + \mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{(\bar{q}g)} + \mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{(q\bar{q})}\,. \tag{3.3}$$

They are in one-to-one correspondence to the case of an identified quark ($q$), anti-quark ($\bar{q}$), and gluon ($g$) and read

$$\mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{q} = N_{qg} B_1^0(1_q, 2_g, 3_q^{\mathrm{id.}}, 4_V) \times \mathrm{d}\Phi_2(k_V, k_3^{\mathrm{id.}}; p_1, p_2) J_1^{(1)}(\{k_3^{\mathrm{id.}}\}; \eta k_3^{\mathrm{id.}})\,, \tag{3.4}$$

$$\mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{\bar{q}} = N_{\bar{q}g} B_1^0(3_{\bar{q}}^{\mathrm{id.}}, 2_g, 1_{\bar{q}}, 4_V) \times \mathrm{d}\Phi_2(k_V, k_3^{\mathrm{id.}}; p_1, p_2) J_1^{(1)}(\{k_3^{\mathrm{id.}}\}; \eta k_3^{\mathrm{id.}})\,, \tag{3.5}$$

$$\mathrm{d}\hat{\sigma}^{\mathrm{LO}}_{g} = N_{q\bar{q}} B_1^0(1_q, 3_g^{\mathrm{id.}}, 2_{\bar{q}}, 4_V) \times \mathrm{d}\Phi_2(k_V, k_3^{\mathrm{id.}}; p_1, p_2) J_1^{(1)}(\{k_3^{\mathrm{id.}}\}; \eta k_3^{\mathrm{id.}})\,. \tag{3.6}$$

### 3.2 Next-to-leading order

In this section, we explicitly construct the NLO corrections to all sub-processes and possible cases of identified partons.

#### 3.2.1 Real matrix element contributions

The real corrections receive contributions from the following initial-state parton-level processes:

$$\mathrm{d}\hat{\sigma}^{\mathrm{R}} = \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(qg)} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(\bar{q}g)} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(q\bar{q})} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(gg)} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(qq')} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(q\bar{q}')} + \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{(\bar{q}\bar{q}')}\,, \tag{3.7}$$

which we decompose into separate identified-parton pieces. The case in which a quark $q$ is identified has contributions from the channels $qg$, $gg$, $qq'$ and $q\bar{q}'$:

$$\begin{aligned}\mathrm{d}\hat{\sigma}^{\mathrm{R}}_{q} = \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big( &\mathcal{M}^{\mathrm{R}}_{q,(qg)} + \mathcal{M}^{\mathrm{R}}_{q,(gg)} + \mathcal{M}^{\mathrm{R}}_{q,(qq')} + \mathcal{M}^{\mathrm{R}}_{q,(q\bar{q}')} \Big) \\ &\times J_1^{(2)}(\{k_3^{\mathrm{id.}}, k_j\}; \eta k_3^{\mathrm{id.}})\, \mathrm{d}\Phi_3(k_V, k_3^{\mathrm{id.}}, k_j; p_1, p_2)\,.\end{aligned} \tag{3.8}$$

The corresponding matrix elements denoted as $\mathcal{M}_q$ can be further colour decomposed according to the colour-ordered matrix elements introduced above:

$$\mathcal{M}^{\mathrm{R}}_{q,(qg)} = N_{qg} N \Big\{ B^0_2(1_q, 2_g, j_g, 3^{\mathrm{id.}}_q, 5_V) + [2_g \leftrightarrow j_g] - \frac{1}{N^2} \tilde{B}^0_2(1_q, 2_g, j_g, 3^{\mathrm{id.}}_q, 5_V) \Big\}, \tag{3.9}$$

$$\begin{aligned} \mathcal{M}^{\mathrm{R}}_{q,(gg)} &= N_{gg} N \sum_{i=u,d} N_i \Big\{ B^0_2(j_{\bar{i}}, 1_g, 2_g, 3^{\mathrm{id.}}_i, 5_V) + [1_g \leftrightarrow 2_g] \\ &\quad - \frac{1}{N^2} \tilde{B}^0_2(j_{\bar{i}}, 1_g, 2_g, 3^{\mathrm{id.}}_i, 5_V) \Big\}, \end{aligned} \tag{3.10}$$

$$\mathcal{M}^{\mathrm{R}}_{q,(qq')} = N_{qq'} \Big\{ C^0_0(1_q, j_{q'}, 2_{q'}, 3^{\mathrm{id.}}_q, 5_V) - \frac{1}{N} D^0_0(1_q, j_{q'}, 2_{q'}, 3^{\mathrm{id.}}_q, 5_V) \Big\}, \tag{3.11}$$

$$\mathcal{M}^{\mathrm{R}}_{q,(q\bar{q}')} = N_{q\bar{q}'} C^0_0(1_q, 2_{\bar{q}'}, j_{\bar{q}'}, 3^{\mathrm{id.}}_q, 5_V). \tag{3.12}$$

Note that in practice, all $N_F$ contributions are split into $N_u$ and $N_d$ types due to the different electroweak couplings of the $u$- and $d$-type quarks to the vector boson. Further, the identification of $q$ and $q'$ has to be performed independently since their flavour correlations are different.

When $q'$ is the identified parton, the real level short-distance contributions read:

$$\begin{aligned} \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{q'} &= \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big( \mathcal{M}^{\mathrm{R}}_{q',(qq')} + \mathcal{M}^{\mathrm{R}}_{q',(q\bar{q})} \Big) \\ &\quad \times J^{(2)}_1(\{k^{\mathrm{id.}}_3, k_j\}; \eta k^{\mathrm{id.}}_3)\, \mathrm{d}\Phi_3(k_V, k^{\mathrm{id.}}_3, k_j; p_1, p_2), \end{aligned} \tag{3.13}$$

with

$$\mathcal{M}^{\mathrm{R}}_{q',(qq')} = N_{qq'} \Big\{ C^0_0(1_q, 3^{\mathrm{id.}}_{q'}, 2_{q'}, j_q, 5_V) - \frac{1}{N} D^0_0(1_q, 3^{\mathrm{id.}}_{q'}, 2_{q'}, j_q, 5_V) \Big\}, \tag{3.14}$$

$$\mathcal{M}^{\mathrm{R}}_{q',(q\bar{q})} = N_{q\bar{q}} \Big\{ \sum_{i=u,d} N_i C^0_0(1_q, 3^{\mathrm{id.}}_i, j_{\bar{i}}, 2_{\bar{q}}, 5_V) - \frac{1}{N} D^0_0(1_q, 3^{\mathrm{id.}}_{\bar{q}}, j_q, 2_{\bar{q}}, 5_V) \Big\}, \tag{3.15}$$

where eq. (3.14) does not contain any singular limits. The contributions when the anti-quark $\bar{q}$ or $\bar{q}'$ are identified are respectively similar to eqs. (3.8)–(3.12) or eqs. (3.13)–(3.15) with the exchange of quarks and anti-quarks in both initial and final states.

We focus next on the real-level partonic contributions where a gluon $g$ is the identified particle. Those contributions arise from the $qg$, $\bar{q}g$ and $q\bar{q}$ initial-state combinations and read:

$$\begin{aligned} \mathrm{d}\hat{\sigma}^{\mathrm{R}}_{g} &= \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big( \mathcal{M}^{\mathrm{R}}_{g,(qg)} + \mathcal{M}^{\mathrm{R}}_{g,(\bar{q}g)} + \mathcal{M}^{\mathrm{R}}_{g,(q\bar{q})} \Big) \\ &\quad \times J^{(2)}_1(\{k^{\mathrm{id.}}_3, k_j\}; \eta k^{\mathrm{id.}}_3)\, \mathrm{d}\Phi_3(k_V, k^{\mathrm{id.}}_3, k_j; p_1, p_2), \end{aligned} \tag{3.16}$$

where the contributions from $\mathcal{M}^{\mathrm{R}}_{g,(qg)}$ and $\mathcal{M}^{\mathrm{R}}_{g,(\bar{q}g)}$ are the same upon the exchange $q \leftrightarrow \bar{q}$. Explicitly, the matrix elements labelled as $\mathcal{M}_g$ from the channels $qg$ and $q\bar{q}$ admit the

following colour decomposition:

$$\mathcal{M}^{\mathrm{R}}_{g,(qg)} = N_{qg} N \Big\{ B^0_2(1_q, 2_g, 3^{\mathrm{id.}}_g, j_q, 5_V) + [3^{\mathrm{id.}}_g \leftrightarrow 2_g] - \frac{1}{N^2} \tilde{B}^0_2(1_q, 2_g, 3^{\mathrm{id.}}_g, j_q, 5_V) \Big\}, \tag{3.17}$$

$$\mathcal{M}^{\mathrm{R}}_{g,(q\bar{q})} = N_{q\bar{q}} \frac{N}{2!} \Big\{ B^0_2(1_q, 3^{\mathrm{id.}}_g, j_g, 2_{\bar{q}}, 5_V) + [3^{\mathrm{id.}}_g \leftrightarrow j_g] - \frac{1}{N^2} \tilde{B}^0_2(1_q, 3^{\mathrm{id.}}_g, j_g, 2_{\bar{q}}, 5_V) \Big\}. \tag{3.18}$$

### 3.2.2 Real subtraction terms

The real-level subtraction terms are made of two categories of terms depending if the unresolved parton $j$ is colour connected to the identified parton or not. If the fragmenting particle is not colour connected to the unresolved parton, the subtraction terms mirror those for $V$ + jet production [23]. These terms are associated to initial–initial kinematical configurations for which we use the notation $\widetilde{1}_p$ and $\widetilde{2}_p$ and no «id.» index on the $X^3_0$ antenna functions. When instead, the fragmenting particle participates actively in the unresolved limit, an initial–final fragmentation subtraction of the type given in eq. (2.9) is needed. These terms include the fragmentation antennae denoted with the «id.» index and we use $\widetilde{1}_p$ or $\widetilde{2}_p$, together with $\widetilde{(3j)}^{\mathrm{id.}}_p$ to indicate the mapped momenta. These subtraction terms have been constructed for the first time for the computations presented in this work. For each of the identified particles, we present the real-level subtraction terms including both categories.

The real-level subtraction terms associated to parton $j$ becoming unresolved and corresponding to an identified quark $q$ denoted as $\mathrm{d}\hat{\sigma}^{\mathrm{S}}_q$ has itself four separate contributions related to the initial states $qg$, $gg$, $qq'$ and $q\bar{q}'$. The sum of those contributions takes the form:

$$\mathrm{d}\hat{\sigma}^{\mathrm{S}}_q = \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big( \mathcal{M}^{\mathrm{S}}_{q(q),(qg)} + \mathcal{M}^{\mathrm{S}}_{q(q),(gg)} + \mathcal{M}^{\mathrm{S}}_{q(q),(qq')} + \mathcal{M}^{\mathrm{S}}_{q(q),(q\bar{q}')} \Big) \times J^{(1)}_1(\{k^{\mathrm{id.}}_3\}; \eta k^{\mathrm{id.}}_3) \mathrm{d}\Phi_3(k_V, k^{\mathrm{id.}}_3, k_j; p_1, p_2). \tag{3.19}$$

The explicit expressions for the contributions from each initial-state channel read:

$$\mathcal{M}^{\mathrm{S}}_{q(q),(qg)} = N_{qg} N \Big\{ D^0_3(1_q, j_g, 2_g)\, B^0_1(\widetilde{1}_q, \widetilde{2}_g, 3^{\mathrm{id.}}_q, 4_V) + d^{0,\mathrm{id.}q}_3(3^{\mathrm{id.}}_q, j_g, 2_g)\, B^0_1(1_q, \widetilde{2}_g, \widetilde{(3j)}^{\mathrm{id.}}_q, 4_V) - \frac{1}{N^2} A^{0,\mathrm{id.}q}_3(3^{\mathrm{id.}}_q, j_g, 1_q)\, B^0_1(\widetilde{1}_q, 2_g, \widetilde{(3j)}^{\mathrm{id.}}_q, 4_V) \Big\}, \tag{3.20}$$

$$\mathcal{M}^{\mathrm{S}}_{q(q),(gg)} = -N_{gg} N_F N \Big(1 + \frac{1}{N^2}\Big) \Big( d^0_3(j_{\bar{q}}, 2_g, 1_g) B^0_1(3^{\mathrm{id.}}_q, \widetilde{1}_g, \widetilde{2}_q, 4_V)) - [\widetilde{1}_q \leftrightarrow \widetilde{2}_g] \Big), \tag{3.21}$$

$$\mathcal{M}^{\mathrm{S}}_{q(q),(q\bar{q}')} \overset{\bar{q}' \leftrightarrow q'}{=} \mathcal{M}^{\mathrm{S}}_{q(q),(qq')} = N_{gq'} E^0_3(1_q, j_{\bar{q}'}, 2_{q'})\, B^0_1(\widetilde{1}_q, \widetilde{2}_g, 3^{\mathrm{id.}}_q, 4_V). \tag{3.22}$$

The contribution where the quark $q'$ is identified has only a contribution from the initial state $q\bar{q}$. It reads:

$$\mathrm{d}\hat{\sigma}^{\mathrm{S}}_{q'} = \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big( \mathcal{M}^{\mathrm{S}}_{g(q'),(q\bar{q})} \Big) J^{(1)}_1(\{k^{\mathrm{id.}}_3\}; \eta k^{\mathrm{id.}}_3) \mathrm{d}\Phi_3(k_V, k^{\mathrm{id.}}_3, k_j; p_1, p_2). \tag{3.23}$$

This equation is related to a contribution in which the flavour of the identified particle in the reduced matrix element is not the same as the one in the real correction. This is usually denoted as an identity changing (IC) subtraction term. In this specific case, the quark $q'$ becomes a gluon $g$ in the reduced matrix element. Further expanding in products of antennae and reduced matrix elements, the individual contribution read:

$$\mathcal{M}^{\mathrm{S}}_{g(q'),(q\bar{q})} = N_{q\bar{q}} N_F E_3^{0,\mathrm{id}.q}(1_q, j_{\bar{q}'}, 3^{\mathrm{id}.}_{q'})\, B_1^0(\widetilde{1}_q, \widetilde{(3j)}^{\mathrm{id}.}_g, 2_{\bar{q}}, 4_V)\,, \tag{3.24}$$

where the flavour changing feature is specified with the labelling $g(q')$ on the left-hand side of the equation.

Finally, we present the explicit expression for the subtraction term involving the identified parton to be the gluon. The general expression is the sum of three terms:

$$\mathrm{d}\hat{\sigma}^{\mathrm{S}}_g = \mathrm{d}\hat{\sigma}^{\mathrm{S}}_{g(g)} + \mathrm{d}\hat{\sigma}^{\mathrm{S}}_{q(g)} + \mathrm{d}\hat{\sigma}^{\mathrm{S}}_{\bar{q}(g)}\,. \tag{3.25}$$

The first term denoted as $\mathrm{d}\hat{\sigma}^{\mathrm{S}}_{g(g)}$ is of identity preserving nature and reads:

$$\begin{aligned}\mathrm{d}\hat{\sigma}^{\mathrm{S}}_{g(g)} &= \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big(\mathcal{M}^{\mathrm{S}}_{g(g),(qg)} + \mathcal{M}^{\mathrm{S}}_{g(g),(\bar{q}g)} + \mathcal{M}^{\mathrm{S}}_{g(g),(q\bar{q})}\Big) \\ &\quad \times J_1^{(1)}(\{k_3^{\mathrm{id}.}\}; \eta k_3^{\mathrm{id}.}) \mathrm{d}\Phi_3(k_V, k_3^{\mathrm{id}.}, k_j; p_1, p_2)\,,\end{aligned} \tag{3.26}$$

with

$$\mathcal{M}^{\mathrm{S}}_{g(g),(\bar{q}g)} \overset{q\leftrightarrow\bar{q}}{=} \mathcal{M}^{\mathrm{S}}_{g(g),(qg)} = -N_{qg} N \Big\{ \Big(1 - \frac{1}{N^2}\Big) A_3^0(1_q, 2_g, j_q)\, B_1^0(\widetilde{1}_q, 3_g^{\mathrm{id}.}, \widetilde{2}_{\bar{q}}, 4_V) \Big\}\,, \tag{3.27}$$

$$\begin{aligned}\mathcal{M}^{\mathrm{S}}_{g(g),(q\bar{q})} &= N_{q\bar{q}} N \Big\{ \Big( d_3^{0,\mathrm{id}.g}(1_q, 3_g^{\mathrm{id}.}, j_g) + [j_g \leftrightarrow 3_g^{\mathrm{id}.}] \Big) B_1^0(\widetilde{1}_q, \widetilde{(3j)}^{\mathrm{id}.}_g, 2_{\bar{q}}, 4_V) \\ &\quad + \Big( d_3^{0,\mathrm{id}.g}(2_{\bar{q}}, i_g, 3_g^{\mathrm{id}.}) + [j_g \leftrightarrow 3_g^{\mathrm{id}.}] \Big) B_1^0(1_q, \widetilde{(3j)}^{\mathrm{id}.}_g, \widetilde{2}_{\bar{q}}, 4_V) \\ &\quad + \frac{1}{N^2} A_3^0(1_q, j_g, 2_{\bar{q}})\, B_1^0(\widetilde{1}_q, 3_g^{\mathrm{id}.}, \widetilde{2}_{\bar{q}}, 4_V) \Big\}\,.\end{aligned} \tag{3.28}$$

The second and third terms of eq. (3.25) are of identity changing nature and we have $\mathrm{d}\hat{\sigma}^{\mathrm{S}}_{\bar{q}(g)} \overset{q\leftrightarrow\bar{q}}{=} \mathrm{d}\hat{\sigma}^{\mathrm{S}}_{q(g)}$. In this case, only the $qg$ ($\bar{q}g$) channel contributes:

$$\mathrm{d}\hat{\sigma}^{\mathrm{S}}_{q(g)} = \alpha_s \frac{\bar{C}(\epsilon)}{C(\epsilon)} \Big(\mathcal{M}^{\mathrm{S}}_{q(g),(qg)}\Big) J_1^{(1)}(\{k_3^{\mathrm{id}.}\}; \eta k_3^{\mathrm{id}.}) \mathrm{d}\Phi_3(k_V, k_3^{\mathrm{id}.}, k_j; p_1, p_2)\,, \tag{3.29}$$

with

$$\begin{aligned}\mathcal{M}^{\mathrm{S}}_{q(g),(qg)} &= N_{qg} N \Big\{ d_3^{0,\mathrm{id}.g}(j_q, 3_g^{\mathrm{id}.}, 2_g)\, B_1^0(1_q, \widetilde{2}_g, \widetilde{(3j)}^{\mathrm{id}.}_q, 4_V) \\ &\quad + \frac{1}{N} A_3^{0,\mathrm{id}.g}(1_q, 3_g^{\mathrm{id}.}, j_q)\, B_1^0(\widetilde{1}_q, 2_g, \widetilde{(3j)}^{\mathrm{id}.}_q, 4_V) \Big\}\,.\end{aligned} \tag{3.30}$$

### 3.2.3 Virtual matrix element contributions

The virtual one-loop contributions arise from the same channels as the ones present at Born level and given in eq. (3.3). Those comprise three initial-state configurations that are associated with the identified partons being a quark, an anti-quark or a gluon,

$$\mathrm{d}\hat{\sigma}^{\mathrm{V}} = \alpha_s \bar{C}(\epsilon) \Big(\mathcal{M}^{\mathrm{V}}_{q,(qg)} + \mathcal{M}^{\mathrm{V}}_{\bar{q},(\bar{q}g)} + \mathcal{M}^{\mathrm{V}}_{g,(q\bar{q})}\Big) J_1^{(1)}(\{k_3^{\mathrm{id}.}\}; \eta k_3^{\mathrm{id}.}) \mathrm{d}\Phi_2(k_V, k_3^{\mathrm{id}.}; p_1, p_2)\,, \tag{3.31}$$

where,

$$\mathcal{M}^{\mathrm{V}}_{q,(qg)} = N_{qg}\, N\Big\{ B^1_1(1_q, 2_g, 3^{\mathrm{id.}}_q, 4_V) - \frac{1}{N^2}\tilde{B}^1_1(1_q, 2_g, 3^{\mathrm{id.}}_q, 4_V) + \frac{N_F}{N}\hat{B}^1_1(1_q, 2_g, 3^{\mathrm{id.}}_q, 4_V)\Big\}\,, \tag{3.32}$$

$$\mathcal{M}^{\mathrm{V}}_{g,(q\bar{q})} = N_{q\bar{q}} N\Big\{ B^1_1(1_q, 3^{\mathrm{id.}}_g, 2_{\bar{q}}, 4_V) - \frac{1}{N^2}\tilde{B}^1_1(1_q, 3^{\mathrm{id.}}_g, 2_{\bar{q}}, 4_V) + \frac{N_F}{N}\hat{B}^1_1(1_q, 3^{\mathrm{id.}}_g, 2_{\bar{q}}, 4_V)\Big\}\,, \tag{3.33}$$

with $\mathcal{M}^{\mathrm{V}}_{q,(\bar{q}g)} \overset{q\leftrightarrow\bar{q}}{=} \mathcal{M}^{\mathrm{V}}_{q,(qg)}$.

The infrared behaviour of the virtual matrix elements $B^1_1$, $\tilde{B}^1_1$ and $\hat{B}^1_1$ is the same regardless of the type of vector boson and can be expressed in terms of the Catani one-loop factorisation formula [24] with the colour-ordered singularity operators $\mathbf{I}^{(1)}_{ij}$ whose forms are given explicitly in the appendix of [22]. For the process at hand, we have three different crossings for the $B$-type matrix element, one for each channel. In the case of the $qg$-initiated process, the pole structure is

$$\mathrm{Poles}\Big(B^1_1(1_q, 2_g, 3_q, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qg}(\epsilon, s_{12}) + \mathbf{I}^{(1)}_{qg}(\epsilon, s_{23})\Big) \times B^0_1(1_q, 2_g, 3_q, 4_V)\,, \tag{3.34}$$

$$\mathrm{Poles}\Big(\tilde{B}^1_1(1_q, 2_g, 3_q, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qq}(\epsilon, s_{13})\Big) \times B^0_1(1_q, 2_g, 3_q, 4_V)\,, \tag{3.35}$$

$$\mathrm{Poles}\Big(\hat{B}^1_1(1_q, 2_g, 3_q, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qg,F}(\epsilon, s_{12}) + \mathbf{I}^{(1)}_{qg,F}(\epsilon, s_{23})\Big) \times B^0_1(1_q, 2_g, 3_q, 4_V)\,, \tag{3.36}$$

which is the same as for the $\bar{q}g$-channel. For $q\bar{q}$-initiated process and in the case of $Z+h$ or for the $q\bar{q}''$ initiated process in the case of $W+h$ production, the pole structure of the virtual matrix element contributions take the form:

$$\mathrm{Poles}\Big(B^1_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qg}(\epsilon, s_{13}) + \mathbf{I}^{(1)}_{qg}(\epsilon, s_{23})\Big) \times B^0_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\,, \tag{3.37}$$

$$\mathrm{Poles}\Big(\tilde{B}^1_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qq}(\epsilon, s_{12})\Big) \times B^0_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\,, \tag{3.38}$$

$$\mathrm{Poles}\Big(\hat{B}^1_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\Big) = 2\Big(\mathbf{I}^{(1)}_{qg,F}(\epsilon, s_{13}) + \mathbf{I}^{(1)}_{qg,F}(\epsilon, s_{23})\Big) \times B^0_1(1_q, 3_g, 2_{\bar{q}}, 4_V)\,. \tag{3.39}$$

### 3.2.4 Virtual subtraction terms

As discussed in section 2.1.2, we construct the virtual-level subtraction term as

$$\mathrm{d}\hat{\sigma}^{\mathrm{T}}(\eta, x_1, x_2, z) = -\int_1 \mathrm{d}\hat{\sigma}^{\mathrm{S}}(\eta, x_1, x_2, z) - \mathrm{d}\hat{\sigma}^{\mathrm{MF}}(\eta, x_1, x_2, z)\,, \tag{3.40}$$

where $\int_1 \mathrm{d}\hat{\sigma}^{\mathrm{S}}$ is the integrated real-level subtraction term and $\mathrm{d}\hat{\sigma}^{\mathrm{MF}}$ the mass-factorisation term. We combine both terms using integrated fragmentation dipoles $\boldsymbol{J}^{(1),\mathrm{id.}p}_2$ as defined in eq. (2.22) as well as the standard integrated dipoles $\boldsymbol{J}^{(1)}_2$ [22].

We present the complete virtual subtraction term separately for the different identified partons. For the identified-quark ($q$) contribution, the virtual subtraction term denoted as $\mathrm{d}\hat{\sigma}^{\mathrm{T}}_q$ reads:

$$\mathrm{d}\hat{\sigma}^{\mathrm{T}}_q = \int \frac{\mathrm{d}x_1}{x_1}\frac{\mathrm{d}x_2}{x_2}\mathrm{d}z\, \alpha_s \bar{C}(\epsilon)\big(\mathcal{M}^{\mathrm{T}}_{q(q),(qg)} + \mathcal{M}^{\mathrm{T}}_{q(q),(gg)} + \mathcal{M}^{\mathrm{T}}_{q(q),(qq')} + \mathcal{M}^{\mathrm{T}}_{q(q),(q\bar{q}')}\big) \times J^{(1)}_1(\{k^{\mathrm{id.}}_3\}; \eta k^{\mathrm{id.}}_3)\mathrm{d}\Phi_2(k_V, k^{\mathrm{id.}}_3; p_1, p_2)\,, \tag{3.41}$$

where

$$\mathcal{M}^{\mathrm{T}}_{q(q),(qg)} = N_{qg}N\Big\{ \big(\boldsymbol{J}^{(1),II}_{2,QG}(s_{12}) + \boldsymbol{J}^{(1),FI,\mathrm{id}.q}_{2,QG}(s_{23}) - \frac{1}{N^2}\boldsymbol{J}^{(1),IF,\mathrm{id}.q}_{2,QQ}(s_{13}) - 2N_F\boldsymbol{J}^{(1),II}_{2,hQG}(s_{12})\big) \times B^0_1(1_q,2_g,3^{\mathrm{id}.}_q,4_V)\Big\} J^{(1)}_1(\{k^{\mathrm{id}.}_3\};\eta k^{\mathrm{id}.}_3)\,, \tag{3.42}$$

$$\mathcal{M}^{\mathrm{T}}_{q(q),(gg)} = N_{gg}N_F\big(N-1/N\big)\Big\{ -\boldsymbol{J}^{(1),II}_{2,GQ,g\to q}(s_{12})B^0_1(1_q,2_g,3^{\mathrm{id}.}_q,4_V) - \boldsymbol{J}^{(1),II}_{2,QG,g\to q}(s_{12})B^0_1(2_q,1_g,3^{\mathrm{id}.}_q,4_V)\Big\}\,, \tag{3.43}$$

$$\mathcal{M}^{\mathrm{T}}_{q(q),(q\bar{q}')} \overset{q\leftrightarrow\bar{q}}{=} \mathcal{M}^{\mathrm{T}}_{q(q),(qq')} = -N_{qq'}\boldsymbol{J}^{(1),II}_{2,QG,q'\to g}(s_{12})\,B^0_1(1_q,2_g,3^{\mathrm{id}.}_q,4_V)\,. \tag{3.44}$$

In the case where the quark $q'$ is the identified particle, the virtual subtraction term $\mathrm{d}\hat{\sigma}^{\mathrm{T}}_{q'}$ reads:

$$\mathrm{d}\hat{\sigma}^{\mathrm{T}}_{q'} = \int \frac{\mathrm{d}x_1}{x_1}\frac{\mathrm{d}x_2}{x_2}\mathrm{d}z\,\alpha_s\bar{C}(\epsilon)\Big(\mathcal{M}^{\mathrm{T}}_{g(q'),(q\bar{q})}\Big)J^{(1)}_1(\{k^{\mathrm{id}.}_3\};\eta k^{\mathrm{id}.}_3)\mathrm{d}\Phi_2(k_V,k^{\mathrm{id}.}_3;p_1,p_2)\,, \tag{3.45}$$

with

$$\mathcal{M}^{\mathrm{T}}_{g(q'),(q\bar{q})} = -N_{qg}N_F\boldsymbol{J}^{(1),IF,\mathrm{id}.q'}_{2,QG,q\leftarrow g}(s_{13})B^0_1(1_q,3^{\mathrm{id}.}_g,2_{\bar{q}},,4_V)\,. \tag{3.46}$$

For the case where $\bar{q}$ and $\bar{q}'$ are the identified partons, the terms have the same structure as for $q$ and $q'$ respectively. Finally, the case in which the gluon is the identified parton is given by:

$$\mathrm{d}\hat{\sigma}^{\mathrm{T}}_{g} = \int \frac{\mathrm{d}x_1}{x_1}\frac{\mathrm{d}x_2}{x_2}\mathrm{d}z\,\alpha_s\frac{\bar{C}(\epsilon)}{C(\epsilon)}\Big(\mathcal{M}^{\mathrm{T}}_{g(g),(qg)} + \mathcal{M}^{\mathrm{T}}_{g(g),(\bar{q}g)} + \mathcal{M}^{\mathrm{T}}_{g(g),(q\bar{q})} + \mathcal{M}^{\mathrm{T}}_{q(g),(qg)} + \mathcal{M}^{\mathrm{T}}_{q(g),(\bar{q}g)}\Big) \times J^{(1)}_1(\{k^{\mathrm{id}.}_3\};\eta k^{\mathrm{id}.}_3)\mathrm{d}\Phi_2(k_V,k^{\mathrm{id}.}_3;p_1,p_2)\,, \tag{3.47}$$

where

$$\mathcal{M}^{\mathrm{T}}_{g(g),(qg)} = -N_{qg}\Big(N-\frac{1}{N}\Big)\boldsymbol{J}^{(1),II}_{2,QQ,g\to q}(s_{12})B^0_1(1_q,3^{\mathrm{id}.}_g,2_{\bar{q}},4_V)\,, \tag{3.48}$$

$$\mathcal{M}^{\mathrm{T}}_{g(g),(q\bar{q})} = -N_{qg}N\Big(\boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,GQ}(s_{23}) + \boldsymbol{J}^{(1),IF,\mathrm{id}.g}_{2,QG}(s_{13}) + \boldsymbol{J}^{(1),II}_{2,QQ}(s_{12}) + N_F\big(\boldsymbol{J}^{(1),IF,\mathrm{id}.q'}_{2,QG,q\leftarrow g}(s_{13}) + \boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,hQG}(s_{13})\big)\Big)B^0_1(1_q,3^{\mathrm{id}.}_g,2_{\bar{q}},4_V)\,, \tag{3.49}$$

$$\mathcal{M}^{\mathrm{T}}_{q(g),(qg)} = N_{qg}N\Big(-\boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,QG,g\leftarrow q}(s_{23}) - \frac{1}{N^2}\boldsymbol{J}^{(1),IF,\mathrm{id}.g}_{2,QQ,g\leftarrow q}(s_{13})\Big)B^0_1(1_q,2_g,3^{\mathrm{id}.}_q,4_V)\,, \tag{3.50}$$

with the contributions from the $\bar{q}g$ channel being the same as the $qg$ channel after the exchange of the quark for an anti-quark.

The virtual subtraction dipoles denoted as $\boldsymbol{J}^{(1)}_2$ can be separated into two categories depending whether there is (or not) a colour connection between the identified particle and the unresolved parton. In the case where there is no colour connection, the virtual subtraction dipoles are all of the initial–initial type. Those are known [22] and recalled

here for completeness:

$$\boldsymbol{J}^{1,II}_{2,QG} = \delta(1-z)\left[\mathcal{D}^{0}_{3,qg}(x_1,x_2) - \delta(1-x_2)\Gamma^{(1)}_{qq}(x_1) - \frac{1}{2}\delta(1-x_1)\Gamma^{(1)}_{gg}(x_2)\right] \tag{3.51}$$

$$\boldsymbol{J}^{(1),II}_{2,QQ} = \delta(1-z)\left[\mathcal{A}^{0}_{3,q\bar{q}}(x_1,x_2) - \delta(1-x_2)\Gamma^{(1)}_{qq}(x_1) - \delta(1-x_1)\Gamma^{(1)}_{qq}(x_2)\right], \tag{3.52}$$

$$\boldsymbol{J}^{(1),II}_{2,QQ,g\to q} = \delta(1-z)\left[-\mathcal{A}^{0}_{3,qg}(x_1,x_2) - \delta(1-x_1)\Gamma^{(1)}_{qg}(x_2)\right], \tag{3.53}$$

$$\boldsymbol{J}^{(1),II}_{2,GQ,g\to q} = \delta(1-z)\left[-\mathcal{D}^{0}_{3,gg}(x_1,x_2) - \delta(1-x_2)\Gamma^{(1)}_{qg}(x_1)\right], \tag{3.54}$$

$$\boldsymbol{J}^{(1),II}_{2,QG,g\to q} = \delta(1-z)\left[-\mathcal{D}^{0}_{3,gg}(x_1,x_2) - \delta(1-x_1)\Gamma^{(1)}_{qg}(x_2)\right], \tag{3.55}$$

$$\boldsymbol{J}^{(1),II}_{2,QG,q'\to g} = \delta(1-z)\left[-\mathcal{E}^{0}_{3,q'q}(x_1,x_2) - \delta(1-x_2)\Gamma^{(1)}_{gq}(x_1)\right], \tag{3.56}$$

$$\boldsymbol{J}^{(1),II}_{2,hQG} = -\frac{1}{2}\delta(1-z)\left[\delta(1-x_1)\Gamma^{(1)}_{gg,F}(x_2)\right]. \tag{3.57}$$

For the case where the identified particle participates in the unresolved limit, the fragmentation dipoles are derived here for the first time. Those are separated according to the nature of the identified parton involved. In the case where a quark is identified, the required dipoles read:

$$\boldsymbol{J}^{(1),FI,\mathrm{id}.q}_{2,QG} = \delta(1-x_1)\left[\mathcal{D}^{0,\mathrm{id}.q}_{3,g}(x_2,z) - \delta(1-x_2)\Gamma^{(1)}_{qq}(z) - \frac{1}{2}\delta(1-z)\Gamma^{(1)}_{gg}(x_2)\right], \tag{3.58}$$

$$\boldsymbol{J}^{(1),IF,\mathrm{id}.q}_{2,QQ} = \delta(1-x_2)\left[\mathcal{A}^{0,\mathrm{id}.q}_{3,q}(x_1,z) - \delta(1-x_1)\Gamma^{(1)}_{qq}(z) - \delta(1-z)\Gamma^{(1)}_{qq}(x_1)\right], \tag{3.59}$$

$$\boldsymbol{J}^{(1),IF,\mathrm{id}.q'}_{2,QG,q\leftarrow g} = \delta(1-x_2)\left[-\mathcal{E}^{0,\mathrm{id}.q'}_{3,q}(x_1,z) + \delta(1-x_1)\Gamma^{(1)}_{qg}(z)\right]. \tag{3.60}$$

For the identified-gluon case, the dipoles used here instead read:

$$\boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,GQ} = \delta(1-x_1)\left[\mathcal{D}^{0,\mathrm{id}.g}_{3,q}(x_2,z) - \delta(1-z)\Gamma^{(1)}_{qq}(x_1) - \frac{1}{2}\delta(1-x_1)\Gamma^{(1)}_{gg}(z)\right], \tag{3.61}$$

$$\boldsymbol{J}^{(1),IF,\mathrm{id}.g}_{2,QG} = \delta(1-x_2)\left[\mathcal{D}^{0,\mathrm{id}.g}_{3,q}(x_1,z) - \delta(1-z)\Gamma^{(1)}_{qq}(x_2) - \frac{1}{2}\delta(1-x_2)\Gamma^{(1)}_{gg}(z)\right], \tag{3.62}$$

$$\boldsymbol{J}^{(1),IF,\mathrm{id}.g}_{2,QQ,g\leftarrow q} = \delta(1-x_2)\left[-\mathcal{A}^{0,\mathrm{id}.g}_{3,q}(x_1,z) + \delta(1-x_1)\Gamma^{(1)}_{gq}(z)\right], \tag{3.63}$$

$$\boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,QG,g\leftarrow q} = \delta(1-x_1)\left[-\mathcal{D}^{0,\mathrm{id}.g}_{3,g}(x_2,z) + \delta(1-x_2)\Gamma^{(1)}_{gq}(z)\right], \tag{3.64}$$

$$\boldsymbol{J}^{(1),FI,\mathrm{id}.g}_{2,hQG} = -\frac{1}{2}\delta(1-x_2)\left[\delta(1-x_1)\Gamma^{(1)}_{gg,F}(z)\right]. \tag{3.65}$$

In the final-state identity-changing cases we use a right-to-left arrow, while in the initial-state identity-changing cases we use a left-to-right arrow. We can sum all virtual subtraction terms presented in section 3.2.4 and the corresponding virtual-level matrix element contributions presented in section 3.2.3 and show that it is free from explicit poles,

$$Poles\Bigg(\mathrm{d}\hat{\sigma}^{\mathrm{V}}(\eta) - \underbrace{\Big(-\int_1 \mathrm{d}\hat{\sigma}^{\mathrm{S}}(\eta) - \mathrm{d}\hat{\sigma}^{\mathrm{MF}}(\eta)\Big)}_{\equiv \mathrm{d}\hat{\sigma}^{\mathrm{T}}(\eta)}\Bigg) = 0\,. \tag{3.66}$$

The finite remainder can then be integrated numerically.

## 4 $Z$ boson in association with a hadron inside a jet

In this section we provide predictions for observables related to the production of charged hadrons inside hadronic jets in association with a leptonically decaying $Z$-boson. We compare our results with the corresponding LHCb measurement [11] at $\sqrt{s} = 13$ TeV and consider both unidentified charged hadron as well as pion production.

### 4.1 Observable definition and setup

Following [11], the LHCb measurement is performed by considering events with at least one reconstructed hadronic jet and by identifying a hadron inside it. In case multiple jets are reconstructed in the event, the one with the largest transverse momentum ($j_1$) is considered. Determining the longitudinal momentum fraction carried by the hadron within the jet,

$$z = \frac{\mathbf{p}_h \cdot \mathbf{p}_{j_1}}{|\mathbf{p}_{j_1}|^2}\,, \tag{4.1}$$

where $\mathbf{p}_h$ and $\mathbf{p}_{j_1}$ are the hadron and the jet three-momenta respectively, the measured observable is given by the normalised distribution

$$F_{\text{exp.}}(z) = \frac{1}{N_{Z+\text{jet}}}\frac{\mathrm{d}N_h(z)}{\mathrm{d}z}, \tag{4.2}$$

where $N_h(z)$ is the number of events with a hadron carrying a longitudinal momentum fraction $z$ inside the leading jet, while the normalization factor $N_{Z+\text{jet}}$ corresponds to the number of $Z+\text{jet}$ events. On the theory side, eq. (4.2) is mirrored by

$$F_{\text{th.}}(z) = \frac{1}{\sigma_{Z+\text{jet}}}\frac{\mathrm{d}\sigma_{Z+h}}{\mathrm{d}z}, \tag{4.3}$$

where the differential distribution present in the numerator require the presence of both a jet and a hadron inside it, whereas the inclusive cross section $\sigma_{Z+\text{jet}}$ only demands the presence of a jet.

The LHCb measurement is performed within the following fiducial volume:

$$\begin{gathered} 20 < p_{T,j_1} < 100 \text{ GeV}, \quad 2 < \eta_{j_1} < 4, \\ p_h > 4 \text{ GeV}, \quad p_{T,h} > 0.25 \text{ GeV}, \quad \Delta R(j_1, h) < 0.5, \\ p_{T,\ell^\pm} > 20 \text{ GeV}, \quad 2.5 < \eta_{\ell^\pm} < 4.5, \quad \Delta R(j_1, \ell) > 0.5, \quad \Delta R(j_1, Z) > \tfrac{7\pi}{8}\,, \end{gathered} \tag{4.4}$$

and jets are reconstructed using the anti-$k_T$ algorithm [26] with cone size $R = 0.5$.

We consider predictions for $F(z)$ in three different $p_{T,j_1}$ intervals: $20 < p_{T,j_1} < 30$ GeV, $30 < p_{T,j_1} < 50$ GeV and $50 < p_{T,j1} < 100$ GeV. We adopt the NNPDF3.1 [27] PDF set, with $\alpha_s = 0.118$ and $n_f^{\max} = 5$: both the PDF and the $\alpha_s$ values are accessed through the LHAPDF library [28]. Lastly, the central renormalisation and factorisation scale are chosen to be equal to

$$\mu_R = \mu_F = \sqrt{m_Z^2 + \sum_{j\in\text{jets}} p_{T,j}^2}\,, \tag{4.5}$$

where the sum runs over all the jets in the event. For the fragmentation scale, on the other hand, we choose the central scale

$$\mu_D = R \cdot p_{T,j_1} \,, \tag{4.6}$$

where $R = 0.5$ is the cone size used in the jet algorithm. This particular choice for the scale $\mu_D$ is motivated by the fact that the hadron in the process definition must be contained within the leading jet, thereby indicating that the resolution scale for the fragmentation process should be related to the jet characteristics. For the numerator, the theoretical uncertainties are determined by the envelope of the common 7-point scale variation, by correlating the variation of $\mu_F$ and $\mu_D$ and otherwise independently halving or doubling the scales and omitting the most extreme variations. In the case of the denominator, we determine the $Z +$ jet cross section at the same central scale as in eq. 4.5.

We provide two categories of predictions depending on the hadron types produced inside the jet recoiling against the $Z$-boson. In section 4.2 we consider the sum of unidentified charged hadrons, including pion, kaon and proton contributions. In section 4.3 we separately consider the pion contribution, the latter being by far the dominant contribution (kaons and protons contribute in a similar way, with the size of their contributions amounting to 20–30% of the pion contribution).

### 4.2 Results for unidentified charged hadron production

Predictions for charged hadron production are computed using three different sets of fragmentation functions: BKK [29], NNFF1.1 [30], DSS07 [31, 32]. BKK and DSS07 have been directly implemented in the NNLOJET infrastructure, while NNFF1.1 is available through the LHAPDF interface. From these, only NNFF1.1 provides error sets, thus we refrain from showing the corresponding errors in our results so that the treatment is the same in the three cases, allowing us to compare better their features. In the following, we provide some details on each of the three fragmentation function sets.

The BKK set [29] was obtained by fits to LEP and HERA data from the early 90's. The unidentified charged hadron FF is defined as the arithmetic sum of the charged pion and kaon contributions. The parton-to-charged pions and parton-to-charged kaons FFs are provided as simple analytic functions which encode the DGLAP evolution at LO and NLO level. The NNFF1.1 set [30] does not consider individual charged hadron contributions, but directly determine an unidentified charged hadron FF through a global NLO fit of a large variety of measurements at $e^+e^-$ colliders, the Tevatron and LHC. In the DSS07 set [31, 32], the determination of the unidentified charged hadron FF is related to the FFs for identified pions, kaons and (anti-)protons by imposing the sum rule:

$$D_a^{h^\pm} = D_a^{\pi^\pm} + D_a^{K^\pm} + D_a^{p/\bar{p}} + D_a^{\mathrm{res.}^\pm} \tag{4.7}$$

as a constraint. In particular, $D_a^{\pi^\pm}, D_a^{K^\pm}$ and $D_a^{p/\bar{p}}$ are individually fitted, and then used as input in a subsequent fit to determine the residual charged hadron contribution $D_a^{\mathrm{res.}^\pm}$. The different FFs are determined through a global NLO fit of data from electron–positron annihilation, proton–proton collisions and deep-inelastic lepton–proton scattering.

Figure 1 presents our predictions for $F(z)$ for three different $p_{T,j1}$ intervals and using the three different FF sets. Predictions using the BKK parametrization are shown on the top row, NNFF1.1 in the middle row, and DSS07 in the bottom row. Each row has three sub-figures corresponding to three different $p_{T,j1}$ intervals: [20, 30], [30, 50], [50, 100] GeV. Each sub-figure is made of two panels: the top panel shows the $F(z)$ distributions, and the bottom one shows the ratio to NLO.

For low $z$ values, we do not expect fixed-order perturbative result to be adequate in describing the data, as this region prone to soft physics, where resummation becomes mandatory.

We first note that when moving from LO to NLO, predictions change in shape, with the size of NLO corrections increasing towards high $z$ values to up to 25%. We also note a reduction of scale uncertainties. We find that predictions using the BKK and DSS07 charged hadron FF best describe the data: for mid $z$ values, both sets show alignment between NLO and data for the three $p_{T,j_1}$ bins, with the exception of the [30, 50] GeV $p_{T,j_1}$ interval where BKK underestimates the data; for high $z$ values, NLO predictions with both BKK and DSS07 fail to describe the data in the highest $p_{T,j_1}$ interval, with the LO predictions being closer to the experimental data points. In the lower $p_{T,j_1}$ cases, the agreement at high $z$ values improves, most notably in the middle [30, 50] GeV $p_{T,j_1}$ interval. In contrast, the results computed using NNFF1.1 are consistently above the data. Furthermore, the shape of the distributions differ for each FF.

It is interesting to have a closer look into the hardest available region, which corresponds to the third interval in transverse momentum, i.e. we focus on the $50 < p_{T,j_1} < 100$ GeV bin and restricting the $z$ range to values $z > 0.02$. This is shown in Fig. 2, where we plot the ratio to data of all three NLO predictions for $F(z)$. We observe a compatibility of the results for the three FFs in the highest and lowest bins of Fig. 2 within the shown range: for $z < 0.06$, data and theory are compatible within uncertainties, while for $z > 0.35$ theory predictions start to be above the data, with larger experimental uncertainty. In the intermediate range of z, BKK and DSS07 seem to provide a better description of data compared to NNFF1.1.

### 4.3 Results for charged pion production

We now move on to $Z$-tagged events associated with the production of charged pions inside jets. We adopt the pion FFs from BKK [29], NNFF1.0 [33] and DSS07 [31]. Note that in this context, we adopt the NNFF1.0 set instead of NNFF1.1 above: this is due to the fact that NNFF1.1 only provides an unidentified charged hadron set, whereas NNFF1.0 provides individual pions, kaons and (anti-)proton FF sets. As for NNFF1.1, also NNFF1.0 is provided via the LHAPDF interface.

Our results for $F(z)$ using the BKK, NNFF1.0 and DSS07 pion FFs are shown in Fig. 3 using the same layout as in Fig. 1.

When comparing Fig. 3 to the analogous plot for unidentified hadrons in Fig. 1, we notice similar features. However, compared to Fig. 1, in the mid $z$ region, the results computed using NNFF1.0 and DSS07 pion FFs show a tendency to overshoot the data, whereas the results using the BKK FF describe the data in a satisfactory manner. In

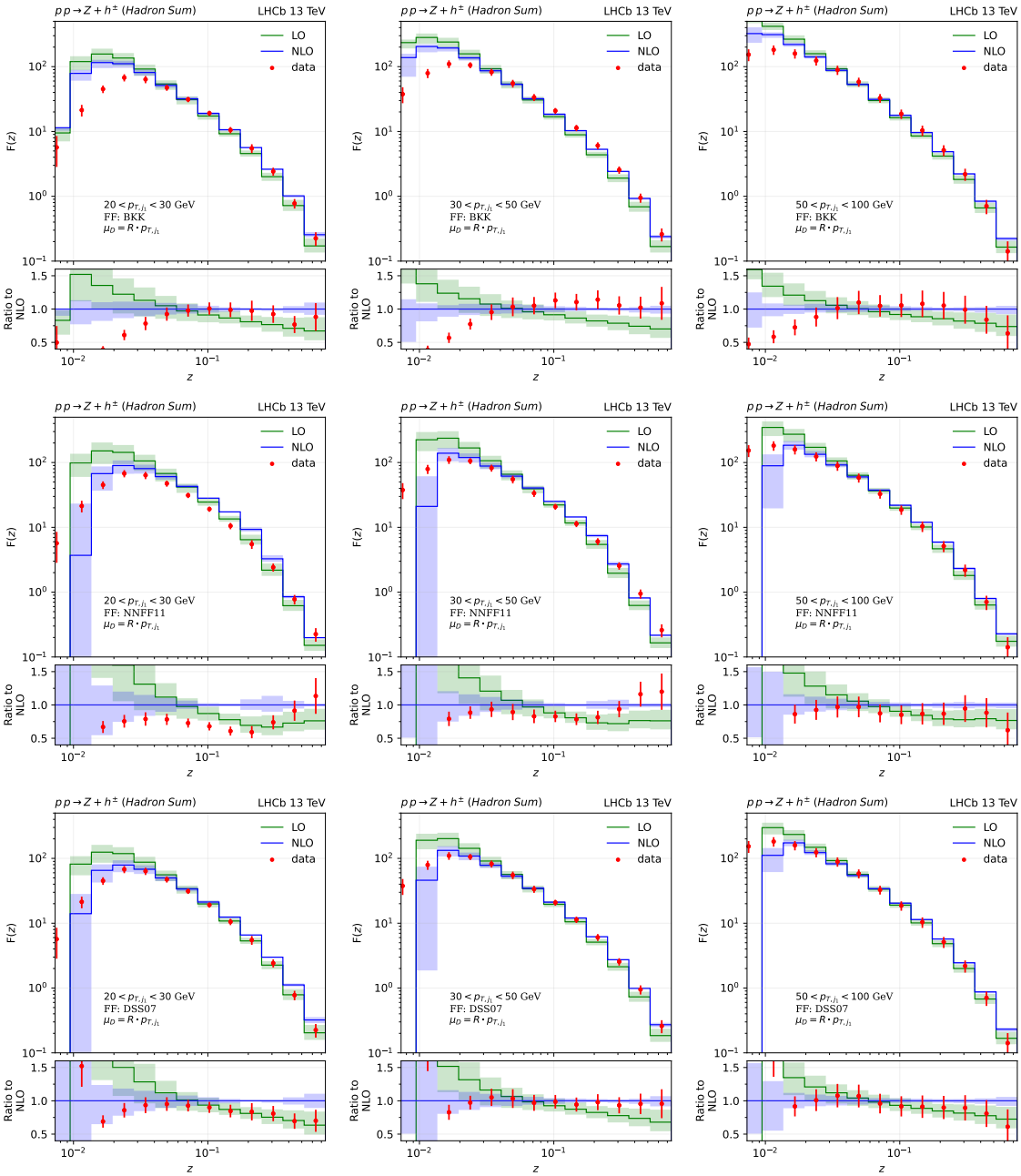


**Figure 1**. Comparison of LO (green) and NLO (blue) results for $F(z)$ obtained with the BKK (top row), NNFF1.1 (middle row) and DSS07 (bottom row) fragmentation functions with LHCb data (red dots) at 13 TeV. From left to right, each column corresponds to the $p_{T,j_1}$ ranges [20, 30], [30, 50] and [50, 100] GeV. In each figure, the top panels show the $F(z)$ distributions and the bottom one shows the ratio to NLO.

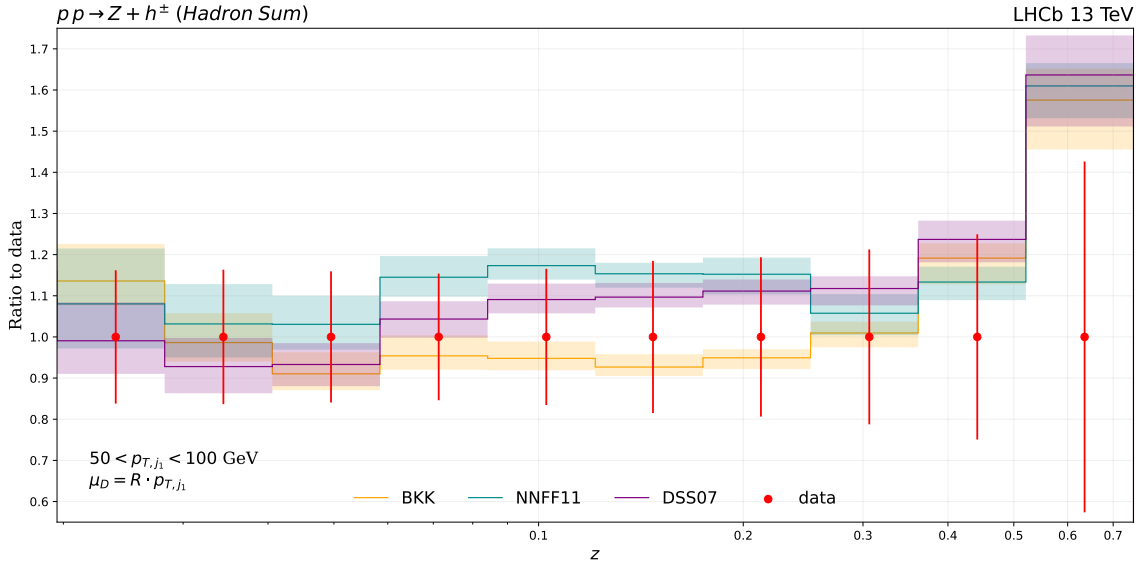


**Figure 2**. Comparison of the NLO results obtained for the different sets of FFs shown in Fig. 1 for $p_{T,j_1} \in [50, 100]$ GeV. Note that when comparing with Fig. 1, the lower $z$ edge is here increased to $z \sim 0.02$.

the high $z$ region, BKK and DSS07 seem to better describe the data compared to the unidentified hadron case, with the BKK set offering the best description. In all three $p_{T,j_1}$ intervals, for intermediate to high values of $z$, we find that the NLO corrections give positive corrections of up to 25% with respect to the LO results.

In order to further study the impact of the different FFs on the theory predictions, in Fig. 4 we compare the $F(z)$ distributions for pion production with the three FF sets in the highest $p_{T,j_1}$ bin. Predictions using BKK fragmentation functions provide the best description of data, while the three sets are compatible among each other in the high $z$-region.

In summary, from our analysis at NLO level, we find that the three FFs considered (BKK, NNFF1.0 or NNFF1.1, DSS07) show qualitative differences in the description of LHCb hadron-in-jet data, with none of the three sets able to describe all the kinematical bins equally well. It is likely that this dataset is able to offer important constraints on FFs when included in global fits. Indeed, such LHCb dataset has been included in a very recent NLO global analysis of FFs [34]: it has been shown how these hadron-in-jet data are useful to better determine the gluon-to-hadron fragmentation functions, which other datasets leave largely unconstrained. As our goal here is to study how publicly available FFs obtained through fits of one-particle inclusive datasets are able to describe this new exclusive dataset, we refrain from presenting results obtained using the FFs of [34].

## 5 $W$ boson in association with a charmed hadron

In this section, we present predictions for the associated production of a $W$ boson with a $D^{(*)}$ meson. More specifically, the following processes are considered: $W^+ + D^-$, $W^- + D^+$,

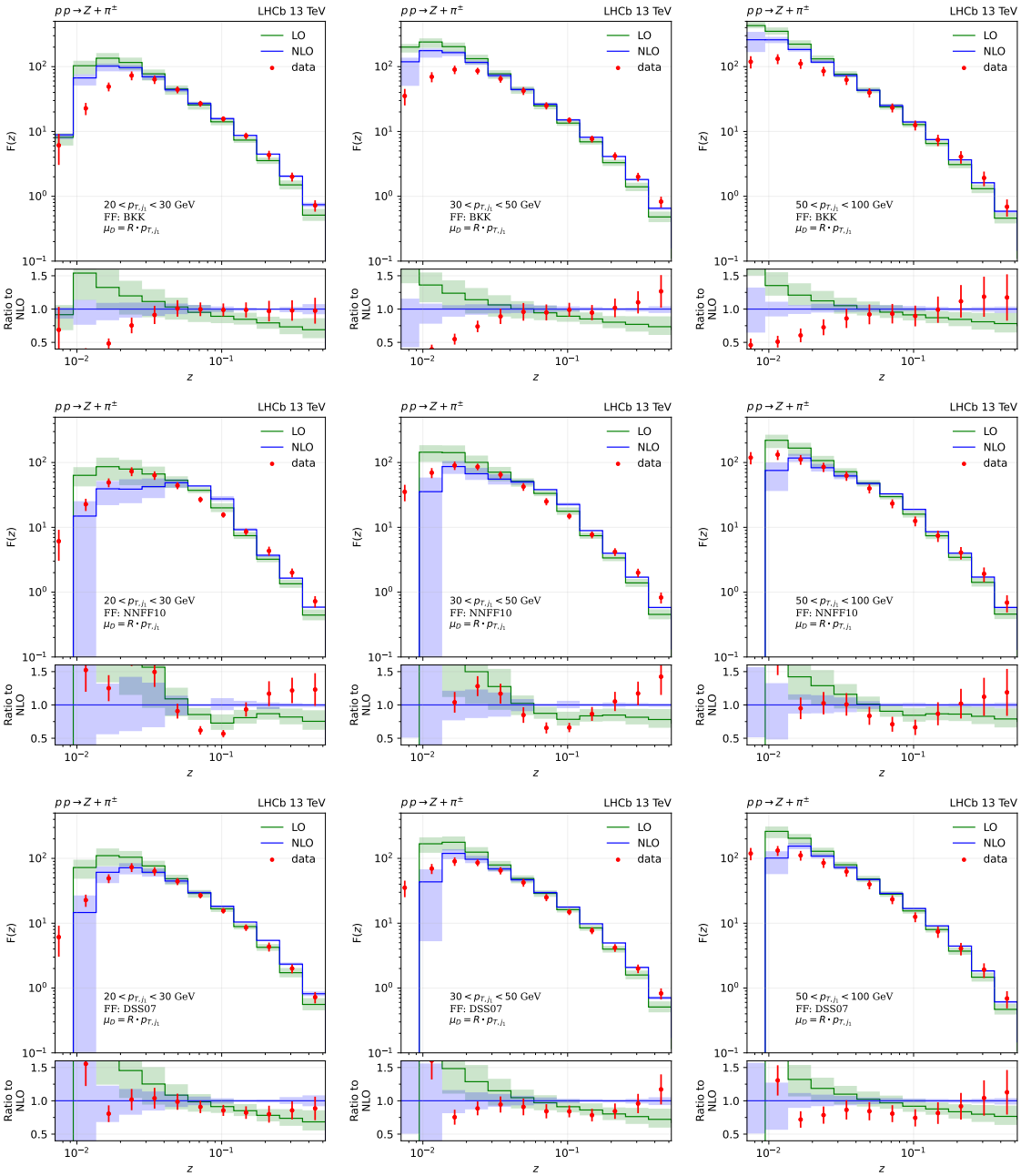


**Figure 3**. Comparison of LO (green) and NLO (blue) results for $F(z)$ obtained with the BKK (top row), NNFF1.0 (middle row) and DSS07 (bottom row) fragmentation functions for pions with LHCb data (red dots) at 13 TeV. From left to right, each column corresponds to the $p_{T,j_1}$ ranges [20, 30], [30, 50] and [50, 100] GeV. The top panel shows the $F(z)$ distributions and the bottom one shows the ratio to NLO.

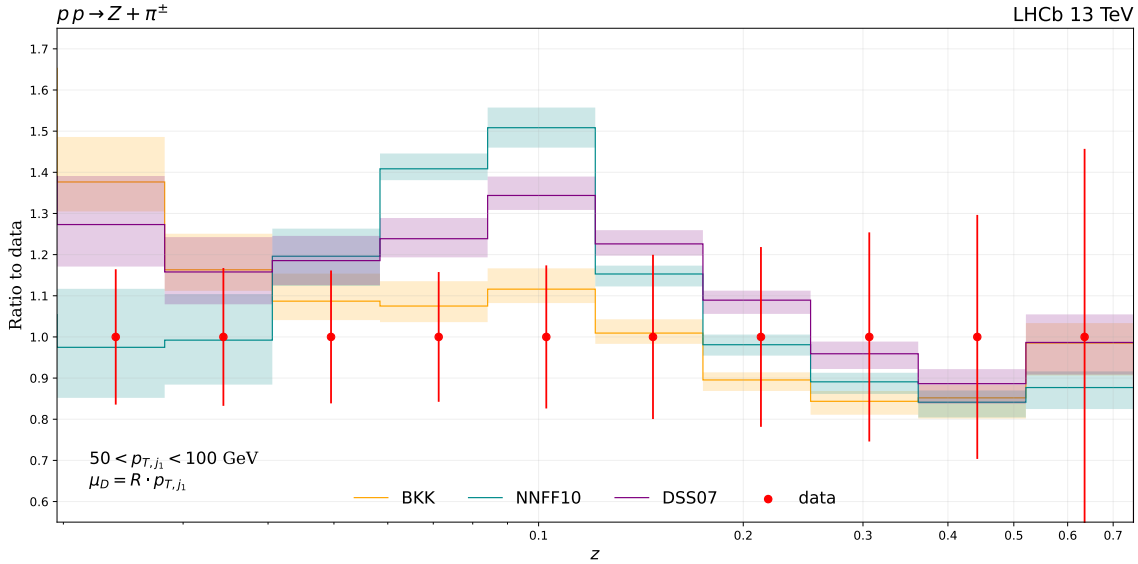


**Figure 4**. Comparison of the NLO results obtained for the different sets of fragmentation functions shown in Fig. 3 in the region $p_{T,j_1} \in [50, 100]$ GeV. Note that when comparing with Fig. 3, the lower $z$ edge is here increased to $z \sim 0.02$.

$W^+ + D^{*-}$ and $W^- + D^{*+}$, with a leptonically decaying $W$ boson. The predictions are compared to data from the ATLAS experiment at 13 TeV [10]. The cross sections measured in [10] are differential either in the transverse momentum $p_{T,h}$ of the $D^{(*)}$ hadron or in the absolute value of the pseudo-rapidity $|\eta_\ell|$ of the lepton from the $W$-boson decay.

In the following, we first provide details about the kinematical cuts and the numerical setup adopted, and we discuss our choice of $D^{(*)}$-meson fragmentation functions. We then divide the presentation in the rest of the section according to the observable considered.

### 5.1 Fiducial cuts and numerical setup

The fiducial region of the ATLAS measurement is defined as follows:

$$p_{T,h} > 8 \text{ GeV}, \quad |\eta_h| < 2.2, \quad p_{T,\ell} > 30 \text{ GeV}, \quad |\eta_\ell| < 2.5, \tag{5.1}$$

without any requirement on the presence of reconstructed jets. As in section 4, we provide predictions using the NNPDF3.1 PDF set [27] with values of $\alpha_s = 0.118$ and $n_f^{\max} = 5$. The central renormalisation, factorisation and fragmentation scales are chosen to be equal to the transverse mass of the $W$ boson:

$$\mu_R = \mu_F = \mu_D = m_T^W \,. \tag{5.2}$$

We estimate theoretical uncertainties with a 7-point scale variation using the same scheme as for the $Z$+hadron results in section 4, i.e. by maintaining $\mu_F = \mu_D$ while varying $\mu_R$ by separately halving or doubling their values, and by discarding pairs of extreme variations. The electroweak parameters are computed in the $G_\mu$-scheme with values

$$\begin{aligned} M_Z &= 91.1876 \text{ GeV}, \quad \Gamma_Z = 2.4952 \text{ GeV}\,, \\ M_W &= 80.379 \text{ GeV}, \quad \Gamma_W = 2.085 \text{ GeV}, \quad G_\mu = 1.1663787 \cdot 10^{-5} \text{ GeV}^{-2}\,. \end{aligned} \tag{5.3}$$

Furthermore, we use a non-diagonal CKM matrix with Wolfenstein parametrisation given by $\lambda = 0.2265$, $A = 0.79$, $\bar{\rho} = 0.141$ and $\bar{\eta} = 0.357$ [35].

In our predictions, we compare the results obtained using two $D^{(*)}$-meson fragmentation function sets, CNO [36] and KKKS08 [37]. We first provide some details concerning the individual sets, given that they feature quite substantial differences in the methodology adopted for their determination.

The CNO set is derived by exploiting the perturbative fragmentation function formalism for heavy quarks [38], in order to resum large logarithms of a hard scale over the charm quark mass. Namely, the NLO perturbative initial condition at the charm-mass scale—further supplemented with next-to-leading-logarithmic soft gluon resummation—is evolved with a NLO DGLAP evolution to the higher scales. A three-parameter non-perturbative component (Colangelo–Nason form supplemented with a hard term) is added on top of the perturbative result, with values for the parameters obtained through a fit to CLEO and BELLE data. See [36] for additional details.

Instead, the KKKS08 set is obtained through a global fit of BELLE, CLEO, ALEPH and OPAL data in a way similar to light-hadron FFs fits, without any perturbative input. Namely, both the charm and bottom FFs are parametrised with a Bowler-like form with three parameters at their respective mass scales, and then they are evolved to higher scales with a NLO DGLAP evolution. Two variants of the KKKS08 are presented in [37]: a zero-mass (ZM) variant, where all quark masses are neglected, and a general-mass (GM) approach, which includes $b$- and $c$-quark finite-mass corrections. In the following, we will present predictions for the GM variant only, in order to have two $D^{(*)}$-meson FF sets as different as possible.

Lastly, in our predictions we must take into account the *opposite sign minus same sign* (OS−SS) prescription applied in the measurement. Such a prescription consists in subtracting the contributions where the $W$ boson and the $D$ hadron have the same sign (SS) to the contributions where they have opposite sign (OS). In order to consistently apply such a prescription in our theoretical predictions, we first compute observables for all the sign combinations of $W$ and $D^{(*)}$ and then we perform the OS−SS subtraction on the resulting distributions and fiducial cross sections. For example

$$\sigma^{\text{OS}-\text{SS}}_{W^-+D^+} \equiv \sigma_{W^-+D^+} - \sigma_{W^-+D^-} \tag{5.4}$$

is the OS−SS fiducial cross section for $W^- + D^+$ production. Note that due the fact that the charm-to-$D^-$ fragmentation function is small but non-zero, the SS piece already contributes at LO. This is in contrast to predictions for $W + c$-jet production [39], where SS contributions only contribute starting from NLO.

### 5.2 Results for the $|\eta_\ell|$ observable

We first consider predictions for differential distributions in $|\eta_\ell|$. Results for $W$-boson plus $D$-hadron and $W$-boson plus $D^*$-hadron production are shown in Fig. 5 and Fig. 6, respectively. In each figure, the OS−SS results for $W^- + D^+$ ($W^- + D^{*+}$) and $W^+ + D^-$ ($W^+ + D^{*-}$) are given on the left and on the right column respectively. The top row

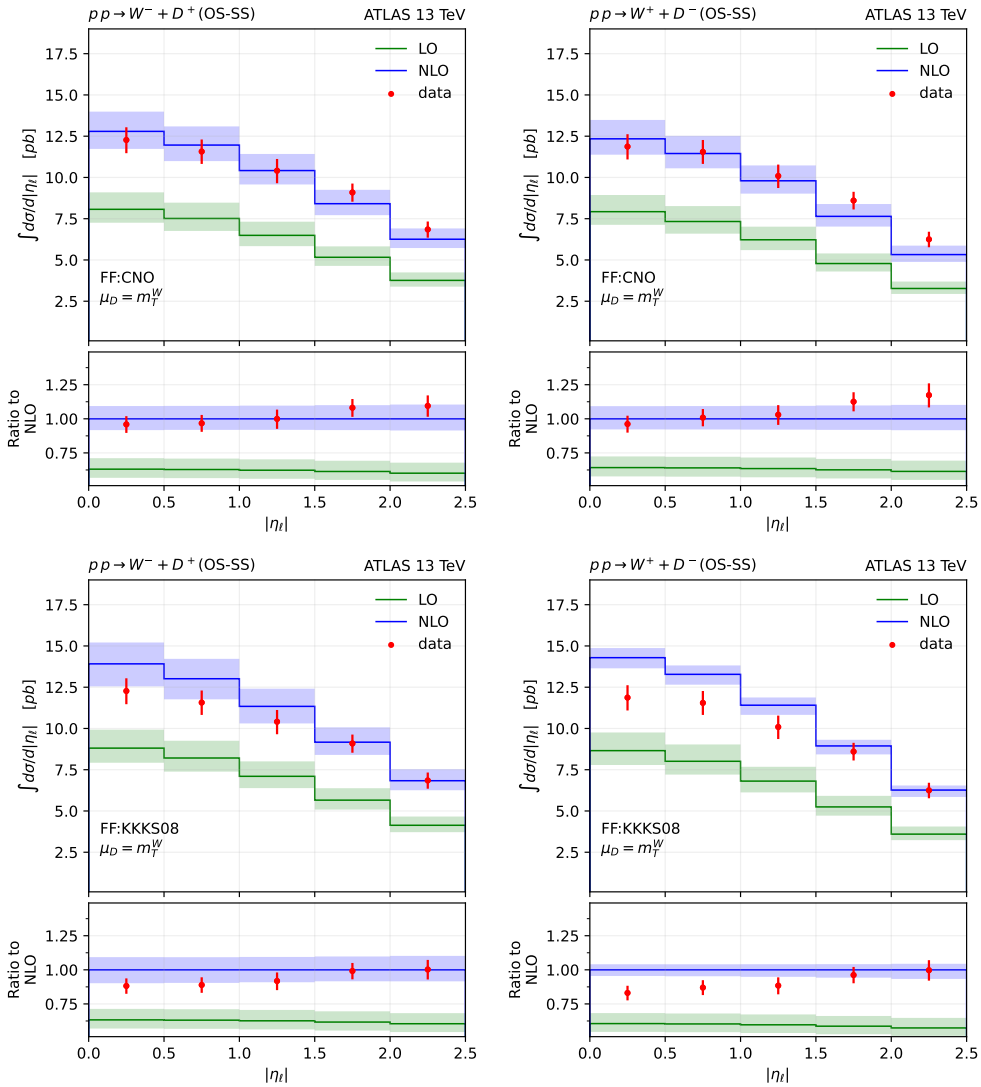


**Figure 5**. Comparison of LO (green) and NLO (blue) predictions with data (red). The $|\eta_\ell|$ distribution is integrated over each differential bin and is shown for all the $D$-hadron cases. The top plots have been produced using the CNO fragmentation function, while the bottom plots using KKKS08. The left and right columns show the results for $W^-$ and $W^+$, respectively. All include the OS−SS prescription.

contains plots with predictions using the CNO set, whereas the bottom row features plots adopting the KKKS08 set. Each plot is divided into two panels, showing the distributions in absolute value (top panel) and their ratio to NLO (bottom panel).

We first note that the NLO correction is about 40% and quite flat across the entire $|\eta_\ell|$ spectrum. The results obtained with the CNO fragmentation functions describe the data better in the central region, while the distributions obtained with KKKS08 instead show better agreement in the forward region. On the other hand, in the $D^*$-hadron case presented in Fig. 6, we observe that the $|\eta_\ell|$ distributions are fairly similar for the two sets of FF.

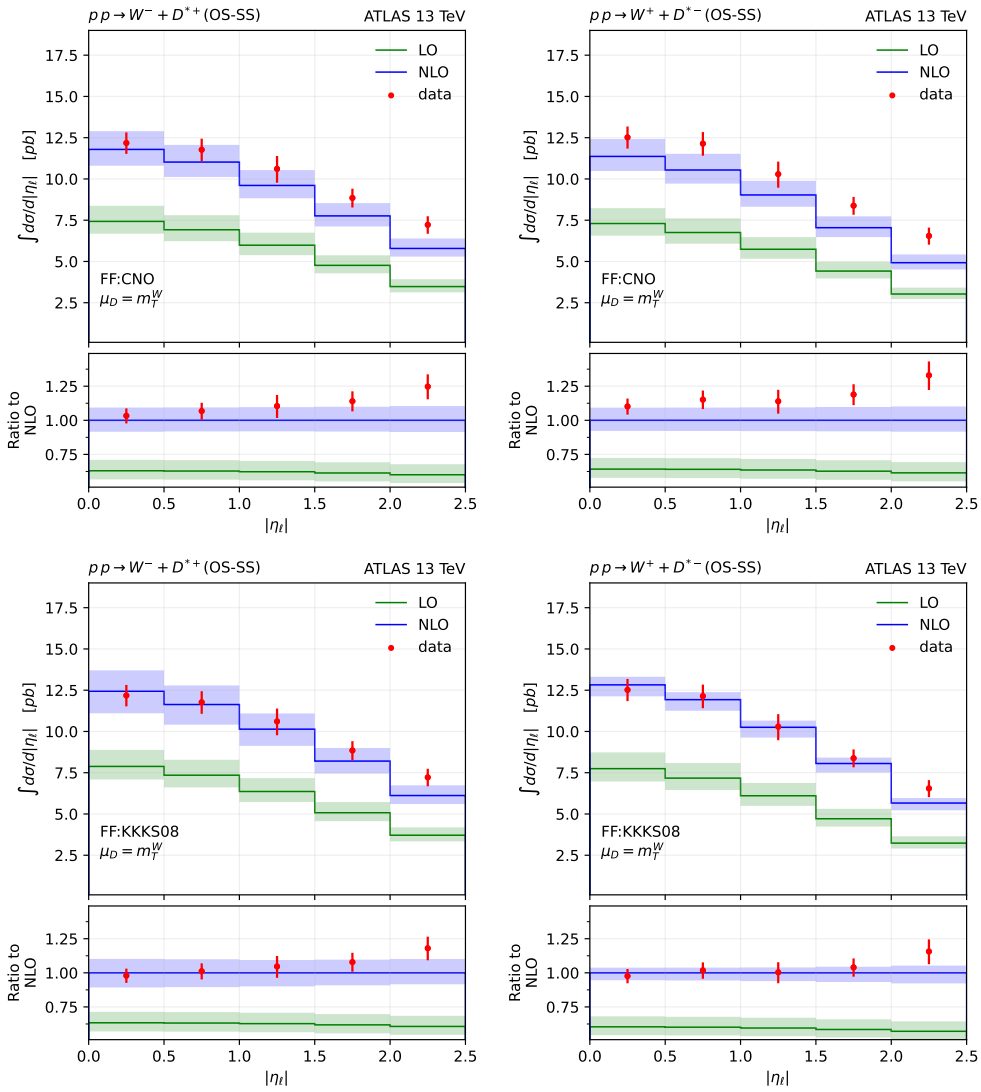


**Figure 6**. Comparison of LO (green) and NLO (blue) predictions with data (red). The $|\eta_\ell|$ distribution is integrated over each differential bin and is shown for all the $D^*$-hadron cases. The top plots have been produced using the CNO fragmentation function, while the bottom plots using KKKS08. The left and right columns show the results for $W^-$ and $W^+$, respectively. All include the OS−SS prescription.

Also in Fig. 6 NLO corrections are very flat in $|\eta_\ell|$ and amount to about 40%. However, the $W^- + D^{*+}$ prediction is in better agreement with the data with respect to the $W^+ + D^{*-}$ process, especially for the CNO fragmentation function. Finally, the last rapidity bin in the forward region is poorly described in both processes by both FFs. Overall, we can see a clear improvement in the quality of description of data moving from LO to NLO.

A direct comparison between the predictions obtained with the CNO and the KKKS08 fragmentation functions is given in Fig. 7 for the $|\eta_\ell|$ distribution. The plot is composed by four panels, each one showing the ratio of the corresponding NLO distribution to data, for

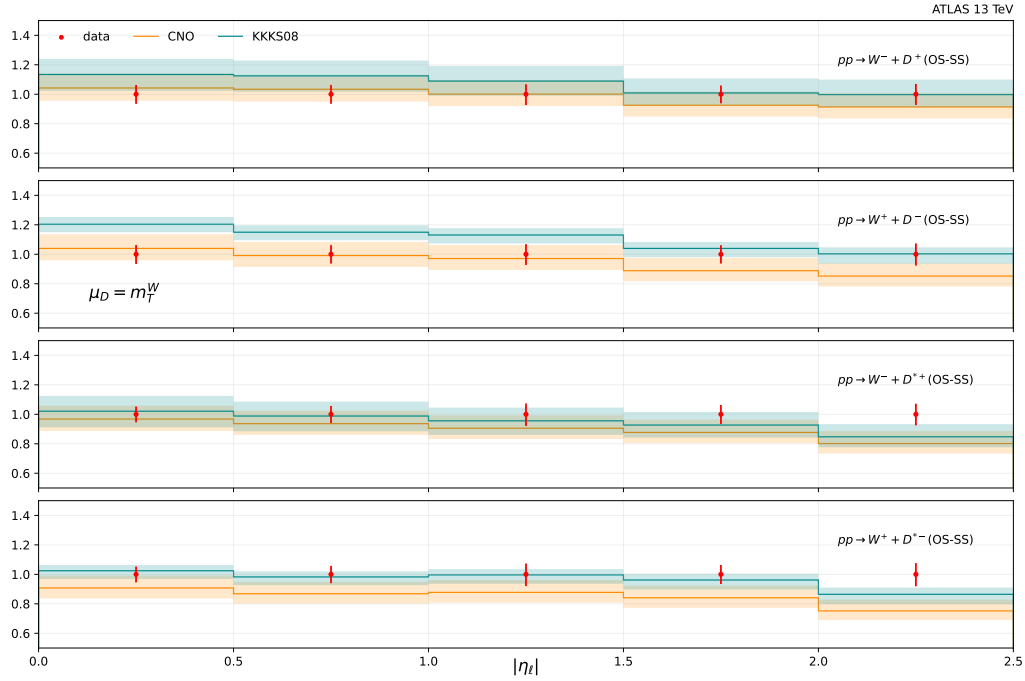


**Figure 7**. Comparison of the NLO results obtained for the different sets of fragmentation functions shown in Fig. 5 and Fig. 6. The plots correspond to the ratio to data of the $|\eta_\ell|$ distribution for $D$- and $D^*$-hadrons.

a different process. From Fig. 7 we can better assess the impact of changing the fragmentation function between CNO and KKKS08 on the $|\eta_\ell|$ distribution which is fairly small, as expected. As we already noticed from the $D$-hadron plots in Fig. 5, the KKKS08 distribution slightly overshoots the data in the central region. We also notice that both the distributions are below the data in the forward region for the $D^*$-hadron case, but KKKS08 is closer to data.

### 5.3 Results for the $p_{T,h}$ observable

In addition to the lepton rapidity $|\eta_\ell|$, we also consider predictions differential in $p_{T,h}$ both for $W$-boson plus $D$-hadron and $W$-boson plus $D^*$-hadron production. They are shown in Fig. 8 and Fig. 9, respectively, with the same format as in Figs. 5 and 6.

Overall, we note that also the $p_{T,h}$ distributions are in good agreement with the data. In the $D$-hadron case shown in Fig. 8, the NLO correction is about a 40% in the whole $p_{T,h}$ region. It is interesting to notice that the results obtained with the CNO fragmentation function are compatible with the data in the full spectrum, while we observe disagreement in the high-$p_{T,h}$ region for the predictions with KKKS08. The $p_{T,h}$ distribution for the $D^*$-hadron case is shown in Fig. 9 and similarly to the $D$-hadron case features a NLO correction of 40%. Here, predictions with both FF sets undershoot the data points in the low-$p_{T,h}$ region, however remaining generally compatible within uncertainties. We again observe some disagreement in the high-$p_{T,h}$ region when the KKKS08 set is adopted.

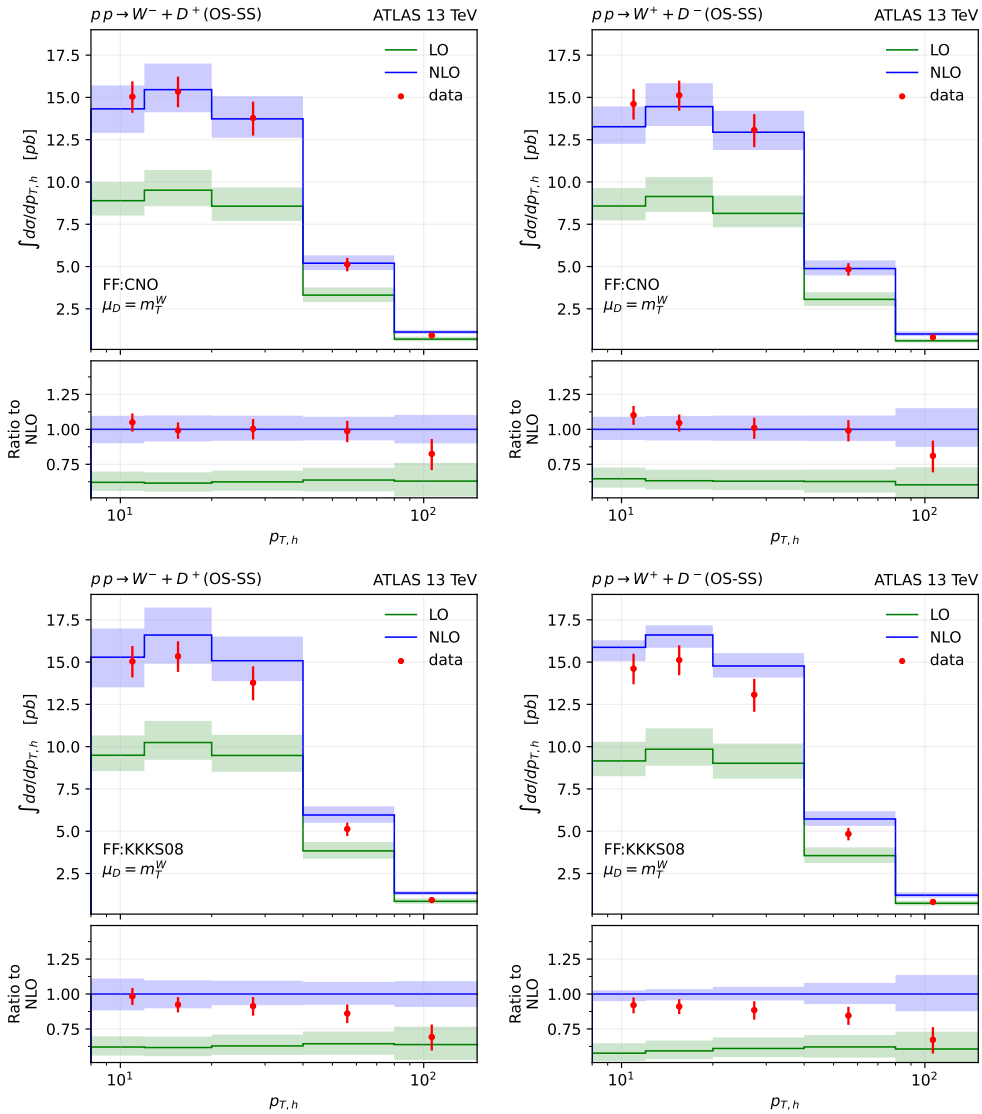


**Figure 8**. Comparison of LO (green) and NLO (blue) predictions with data (red). The $p_{T,h}$ distribution is integrated over each differential bin and is shown for all the $D$-hadron cases. The top plots have been produced using the CNO fragmentation function, while the bottom plots using KKKS08. The left and right columns show the results for $W^-$ and $W^+$, respectively. All include the OS−SS prescription.

A direct comparison of the $p_{T,h}$ distribution obtained with the two sets of fragmentation functions is shown in Fig. 10. From this plots it is easier to notice that the $p_{T,h}$ distribution is better described by the prediction obtained using the CNO fragmentation function in the low $p_{T,h}$ region, while the results obtained with KKKS08 are closer to data in the high-$p_{T,h}$ region, especially for the $D^*$-hadron cases. However, it is interesting to notice that the difference between the two seems to be just a shift that does not affect the shape of the distributions.

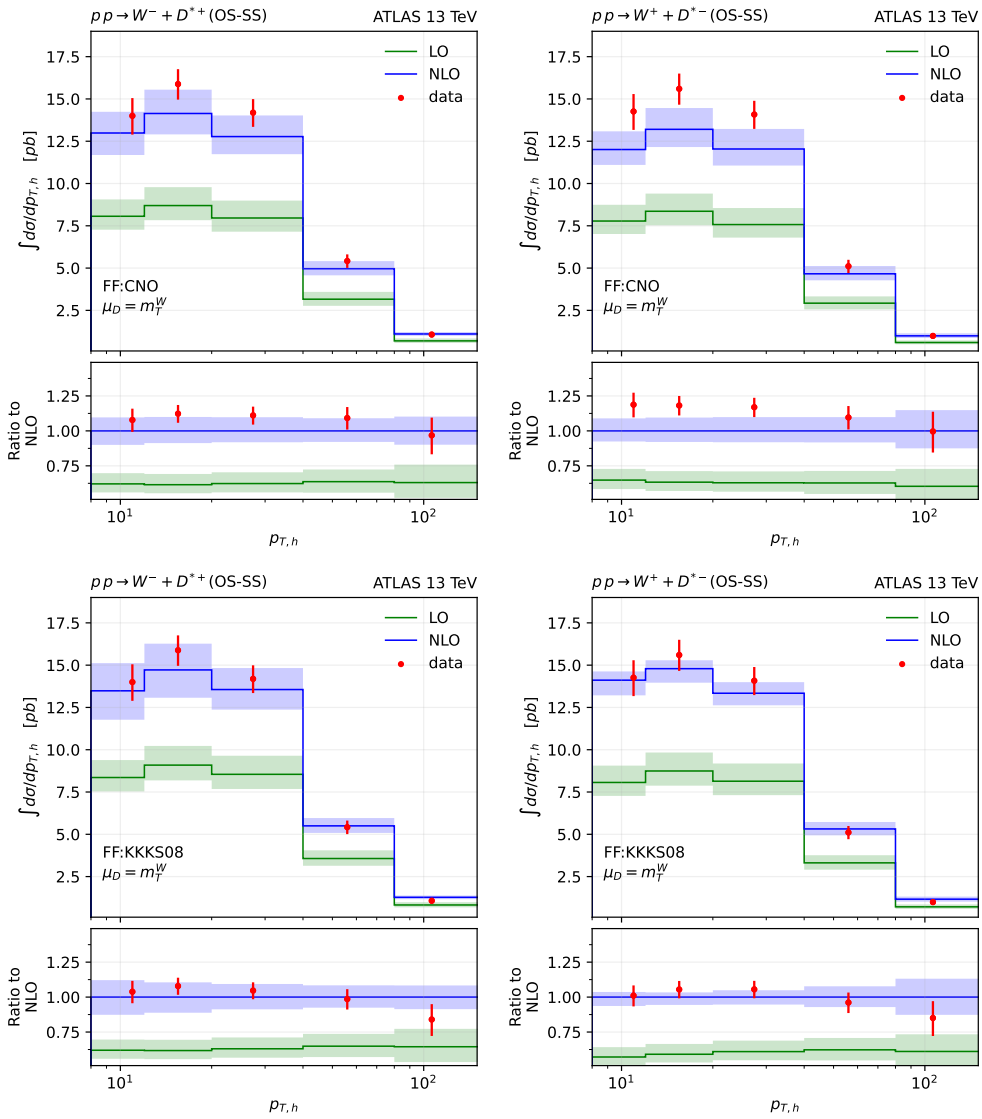


**Figure 9**. Comparison of LO (green) and NLO (blue) predictions with data (red). The $p_{T,h}$ distribution is integrated over each differential bin and is shown for all the $D^*$-hadron cases. The top plots have been produced using the CNO fragmentation function, while the bottom plots using KKKS08. The left and right columns show the results for $W^-$ and $W^+$, respectively. All include the OS−SS prescription.

## 6 Conclusions

In this paper, we have provided theoretical predictions for observables related to identified hadron production at LHC. In the first part, we have described how the antenna formalism has been extended to deal with infrared singularities associated to identified parton production in the short-distance cross section. We have detailed the analytical ingredients necessary for the computation of observables related to the production of a hadron in association with an electroweak boson decaying leptonically at NLO accuracy. In the second

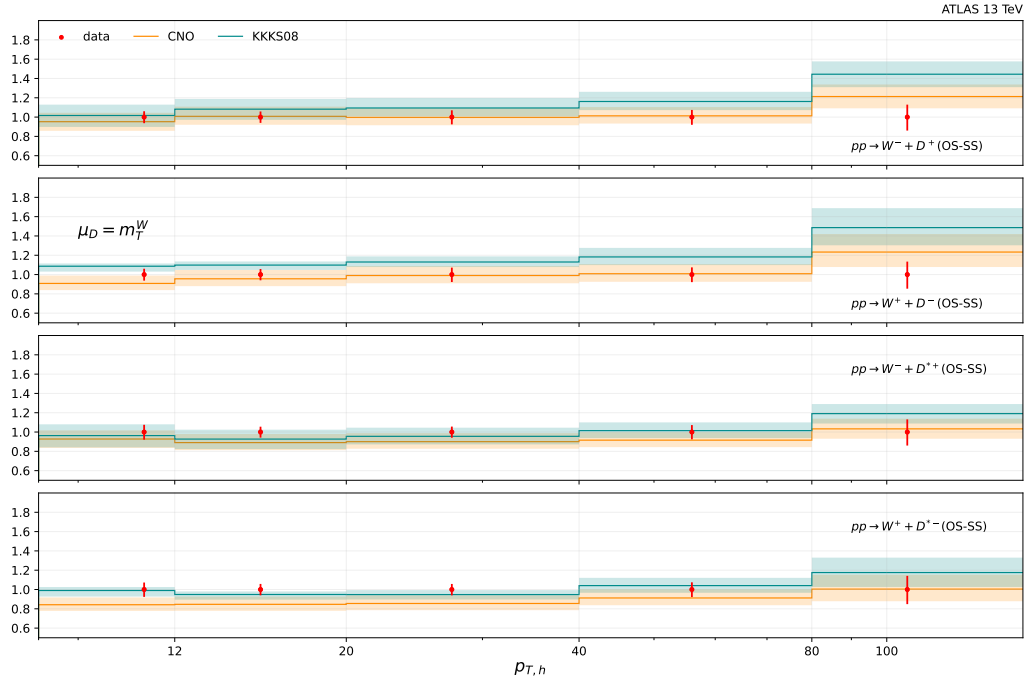


**Figure 10**. Comparison of the NLO results obtained for the different sets of fragmentation functions shown in Fig. 8 and Fig. 9 The plots corresponds to the ratio to data of the $p_{T,h}$ distribution for $D$ and $D^*$-hadrons.

part, we have performed a detailed comparison between our predictions and experimental data for both the production of a $Z$ boson in association with charged hadrons inside a jet, as measured by the LHCb collaboration [11], as well as the production of a $W$ boson in association with charmed hadrons, as measured by the ATLAS collaboration in [10]. In our predictions, we have considered various choices of light- or heavy-quark fragmentation functions.

In observables related to the production of a hadron inside a jet in association with a $Z$-boson, we have found that our results depend significantly on the choice of the fragmentation function, highlighting the importance of including this data in FF fits. The agreement of theory predictions with experimental data is also largely determined by the kinematical region considered, as for low values of the longitudinal momentum fraction of the hadron inside the jet we cannot expect perturbative fixed-order calculations to provide a fair description of data. It would be interesting to match our NLO calculations with resummation, possibly including non-perturbative modelling, in order to enlarge the range of validity of our theoretical predictions.

Predictions for $W$-boson production in association with a charmed hadron have instead been found to be less dependent on the heavy quark fragmentation function adopted. We have shown that two rather different FF sets (CNO and KKKS08) give similar sizeable NLO corrections, which bring theory predictions in good agreement with data, for both the lepton rapidity $|\eta_\ell|$ and the hadron transverse momentum $p_{T,h}$ distributions. Both CNO and KKKS08 sets contain parameters that have been fitted to $e^+e^-$ data and it is reassuring

to observe that these FF sets are also able to provide a reasonable description of LHC data.

Of course, it would be desirable to push our theory predictions to NNLO and the work presented in this paper is a first step towards this direction. Consistent NNLO fixed-order predictions would also require the availability of NNLO heavy quark fragmentation functions. Within the perturbative fragmentation function framework, choices related to the delicate treatment of non-perturbative effects in association with soft gluon resummation have been shown to give rather different results [40]. This behaviour is also enhanced by the small value of the charm mass, which is barely above the non-perturbative energy scale. Hence, pushing the accuracy of predictions for identified hadron production at the LHC to NNLO will require not only technical work within the antenna subtraction formalism but also improvements on the conceptual side to better understand the heavy quark fragmentation dynamics.

## Acknowledgments

We would like to thank Luca Rottoli for useful discussions on the hadron fragmentation functions belonging to the sets NNFF1.0 and NNFF1.1. We are grateful to Leonardo Bonino for a careful reading of the manuscript. This research was supported by the Swiss National Science Foundation (SNF) under contract 200021-197130. Numerical simulations were facilitated by the HighPerformance Computing group at ETH Zürich and the Swiss National Supercomputing Centre (CSCS) under project ID ETH5f.

## References

[1] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[2] F. Aversa, P. Chiappetta, M. Greco and J.P. Guillet, *QCD Corrections to Parton-Parton Scattering Processes*, *Nucl. Phys. B* **327** (1989) 105.

[3] P.J. Rijken and W.L. van Neerven, *Higher order QCD corrections to the transverse and longitudinal fragmentation functions in electron - positron annihilation*, *Nucl. Phys. B* **487** (1997) 233 [hep-ph/9609377].

[4] A. Mitov, S. Moch and A. Vogt, *Next-to-Next-to-Leading Order Evolution of Non-Singlet Fragmentation Functions*, *Phys. Lett. B* **638** (2006) 61 [hep-ph/0604053].

[5] L. Bonino, T. Gehrmann and G. Stagnitto, *Semi-inclusive deep-inelastic scattering at NNLO in QCD*, 2401.16281.

[6] ATLAS collaboration, *Measurement of the production of a $W$ boson in association with a charm quark in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **05** (2014) 068 [1402.6263].

[7] CMS collaboration, *Measurement of associated production of a $W$ boson and a charm quark in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 269 [1811.10021].

[8] CMS collaboration, *Measurements of the associated production of a $W$ boson and a charm quark in proton–proton collisions at $\sqrt{s} = 8$ TeV* , *Eur. Phys. J. C* **82** (2022) 1094 [2112.00895].

[9] CMS collaboration, *Measurement of the production cross section for a W boson in association with a charm quark in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **84** (2024) 27 [2308.02285].

[10] ATLAS collaboration, *Measurement of the production of a W boson in association with a charmed hadron in pp collisions at* $\sqrt{s} = 13$ TeV *with the ATLAS detector*, *Phys. Rev. D* **108** (2023) 032012 [2302.00336].

[11] LHCb collaboration, *Multidifferential study of identified charged hadron distributions in Z-tagged jets in proton-proton collisions at* $\sqrt{s}$ *=13 TeV*, *Phys. Rev. D* **108** (2023) L031103 [2208.11691].

[12] S. Catani and M.H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323].

[13] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [1804.10017].

[14] M.L. Czakon, T. Generet, A. Mitov and R. Poncelet, *B-hadron production in NNLO QCD: application to LHC* $t\bar{t}$ *events with leptonic decays*, *JHEP* **10** (2021) 216 [2102.08267].

[15] T. Gehrmann and R. Schürmann, *Photon fragmentation in the antenna subtraction formalism*, *JHEP* **04** (2022) 031 [2201.06982].

[16] X. Chen, T. Gehrmann, E.W.N. Glover, M. Höfer, A. Huss and R. Schürmann, *Single photon production at hadron colliders at NNLO QCD with realistic photon isolation*, *JHEP* **08** (2022) 094 [2205.01516].

[17] T. Gehrmann and G. Stagnitto, *Antenna subtraction at NNLO with identified hadrons*, *JHEP* **10** (2022) 136 [2208.02650].

[18] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[19] C. Liu, X. Shen, B. Zhou and J. Gao, *Automated calculation of jet fragmentation at NLO in QCD*, *JHEP* **09** (2023) 108 [2305.14620].

[20] G. Bevilacqua, M.V. Garzelli, A. Kardos and L. Toth, *W + charm production with massive c quarks in PowHel*, *JHEP* **04** (2022) 056 [2106.11261].

[21] S. Ferrario Ravasio and C. Oleari, *NLO + parton-shower generator for Wc production in the POWHEG BOX RES*, *Eur. Phys. J. C* **83** (2023) 684 [2304.13791].

[22] J. Currie, E.W.N. Glover and S. Wells, *Infrared Structure at NNLO Using Antenna Subtraction*, *JHEP* **04** (2013) 066 [1301.4693].

[23] A. Daleo, T. Gehrmann and D. Maitre, *Antenna subtraction with hadronic initial states*, *JHEP* **04** (2007) 016 [hep-ph/0612257].

[24] S. Catani, *The Singular behavior of QCD amplitudes at two loop order*, *Phys. Lett. B* **427** (1998) 161 [hep-ph/9802439].

[25] LHCb collaboration, *Study of Z Bosons Produced in Association with Charm in the Forward Region*, *Phys. Rev. Lett.* **128** (2022) 082001 [2109.08084].

[26] M. Cacciari, G.P. Salam and G. Soyez, *The anti-*$k_t$ *jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189].

[27] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [1706.00428].

[28] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [1412.7420].

[29] J. Binnewies, B.A. Kniehl and G. Kramer, *Pion and kaon production in e+ e- and e p collisions at next-to-leading order*, *Phys. Rev. D* **52** (1995) 4947 [hep-ph/9503464].

[30] NNPDF collaboration, *Charged hadron fragmentation functions from collider data*, *Eur. Phys. J. C* **78** (2018) 651 [1807.03310].

[31] D. de Florian, R. Sassot and M. Stratmann, *Global analysis of fragmentation functions for pions and kaons and their uncertainties*, *Phys. Rev. D* **75** (2007) 114010 [hep-ph/0703242].

[32] D. de Florian, R. Sassot and M. Stratmann, *Global analysis of fragmentation functions for protons and charged hadrons*, *Phys. Rev. D* **76** (2007) 074033 [0707.1506].

[33] NNPDF collaboration, *A determination of the fragmentation functions of pions, kaons, and protons with faithful uncertainties*, *Eur. Phys. J. C* **77** (2017) 516 [1706.07049].

[34] J. Gao, C. Liu, X. Shen, H. Xing and Y. Zhao, *Simultaneous Determination of Fragmentation Functions and Test on Momentum Sum Rule*, 2401.02781.

[35] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[36] M. Cacciari, P. Nason and C. Oleari, *A Study of heavy flavored meson fragmentation functions in e+ e- annihilation*, *JHEP* **04** (2006) 006 [hep-ph/0510032].

[37] T. Kneesch, B.A. Kniehl, G. Kramer and I. Schienbein, *Charmed-meson fragmentation functions with finite-mass corrections*, *Nucl. Phys. B* **799** (2008) 34 [0712.0481].

[38] B. Mele and P. Nason, *Next-to-leading QCD calculation of the heavy quark fragmentation function*, *Phys. Lett. B* **245** (1990) 635.

[39] A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, A. Huss, A.R. Garcia and G. Stagnitto, *Precise QCD predictions for W-boson production in association with a charm jet*, *Eur. Phys. J. C* **84** (2024) 361 [2311.14991].

[40] L. Bonino, M. Cacciari and G. Stagnitto, *Heavy Quark Fragmentation in $e^+e^-$ Collisions to NNLO+NNLL Accuracy in Perturbative QCD*, 2312.12519.